

POLITECHNIKA ŚLĄSKA W GLIWICACH
WYDZIAŁ INŻYNIERII BIOMEDYCZNEJ

Multimodalny system śledzenia i rejestracji
w zabiegach biopsji gruboigłowej
gruczołu sutkowego

— ROZPRAWA DOKTORSKA —

AUTOR

mgr inż. Agata Maria Wijata

PROMOTOR

prof. dr hab. inż. Dominik Spinczyk

PROMOTOR POMOCNICZY

dr inż. lek. Bartłomiej Pyciński

Zabrze 2022

Składam wyrazy podziękowania Panom Promotorom niniejszej pracy: prof. dr. hab. inż. Dominikowi Spinczykowi za poświęcony czas oraz dr. inż. lek. Bartłomiejowi Pycińskiemu za opiekę naukową, zaangażowanie i wsparcie.

Wyrazy wdzięczności kieruję do Pani prof. dr. hab. inż. Ewy Piętki oraz Koleżanek i Kolegów z Katedry Informatyki Medycznej i Sztucznej Inteligencji za okazaną życzliwość oraz pomoc w naukowych i życiowych wyzwaniach.

Dziękuję mojemu Mężowi za spokój i cierpliwość, Rodzicom za opiekę i zaufanie.

Serdecznie dziękuję za uśmiech, kibicowanie i nieustające wsparcie mojej Rodzinie i Przyjaciółom.

Szczególne podziękowania składam Panu dr. hab. inż. Jakubowi Nalepie za pokazanie, że można inaczej.

Spis treści

1.	<i>Wprowadzenie</i>	1
1.1	Nowotwór złośliwy gruczołu sutkowego	1
1.2	Biopsja gruboigłowa gruczołu sutkowego	3
1.3	Ultrasonografia	4
1.3.1	Powstawanie obrazu USG	5
1.3.2	Zastosowanie obrazowania USG	5
1.3.3	Typy obrazowania USG	6
1.4	Systemy śledzenia	6
1.5	Przedstawienie problemu, cel i zakres prac	7
1.6	Układ rozprawy	9
2.	<i>Istniejące rozwiązania</i>	11
2.1	Diagnostyka radiologiczna gruczołu sutkowego	11
2.1.1	Mammografia	12
2.1.2	Ultrasonografia	12
2.2	Segmentacja i klasyfikacja zmian ogniskowych w obrazie USG	14
2.2.1	Segmentacja zmian ogniskowych	14
2.2.2	Klasyfikacja złośliwości zmian ogniskowych	20
2.3	Detekcja i segmentacja igły w obrazie USG	22
2.3.1	Detekcja igły	22
2.3.2	Segmentacja igły	26
2.4	Systemy śledzenia	29
2.4.1	Optyczny system śledzenia Polaris	33
2.4.2	Elektromagnetyczny system śledzenia	34
2.5	Podsumowanie	35
3.	<i>System rejestracji danych</i>	37
3.1	Konstrukcja stanowiska pomiarowego	37
3.2	Stanowisko pomiarowe – integracja akwizycji danych	41
3.3	Zintegrowany system śledzenia narzędzi i obrazu USG	43
3.3.1	Działanie systemu śledzenia	44
3.3.2	Kalibracja wskaźnika w systemach śledzenia	46

3.3.3	Kalibracja pistoletu biopsyjnego	47
3.3.4	Kalibracja czasowa	50
3.3.5	Kalibracja przestrzenna obrazu ultrasonograficznego	51
3.3.6	Kalibracja układów współrzędnych systemów śledzenia położenia	55
3.4	Podsumowanie	56
4.	<i>Klasyfikacja i segmentacja struktur w obrazach USG</i>	57
4.1	Przetwarzanie wstępne	57
4.2	Konwolucyjne sieci neuronowe	59
4.3	Elementy architektury sieci CNN	62
4.3.1	Warstwa wejściowa	62
4.3.2	Warstwa konwolucyjna	63
4.3.3	Warstwa głosująca	64
4.3.4	Warstwa transponowanej konwolucji	66
4.3.5	Warstwa w pełni połączona	67
4.3.6	Warstwa klasyfikująca	67
4.3.7	Warstwa klasyfikacji pikseli	68
4.3.8	Funkcja aktywacji typu ReLU	69
4.3.9	Funkcja aktywacji typu <i>softmax</i>	70
4.3.10	Mechanizm normalizacji	70
4.3.11	Regularyzator typu <i>dropout</i>	71
4.4	Sieć CNN klasyfikująca obecność igły	71
4.4.1	Architektura sieci klasyfikującej obecność igły	71
4.4.2	Zbiory treningowe, walidacyjne i testowe	73
4.4.3	Augmentacja danych	74
4.4.4	Proces treningu sieci klasyfikującej	76
4.5	Sieć CNN klasyfikująca zmianę ogniskową	79
4.5.1	Architektura sieci klasyfikującej obecność i typ zmiany ogniskowej	80
4.5.2	Zbiory treningowe, walidacyjne i testowe	80
4.5.3	Augmentacja danych	81
4.5.4	Proces treningu sieci klasyfikującej	82
4.6	Sieć CNN do segmentacji semantycznej regionu igły biopsyjnej	82
4.6.1	Architektura sieci do segmentacji semantycznej	83
4.6.2	Zbiory treningowe, walidacyjne i testowe	85
4.6.3	Augmentacja danych	86
4.6.4	Proces treningu sieci segmentującej	86
4.6.5	Wyznaczanie trajektorii igły za pomocą transformaty Ra- dona	87
4.7	Komitety klasyfikujące do segmentacji zmiany ogniskowej	87
4.7.1	Filtracja dwuwymiarowych obrazów ultrasonograficznych	88

4.7.2	Metody oceny jakości filtracji obrazów	91
4.7.3	Dobór metody filtracji i jej parametrów	95
4.7.4	Segmentacja zmiany ogniskowej metodami klasycznymi	96
4.7.5	Komitety klasyfikujące	99
4.8	Ewaluacja metod klasyfikacji i segmentacji obrazów USG . . .	100
4.8.1	Miary oceny klasyfikacji	102
4.8.2	Miary oceny segmentacji	106
4.8.3	Miary dokładności detekcji igły	107
4.9	Podsumowanie	110
5.	<i>Eksperymenty i wyniki</i>	113
5.1	System rejestracji danych	113
5.1.1	Pomiary dokładności systemów śledzenia	113
5.1.2	Kalibracja wskaźnika	114
5.1.3	Kalibracja pistoletu biopsyjnego	115
5.1.4	Dokładność rejestracji fantomu	116
5.1.5	Kalibracja głowicy ultrasonograficznej	117
5.2	Dane źródłowe	119
5.2.1	Obrazy USG z zapisem biopsji gruboigłowej	119
5.2.2	Obrazy USG ze zmianami ogniskowymi	123
5.3	Klasyfikacja i segmentacja struktur w obrazie USG	126
5.3.1	Klasyfikacja obrazów USG z igłą biopsyjną	126
5.3.2	Klasyfikacja obrazów USG ze zmianą ogniskową	142
5.3.3	Detekcja igły biopsyjnej w obrazach USG	146
5.3.4	Segmentacja zmiany ogniskowej w obrazach USG	156
5.4	Podsumowanie	164
6.	<i>Podsumowanie</i>	167
6.1	Praktyczne zastosowanie systemu	168
6.2	Plany dalszych prac	168
A.	<i>Dodatek – tabele z wynikami</i>	171
	<i>Bibliografia</i>	202

Spis rysunków

1.1	Statystyka zachorowalności oraz śmiertelności z powodu nowotworu gruczołu sutkowego w Polsce w latach 1999-2019	2
1.2	Schemat zabiegu biopsji gruczołu sutkowego pod kontrolą ultrasonograficzną	4
2.1	Liczba publikacji poświęcona tematyce raka piersi. Stan na wrzesień 2022	11
2.2	Liczba publikacji z zakresu technik obrazowania zmian ogniskowych gruczołu sutkowego, na podstawie wyszukiwania hasłami <i>breast cancer imaging</i> i <i>breast cancer ultrasound</i> . Stan na wrzesień 202	12
2.3	Kamera stereoskopowa Polaris Vicra	33
2.4	Pole widzenia kamery stereoskopowej optycznego systemu śledzenia Polaris Vicra firmy NDI	34
2.5	System elektromagnetyczny Aurora firmy NDI	35
2.6	Kształt obszaru śledzenia systemu Aurora	36
2.7	Rzut boczny dwóch typów obszarów śledzenia systemu Aurora	36
3.1	Stanowisko pomiarowe	38
3.2	Układy współrzędnych obiektów i transformacje pomiędzy obiektami a znacznikami	39
3.3	Elementy pliku konfiguracyjnego aplikacji PlusServer	43
3.4	Plik konfiguracyjny aplikacji PlusServer	44
3.5	Etapy kalibracji stanowiska pomiarowego	45
3.6	Porównanie znaczników systemów optycznego i elektromagnetycznego	45
3.7	Znaczniki optycznego systemu Polaris	48
3.8	Znaczniki optycznego systemu Aurora	49
3.9	Pistolet Magnum z igłą przeznaczony do zabiegu biopsji gruczołowej	49
3.10	Markery umieszczone na śledzonych narzędziach	52
3.11	Fantom żyłkowy	53
3.12	Przestrzenna kalibracja obrazu USG	53

3.13	Przekrój poprzeczny przez żyłki fantomu	54
3.14	Rejestracja tożsamyh punktów w dwóch układach współrzędnych	55
3.15	Etapy kalibracji stanowiska pomiarowego	56
4.1	Schemat proponowanego całościowego przetwarzania obrazów ultrasonograficznych	58
4.2	Przykładowy obraz ultrasonograficzny zmiany ogniskowej poddawany przetwarzaniu	59
4.3	Schematyczne przedstawienie elementów architektury konwulucyjnej sieci neuronowej	63
4.4	Przykład operacji splotu	64
4.5	Przykład operacji maxpooling	65
4.6	Przykład operacji averagepooling	65
4.7	Przykład operacji próbkowania w górę	66
4.8	Szacowanie prawdopodobieństwa przynależności do klasy w architekturze sieci neuronowej	67
4.9	Architektura sieci CNN klasyfikującej obecność igły	72
4.10	Techniki augmentacji przez filtrację	77
4.11	Schemat treningu klasyfikatora obecności igły	79
4.12	Przykłady typów obrazów ultrasonograficznych tkanek zdrowej i patologicznych	79
4.13	Architektura sieci CNN wieloklasowej klasyfikującej typ zmiany ogniskowej	81
4.14	Schemat proponowanej metody detekcji i lokalizacji igły biopsyjnej	83
4.15	Architektura konwulucyjnej sieci neuronowej do segmentacji obszaru igły	84
4.16	Schemat treningu metody segmentacji z wykorzystaniem sieci CNN	87
4.17	Przykłady filtracji obrazów ultrasonograficznych zmiany ogniskowej gruczołu sutkowego	92
4.18	Ocena metod filtracji obrazów ultrasonograficznych zmiany ogniskowej gruczołu sutkowego	96
4.19	Algorytm wododziałowy	97
4.20	Komitet klasyfikatorów klasycznych metod segmentacji zmiany ogniskowej	100
4.21	Komitet klasyfikatorów klasycznych metod segmentacji zmiany ogniskowej	101
4.22	Podsumowanie metod ewaluacji z zakresu przetwarzania obrazów ultrasonograficznych	102
4.23	Tablica pomyłek	103
4.24	Przykład konstrukcji krzywej ROC	105
4.25	Definicja klasyfikacji pikseli	106

4.26	Wyjaśnienie parametrów oceny wyznaczonej trajektorii igły . . .	109
5.1	Wykresy przebiegów znormalizowanych sygnałów położenia obrazu	117
5.2	Zmiana ogniskowa echoujemna i echodatnia	121
5.3	Obrys ekspercki igły	122
5.4	Podsumowanie wpływu zmiany wartości współczynnika tempa uczenia i liczby epok na efektywność autorskiego klasyfikatora obecności igły w obrazie	129
5.5	Podsumowanie wpływu zmiany wartości prawdopodobieństwa p regularyzatora typu <i>dropout</i> na efektywność autorskiego klasyfikatora obecności igły w obrazie	131
5.6	Podsumowanie wpływu ważenia klas podczas procesu treningu na efektywność autorskiego klasyfikatora obecności igły w obrazie	132
5.7	Podsumowanie wpływu ważenia klas podczas procesu treningu na efektywność autorskiego klasyfikatora obecności igły w obrazie	132
5.8	Wpływ wymiarów obrazu wejściowego i liczby epok na efektywność autorskiego klasyfikatora obecności igły w obrazie	134
5.9	Wpływ wymiarów obrazu wejściowego i liczby epok na efektywność autorskiego klasyfikatora obecności igły w obrazie, trenowanego metodą dostrajania	136
5.10	Wykresy pudełkowe efektywności autorskiej metody klasyfikacji obecności igły biopsyjnej w obrazie ultrasonograficznym wyznaczone dla modeli trenowanych z uwzględnieniem mechanizmu wcześniejszego zatrzymania oryginalną bazą obrazów	138
5.11	Wykresy pudełkowe efektywności autorskiej metody klasyfikacji obecności igły biopsyjnej w obrazie ultrasonograficznym wyznaczone dla modeli trenowanych z uwzględnieniem mechanizmu wcześniejszego zatrzymania oraz z wykorzystaniem techniki dostrajania	140
5.12	Rozkład czasów klasyfikacji obrazów USG o różnej wielkości za pomocą autorskiej metody klasyfikacji obecności igły biopsyjnej	141
5.13	Wykresy pudełkowe efektywności autorskiej metody klasyfikacji obecności igły biopsyjnej w obrazie ultrasonograficznym wyznaczone dla modeli trenowanych z uwzględnieniem mechanizmu wcześniejszego zatrzymania oraz z wykorzystaniem techniki dostrajania	143
5.14	Wykresy pudełkowe czasu klasyfikacji obrazu ultrasonograficznego za pomocą autorskiej metody klasyfikacji obecności igły biopsyjnej i metod referencyjnych	144

5.15	Wykresy pudełkowe efektywności autorskiej metody klasyfikacji obecności zmiany ogniskowej w obrazie ultrasonograficznym wyznaczone dla modeli trenowanych z uwzględnieniem mechanizmu wcześniejszego zatrzymania oryginalną bazą obrazów . . .	147
5.16	Wykresy pudełkowe czasu klasyfikacji obrazu ultrasonograficznego za pomocą autorskiej metody klasyfikacji obecności zmiany ogniskowej i metod referencyjnych	148
5.17	Najlepszy wynik segmentacji igły autorską metodą w ocenie za pomocą miary IoU	148
5.18	Wynik autorskiej metody segmentacji odpowiadający średnim wartościom $\Delta\lambda$ i $\Delta\theta$	149
5.19	Nieprawidłowy wynik segmentacji igły autorską metodą	149
5.20	Przykładowe wyniki jakościowe i ilościowe ($\Delta\lambda$, $\Delta\theta$) dla testowanych metod detekcji igły biopsyjnej.	153
5.21	Histogramy efektywności automatycznej detekcji trajektorii igły $\Delta\lambda$ i $\Delta\theta$ dla autorskiej metody i metod referencyjnych	157
5.22	Najlepszy wynik segmentacji zmiany ogniskowej echoujemnej autorską metodą	158
5.23	Najlepszy wynik segmentacji zmiany ogniskowej echododatniej autorską metodą	158
5.24	Średni wynik segmentacji zmiany ogniskowej echoujemnej . . .	163
5.25	Średni wynik segmentacji zmiany ogniskowej echododatniej . .	164

Spis tabel

2.1	Podsumowanie metod segmentacji zmian ogniskowych w obrazach ultrasonograficznych 2D	14
2.2	Podsumowanie baz danych zmian ogniskowych 2D oraz miar z otrzymanymi wartościami	15
2.3	Podsumowanie metod segmentacji zmian ogniskowych w obrazach ultrasonograficznych 3D	17
2.4	Podsumowanie baz danych zmian ogniskowych 3D oraz miar z otrzymanymi wartościami	18
2.5	Podsumowanie metod klasyfikacji zmian ogniskowych w obrazach ultrasonograficznych	23
2.6	Podsumowanie baz danych zmian ogniskowych oraz miar z otrzymanymi wartościami	24
2.7	Podsumowanie metod detekcji igły w obrazach ultrasonograficznych 2D	27
2.8	Podsumowanie metod detekcji igły w obrazach ultrasonograficznych 3D	28
2.9	Podsumowanie baz danych 2D oraz miar z otrzymanymi wartościami	30
2.10	Podsumowanie baz danych 3D oraz miar z otrzymanymi wartościami	31
3.1	Zestawienie systemów śledzenia	41
3.2	Współrzędne sfer wskaźnika w lokalnym układzie współrzędnych [mm	47
3.3	Współrzędne sfer markera umieszczonego na powierzchni pistoletu biopsyjnego	50
4.1	Wybrane pretrenowane sieci CNN i ich parametry	61
4.2	Testowane parametry różnych metod filtracji (pierwsza wartość to początek testowanego zakresu dla danego parametru, ostatnia koniec, natomiast środkowa to krok).	96

5.1	Średnie błędy lokalizacji znacznika w zależności od odległości od urządzenia śledzącego	114
5.2	Średnie błędy lokalizacji znacznika w brzeźnym obszarze pola widzenia urządzenia śledzącego	115
5.3	Średnie błędy kalibracji wskaźników stosowanych w systemach nawigacji	115
5.4	Średnie błędy kalibracji pistoletu biopsyjnego	116
5.5	Powtarzalność kalibracji (CR) głowicy ultrasonograficznej	119
5.6	Obrazy USG wykorzystane w zakresie oceny metod detekcji igły biopsyjne	120
5.7	Podsumowanie parametrów igły widocznej w dwuwymiarowych obrazach w połączonej bazie danych	123
5.8	Liczebności pięciu źródłowych zbiorów obrazów USG	124
5.9	Parametry geometryczne zmian ogniskowych echoujemnych i echodatnich. P – pole powierzchni zmiany ogniskowej, L_1, L_2 – długości osi głównych. R_θ – kierunek obiektu. P_{wzg} – procentowy obszar pola powierzchni zmiany ogniskowej względem całego obrazu.	126
5.10	Podsumowanie testowanych hiperparametrów autorskiego klasyfikatora	127
5.11	Podsumowanie statystyk dla wyników segmentacji regionu igły dla proponowanej metody (PM) oraz metod porównawczych: Lee i wsp. [89] i Hatt i wsp. [55].	150
5.12	Błędy detekcji trajektorii igły oraz średni czas analizy obrazu dla autorskiej metody (PM), metody Lee i Hatt	152
5.13	Podsumowanie statystyk błędów detekcji trajektorii igły na wyniku segmentacji autorską architekturą sieci (PM), metodzie Lee i wsp. i metodzie Hatt i wsp.	154
5.14	Podsumowanie statystyk parametrów detekcji trajektorii igły λ i θ dla metod detekcji	154
5.15	Podsumowanie statystyk dla wyników segmentacji regionu zmiany ogniskowej niezależnie od jej typu znanymi metodami klasycznej segmentacji obiektów w obrazach.	160
5.16	Podsumowanie statystyk dla wyników segmentacji regionu zmiany ogniskowej z wykorzystaniem komitetu klasyfikującego na podstawie klasycznych metod segmentacji.	162
A.1	Wartości p wyznaczone za pomocą testu Shapiro Wilka dla miar oceniających segmentację realizowaną za pomocą autorskiej metody (PM) oraz metod referencyjnych	171

- A.2 Wyniki testu Wilcoxona dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary ACC_{glob} obliczonej dla autorskiej metody segmentacji oraz metod porównawczych 171
- A.3 Wyniki testu Wilcoxona dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary ACC_{mean} obliczonej dla autorskiej metody segmentacji oraz metod porównawczych 172
- A.4 Wyniki testu Wilcoxona dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary IoU obliczonej dla autorskiej architektury, metod Lee i wsp. oraz Hatt i wsp. 172
- A.5 Wyniki testu Wilcoxona dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary IoU_w obliczonej dla autorskiej architektury, metod Lee i wsp. oraz Hatt i wsp. 172
- A.6 Wyniki testu Wilcoxona dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary F_1 wyznaczonej dla autorskiej architektury, metod Lee i wsp. oraz Hatt i wsp. 172
- A.7 Wyniki testu Shapiro Wilka (p -wartości) parametrów λ oraz θ oraz błędów trajektorii igły $\Delta\lambda$ i $\Delta\theta$ wyznaczone dla złotego standardu (GT), autorskiej metody detekcji igły (PM) oraz metod porównawczych 173
- A.8 Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla parametru λ określonego dla trajektorii igły wyznaczonej za pomocą złotego standardu (GT) oraz autorskiej metody detekcji igły oraz metod porównawczych 173
- A.9 Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla parametru θ określonego dla trajektorii igły wyznaczonej za pomocą złotego standardu (GT) oraz autorskiej metody detekcji igły oraz metod porównawczych 173
- A.10 Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla błędu dystansu detekcji trajektorii $\Delta\lambda$ określonego dla trajektorii igły wyznaczonej za pomocą autorskiej metody detekcji igły oraz metod porównawczych 173

A.11 Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla błędu kąta przy detekcji trajektorii $\Delta\theta_R$ określonego dla trajektorii igły wyznaczonej za pomocą autorskiej metody detekcji igły oraz metod porównawczych	174
A.12 Wartości p wyznaczone za pomocą testu Shapiro Wilka dla miar segmentacji zmian ogniskowych niezależnie od jej typu	174
A.13 Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla dokładności segmentacji zmian ogniskowych niezależnie od jej typu	174
A.14 Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla współczynnika wyników prawdziwie pozytywnych segmentacji zmian ogniskowych niezależnie od jej typu . .	175
A.15 Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla współczynnika wyników prawdziwie negatywnych segmentacji zmian ogniskowych niezależnie od jej typu . .	175
A.16 Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla miary IoU segmentacji zmian ogniskowych niezależnie od jej typu	175
A.17 Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla miary F_1 segmentacji zmian ogniskowych niezależnie od jej typu	176
A.18 Podsumowanie statystyk dla wyników segmentacji regionu odpowiednio zmiany ogniskowej echoujemnej i echodatniej klasycznymi metodami segmentacji	177
A.19 Podsumowanie statystyk dla wyników segmentacji regionu zmiany ogniskowej echoujemnej i echodatniej metodą komitetu klasyfikującego na podstawie wyników klasycznych metod segmentacji	178

Spis skrótów i oznaczeń

AC	<i>active contour</i> – metoda aktywnego konturu
ACC	<i>accuracy</i> – dokładność
AF	<i>anisotropic diffusion</i> – dyfuzja anizotropowa
ARE	<i>averaged radial error</i> – średni błąd promieniowy
AUC	<i>area under ROC curve</i> – pole powierzchni pod krzywą ROC
BF	<i>bilateral filter</i> – filtracja bilateralna
CE	<i>cross-entropy</i> – entropia krzyżowa
CNB	<i>core needle biopsy</i> – biopsja gruboigłowa
CNN	<i>convolutional neural network</i> – konwolucyjne sieci neuronowe
CR	<i>calibration Reproducibility</i> – powtarzalność kalibracji
CV	<i>k-fold cross-validation</i> – k-krotna walidacja krzyżowa
DI	<i>Dice index</i> – indeks Dice’a
DOF	<i>dimension of freedom</i> – stopień swobody
EMTS	<i>electromagnetic tracking system</i> – elektromagnetyczny system śledzenia
FG	<i>field generator</i> – generator pola elektromagnetycznego
FN	<i>false negative</i> – fałszywie ujemne
FP	<i>false positive</i> – fałszywie dodatnie
PR	<i>false positive rate</i> – współczynnik wyników fałszywie pozytywnych
F₁	– miara F_1
GT	<i>ground truth</i> – złoty standard

GVF	<i>gradient vector flow</i> – metoda aktywnego konturu uwzględniająca obraz gradientowy
HD	<i>Hausdorff distance</i> – odległość Hausdorffa
IoU	<i>intersection over union</i> - przecięcie sumy
JI	<i>Jaccard index</i> – indeks Jaccarda
λ_R	– najmniejsza odległość pomiędzy P_C a szacowaną trajektorią
$\lambda_{R_{GT}}$	– najmniejsza odległość pomiędzy P_C a trajektorią złotego standardu
M	<i>mean</i> – wartość średnia
MAE	<i>mean absolute error</i> – średni błąd bezwzględny
max	<i>maximum</i> – wartość maksymalna
Me	<i>median</i> – mediana
MCC	<i>Matthews correlation coefficient</i> – współczynnik korelacji Matthews
min	<i>minimum</i> – wartość minimalna
ML	<i>machine learning</i> – uczenie maszynowe
NLSR	<i>needle localization success rate</i> – wskaźnik sukcesu lokalizacji igły
NLM	<i>non-local means</i> – filtracją metodą średnich nielokalnych
NPV	<i>negative predictive value</i> – wartość predykcyjna ujemnego wyniku
OTS	<i>optical tracking system</i> – optyczny system śledzenia
P_C	– centralny punkt obrazu
PM	– autorska metoda
PPV	<i>positive predictive value</i> – precyzja przewidywania pozytywnego
PRE	<i>precision</i> – precyzja
PCA	<i>principal component analysis</i> – analiza głównych składowych
RG	<i>region growing</i> – metoda rozrostu obszaru
RMSE	<i>root mean square error</i> – błąd średniokwadratowy

ROI	<i>region of interest</i> – region zainteresowania
ROC	<i>receiver operating characteristic</i> – krzywa charakterystyki roboczej odbiorcy
SD	<i>standard deviation</i> – odchylenie standardowe
SEN	<i>sensitivity</i> – wrażliwość
SIR	<i>similarity ratio</i> – współczynnik podobieństwa
SPE	<i>specificity</i> – specyficzność
SVM	<i>support vector machine</i> – maszyna wektorów nośnych
θ_R	– kąt pomiędzy osią horyzontalną obrazu a trajektorią wyznaczoną na podstawie wyniku segmentacji
θ_{RGT}	– kąt pomiędzy osią horyzontalną obrazu a trajektorią wyznaczoną na podstawie złotego standardu
TN	<i>true negative</i> – prawdziwie ujemne
TP	<i>true positive</i> – prawdziwie dodatnie
TPR	<i>true positive rate</i> – współczynnik wyników prawdziwie pozytywnych
WS	<i>watershed segmentation</i> – algorytm wododziałowy
USG	– obrazowanie ultrasonograficzne

1. Wprowadzenie

Nowotwory złośliwe są w Polsce i w Europie drugą co do częstości przyczyną zgonów ustępując jedynie chorobom układu sercowo-naczyniowego [46]. Obejmują one bardzo zróżnicowaną grupę chorób, zatem nie istnieje i nigdy nie powstanie jedno uniwersalne „lekarstwo na raka”. Choroby te stanowią tak duży problem badawczy, kliniczny czy społeczny, że konieczne jest prowadzenie prac o szerokim zakresie interdyscyplinarnym zmierzających do jak najlepszego ich poznania.

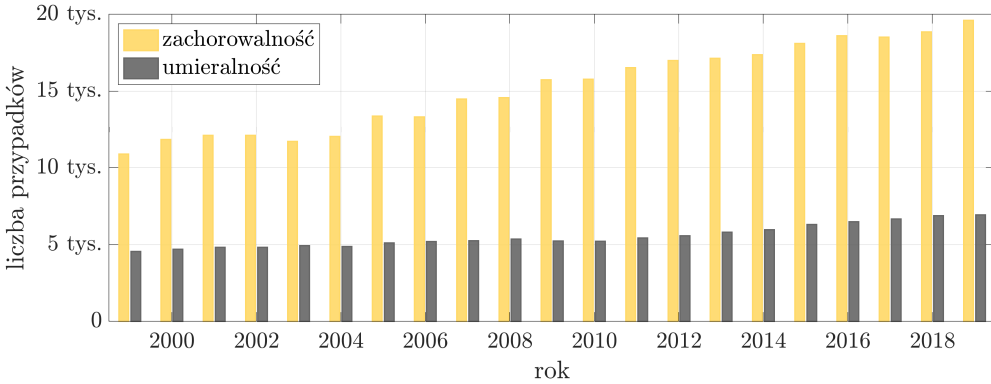
1.1 Nowotwór złośliwy gruczołu sutkowego

Nowotwór złośliwy gruczołu sutkowego (inaczej nowotwór złośliwy sutka lub piersi) jest chorobą, w której w wyniku niekontrolowanego podziału komórek tego narządu tworzy się zmiana guzowata. Powstała zmiana ogniskowa konkuruje z sąsiadującymi zdrowymi tkankami o zaspokajanie swoich zwiększonych potrzeb metabolicznych, co wpływa na stan całego organizmu prowadząc do jego wyniszczenia [84].

Określenie „nowotwór złośliwy piersi” nie oznacza tej samej choroby co „rak piersi”. To drugie pojęcie odnosi się wyłącznie do tych nowotworów złośliwych, które rozwijają się z tkanki nabłonkowej. Zatem raki piersi stanowią (dominujący) podzbiór wszystkich nowotworów złośliwych tego narządu. Polskie nazewnictwo nie jest tutaj zgodne z terminologią anglojęzyczną, gdyż angielskie pojęcie *cancer* oznacza każdy nowotwór złośliwy, a polski „rak” jest tłumaczony jako *carcinoma*.

Większość przypadków raków sutka rozwija się z komórek przewodów wyprowadzających albo komórek gruczołu mlecznego. Z tego powodu choroba ta znacząco częściej dotyczy kobiet niż mężczyzn. Szacuje się, że w 2019 r. W Stanach Zjednoczonych Ameryki wystąpiło 316 700 nowych przypadków nowotworu gruczołu sutkowego wśród kobiet oraz około 2 670 wśród mężczyzn. Nowotwór ten był wówczas przyczyną śmierci 41 760 kobiet i 500 mężczyzn [14].

Nowotwór złośliwy gruczołu sutkowego jest najczęściej występującym rodzajem nowotworu złośliwego wśród kobiet w Polsce. Według danych Krajowego Rejestru Nowotworów [186] w 2018 roku rozpoznano 18 869 przypadków tej choroby, co stanowiło niemal 20% wszystkich rozpoznanych nowotworów



Rys. 1.1: Statystyka zachorowalności oraz śmiertelności z powodu nowotworu gruczołu sutkowego w Polsce w latach 1999-2019¹.

złośliwych w grupie kobiet. U mężczyzn natomiast liczba przypadków wyniosła 154, co stanowiło niespełna 2‰ zachorowań w tej populacji. Przypadki zachorowalności oraz śmierci z powodu nowotworu gruczołu sutkowego w latach 1999-2019 na terenie Polski, przedstawiono na rysunku 1.1.

Według danych Światowej Organizacji Zdrowia (WHO), w roku 2020 na całym świecie chorobę nowotworową zdiagnozowano u ponad 19,3 mln pacjentów i była przyczyną śmierci prawie 10 mln osób [187]. Nowotwór złośliwy piersi dotyczył 2,2 miliona osób i był to najczęstszy nowotwór złośliwy na świecie obok nowotworu złośliwego płuca. Z jego powodu zmarło wówczas 685 tysięcy osób [160]. Na przestrzeni ostatnich 25 lat obserwowana jest globalna tendencja wzrostowa w liczbie zarówno zachorowań, jak i zgonów z powodu nowotworu złośliwego gruczołu sutkowego [6, 33].

Klasyfikacja nowotworów gruczołu sutkowego obejmuje lokalizację, typ histopatologiczny, stopień złośliwości histologicznej (ang. *grading*), stadium zaawansowania klinicznego (ang. *staging*) według klasyfikacji TNM. American Joint Committee on Cancer definiuje cztery stopnie oceny stopnia zaawansowania od najłagodniejszego (stopień I) do najpoważniejszego (stopień IV). Obecna klasyfikacja WHO z 2019 roku kładzie większy nacisk również na ocenę ekspresji genowej i zaburzenia molekularne nowotworu [188].

Mimo, że nie jest możliwe przewidzenie czasu przeżycia od chwili rozpoznania konkretnego przypadku tej choroby, to globalnie 5-letnie przeżycie wśród kobiet waha się od 87% w stopniu I choroby do 13% w stopniu IV [23]. Wczesna diagnoza jest zatem istotnym aspektem w przypadku nowotworu gruczołu sutkowego (jak również każdego innego nowotworu), ponieważ pozwala pacjentkom i pacjentom na wczesne rozpoczęcie procesu terapeutycznego, a tym samym lepsze rokowanie i wyższy wskaźnik przeżycia.

¹ źródło: Krajowy Rejestr Nowotworów <http://onkologia.org.pl>

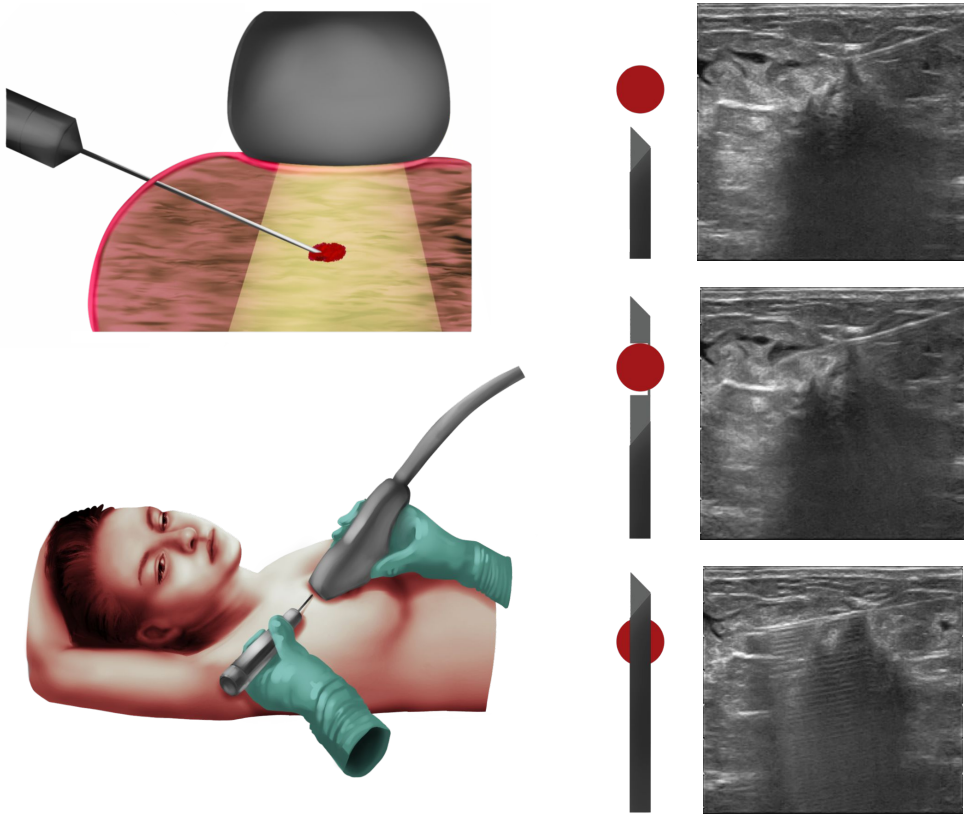
Według Światowej Organizacji Zdrowia (WHO), globalny niedobór pracowników ochrony zdrowia osiągnie 12,9 mln do 2035 roku [52]. Oznacza to wysokie prawdopodobieństwo niewystarczającej liczby lekarzy radiologów oceniających obrazy medyczne zawierające zmiany ogniskowe. Zwiększy to obciążenie pracą dostępnych radiologów i może prowadzić do opóźnień w leczeniu, i z tego rokowania pacjentów.

1.2 Biopsja gruboigłowa gruczołu sutkowego

Na przestrzeni lat opracowano liczne metody diagnostyki i terapii nowotworów piersi. Podstawowe parametry oceny progresji klinicznej choroby stanowią rozmiar, lokalizacja i typ zmiany ogniskowej [127], które określane są z wykorzystaniem zabiegu biopsji gruboigłowej (ang. *core needle biopsy, CNB*). Zabieg ten wykonywany jest z użyciem tzw. pistoletu biopsyjnego, który jest wyposażony w igłę o średnicy 1,2–2,3 mm [164]. Pod kontrolą ultrasonograficzną igła biopsyjna wprowadzana jest do gruczołu sutkowego, a następnie pobierany jest fragment tkanki zmiany ogniskowej i jej otoczenia, który później poddawany jest ocenie histopatologicznej [163]. Jednorazowo zazwyczaj pobiera się od czterech do sześciu wycinków. Dla pogłębienia diagnostyki i ułatwienia późniejszego zabiegu chirurgicznego, po pobraniu wycinków, w to samo miejsce gruczołu sutkowego wprowadzany jest metalowy marker (tzw. kotwica). Identyfikacja lokalizacji pobrania próbek realizowana jest następnie poprzez wykonanie badania mammograficznego i określenie położenia kotwicy [98]. Niestety istnieje ryzyko, że zostanie ona pozostawiona niedokładnie w punkcie pobrania materiału oraz że ulegnie ona przemieszczeniu w obrębie gruczołu [164].

Trudność w prawidłowym wykonaniu badania biopsji gruboigłowej wynika z braku pełnej kontroli wzrokowej nad lokalizacją igły w obrazie USG. Spowodowane to jest deformowaniem i przemieszczaniem zmiany ogniskowej w trakcie wykonywania badania. Ponadto obraz USG jest w wysokim stopniu zaszumiony, a sama igła często nie znajduje się w płaszczyźnie obrazowania, przez co nie jest widoczna w obrazie. Wsparciem podczas zabiegu biopsji mogą być przystawki do głowicy ultrasonograficznej, za pomocą których igła jest pozycjonowana w płaszczyźnie obrazu, a także głowice ultrasonograficzne z wbudowanym centralnym kanałem biopsyjnym. Rozwiązania te jednak nie zawsze mogą być zastosowane ze względu na anatomiczną lokalizację zmiany.

Do zalet zabiegu biopsji gruboigłowej należy zaliczyć brak konieczności stosowania promieniowania rentgenowskiego oraz niski koszt badania. W przypadku wykonywania zabiegu przez doświadczonego lekarza radiologa, czas trwania zabiegu jest krótki (poniżej 10 minut). Czułość badania w dużej mierze uzależniona jest od liczby i lokalizacji pobranych wycinków tkanki.



Rys. 1.2: Schemat zabiegu biopsji gruboigłowej gruczołu sutkowego pod kontrolą ultrasonograficzną — ilustracja Monika Boroń.

1.3 Ultrasonografia

Ultrasonografia (ang. *ultrasonography*, USG) stanowi jedną z podstawowych technik obrazowania medycznego, które umożliwiają prezentację fizjologicznych i patologicznych struktur w ludzkim ciele [58]. USG jest stosowane często jako pierwsze badanie obrazowe w ocenie narządów wewnętrznych, naczyń krwionośnych, układu mięśniowo-szkieletowego czy w diagnostyce położniczej. Obrazy ultrasonograficzne wspomagają również diagnostykę chorób tarczycy [106], wątroby [74] lub zmian skórnych [35].

Rozpowszechnienie USG spowodowane jest licznymi zaletami tego badania w porównaniu do innych metod obrazowania medycznego: tomografii komputerowej (TK) i obrazowania rezonansu magnetycznego (RM). Zalety te to możliwość obserwacji ciała w czasie rzeczywistym, brak występowania rentgenowskiego promieniowania jonizującego (tylko TK) oraz znacznie niższe koszty realizacji badania.

Ponadto aparat USG jest urządzeniem mobilnym o niskim poborze mocy elektrycznej oraz niewymagającym specjalistycznych pomieszczeń do swej pracy. Cechy te pozwalają na szerokie zastosowanie ultrasonografii zarówno w zakładach opieki zdrowotnej jak i w warunkach domowych [13].

1.3.1 Powstawanie obrazu USG

W badaniu USG dwuwymiarowy obraz uzyskiwany jest w czasie rzeczywistym w wyniku ciągłej naprzemiennej emisji i detekcji fal ultradźwiękowych. Fala o częstotliwości z zakresu od ok. 1 do 20 MHz emitowana jest w kierunku badanych narządów [58]. Na granicach tkanek, które wykazują różnice w impedancji akustycznej, powstają echa ultradźwiękowe, które powracają do głowicy. Amplituda fali powracającej (jej echo) zostaje odebrana przez głowicę, w której kryształach piezoelektrycznych energia mechaniczna fali ulega konwersji na sygnał elektryczny. Z sygnału tego finalnie jest rekonstruowany obraz. Może on zawierać artefakty, wśród których najczęstsze są cienie akustyczne, cienie brzegowe [102], szum multiplikatywny [36], pozorne wzmocnienie czy wielokrotne odbicia [8].

Technika obrazowania ultrasonograficznego cechuje się bardzo wysoką dokładnością, gdyż umożliwia rozróżnienie struktur, których wymiar wynosi już 0,1 mm. Dokładność taka jednak jest osiągalna jedynie w obrazowaniu struktur powierzchniowych; wraz z głębokością maleje rozdzielczość obrazowania. Jakość obrazu jest również uzależniona od innych czynników, do których należą liczba i wielkość elementów elektronicznych głowicy, częstotliwość emitowanej fali czy kąt rozchodzenia wiązki fali w głowicy konweksowej.

1.3.2 Zastosowanie obrazowania USG

Obrazowanie ultrasonograficzne jest w medycynie szeroko stosowane, m. in. w diagnostyce narządów jamy brzucha, struktur serca (echokardiografia), naczyń krwionośnych (USG dopplerowskie), gałki ocznej, a także w badaniach prenatalnych i w położnictwie. Stosowane jest również jako technika wspomagająca przeprowadzanie zabiegów biopsji, szczególnie gruczołu sutkowego, stercza czy tarczycy, ponadto pod kontrolą USG przeprowadza się niektóre przewodowe znieczulenia anestetyczne. Struktury wypełnione gazem (np. płuca) oraz kości stanowią ograniczenie dla badania ultrasonograficznego, ze względu na prawie całkowite odbicie fali ultradźwiękowej na granicy tych ośrodków [58].

Z użyciem obrazowania USG związane są również inne trudności, wśród których wymienić można wysoką zależność pomiędzy jakością przeprowadzonego badania a doświadczeniem i umiejętnościami lekarza przeprowadzającego badanie [22]. Nie bez znaczenia pozostają również znaczne różnice jakości obrazów tej samej okolicy anatomicznej uzyskiwane różnymi aparatami. Zmienność

osobnicza pacjentów jest kolejnym czynnikiem przekładającym się na jakość wykonywanego badania. Szczególnie w przypadku przeprowadzania diagnostyki onkologicznej określanie charakteru zmian ogniskowych jedynie za pomocą obrazowania ultrasonograficznego zwykle nie jest wystarczające. Z tego powodu rozwijane są obecnie metody automatycznej lub półautomatycznej analizy obrazów ultrasonograficznych, co może pomóc w rozwiązaniu choć niektórych z powyższych problemów [176].

1.3.3 Typy obrazowania USG

W rezultacie wybranego typu konwersji sygnału do obrazu możliwe jest otrzymanie różnych typów obrazu ultrasonograficznego. Najszerszej stosowanym jest obraz typu B (ang. *brightness*) pozwalający na wizualizację dwuwymiarowego przekroju struktur w skali szarości. Prezentacja typu A (ang. *amplitude*) obecnie stosowana jest przede wszystkim w okulistyce, jest to zapis amplitudy sygnału w funkcji głębokości wzdłuż jednej osi. Przepływ krwi w naczyniach obrazowany jest za pomocą ultrasonografii dopplerowskiej (typ D) oraz w badaniach z użyciem kontrastu (ang. *contrast-enhanced ultrasound*, CEUS). W prezentacji typu M (ang. *motion*) wyświetlany jest obraz pochodzący z jednej osi, prezentowany jako kolejne kolumny pikseli odkładane z pewnym opóźnieniem. Wykorzystuje się to przede wszystkim w diagnostyce kardiologicznej. Innym rodzajem obrazowania jest elastografia, która pozwala na określenie stopnia elastyczności tkanek poprzez śledzenie na żywo stopnia ich odkształcenia pod wpływem ucisku. Dzięki głowicom matrycowym i akwizycji sygnału z całego wolumenu pod głowicą, realizowane jest obrazowanie trójwymiarowe oraz czterowymiarowe [58].

1.4 Systemy śledzenia

Systemy komputerowego wspomaganie diagnostyki medycznej (ang. *computer-aided diagnosis*, CAD) stanowią integralną część nowoczesnej opieki nad pacjentem. Efekt ten jest związany z szeregiem oczekiwanych korzyści w porównaniu do konwencjonalnych podejść poprzez: zwiększenie dokładności, zmniejszenie powikłań i skrócenie czasu interwencji. Korzyści wykorzystywania systemów wspomaganie diagnostyki zostały wykazane w różnych zabiegach, szczególnie z zakresu neurochirurgii oraz ortopedii [86, 170]. W zakresie systemów komputerowego wspomaganie diagnostyki medycznej stosowane są techniki terapii oparte na nawigacji obrazowej (ang. *image-guided therapy*, IGT), które opierają się na rejestracji danych pacjenta przedoperacyjnego (model pacjenta przygotowany za pomocą np. tomografii komputerowej lub rezonansu magnetycznego) i danych z badań śródoperacyjnych (np. ultrasonografii) podczas zabiegu operacyjnego. Bazując na tej informacji system wspomaganie może

zapewnić swobodną nawigację lub wskazówki do pozycjonowania śledzonego instrumentarium medycznego [42], którego przykładem może być pistolet biopsyjny w trakcie zabiegu biopsji gruboigłowej lokalizowany względem położenia głowicy ultrasonograficznej.

1.5 Przedstawienie problemu, cel i zakres prac

Prawidłowa diagnostyka zmian ogniskowych gruczołu sutkowego wymaga zastosowania różnych technik obrazowania medycznego. Jednym z ważniejszych zabiegów diagnostycznych jest biopsja gruboigłowa wykonywana pod kontrolą ultrasonograficzną, w trakcie której do gruczołu sutkowego wprowadzana jest igła biopsyjna, którą radiolog pobiera od czterech do sześciu próbek tkanki do dalszej diagnostyki. Próbkę te mogą być oceniane histopatologicznie, immunohistochemicznie i molekularnie i na ich podstawie stawiana jest ostateczna diagnoza i wybierany sposób leczenia [135, 164]. Poprawność rozpoznania w dużej mierze zależy od precyzji lokalizacji pobrania bioptatu, i w razie poprawnie pobranych próbek dokładność diagnostyczna biopsji gruboigłowej przekracza 95% [39].

Poprawne pobranie fragmentu tkanki wymaga dokładnego wprowadzenia igły biopsyjnej do określonej strefy zmiany ogniskowej. Lekarz wykonujący biopsję korzysta z pomocy dwuwymiarowego, zaszumionego obrazu ultrasonograficznego, w którym wizualizacja igły nie zawsze jest poprawna. Igła jest wyraźnie widoczna jedynie wówczas, gdy znajduje się w całości w płaszczyźnie obrazu. Jeśli igła leży równolegle do czoła głowicy, występuje silne lustrzane odbicie, które poprawia jakość obrazu igły, jednak może powodować artefakt pogłosu. Dodatkowo sąsiednie, wyraźne podłużne, struktury anatomiczne (np. granice powięzi) mogą utrudniać prawidłowe umiejscowienie igły w obrazie [66].

Segmentacja zmiany ogniskowej w obrazie USG na żywo, podczas przeprowadzanej biopsji, w połączeniu z lokalizacją trajektorii igły w obrazie, doprowadzi do ułatwienia zabiegu biopsji gruboigłowej. Opisana powyżej analiza problemu stanowi podstawę do sformułowania tezy badawczej niniejszej pracy:

Wykorzystanie obrazów ultrasonograficznych zebranych przy użyciu systemu śledzenia i rejestracji pozwala na określenie trajektorii igły biopsyjnej oraz na detekcję zmiany ogniskowej w zabiegu biopsji gruboigłowej gruczołu sutkowego.

Do wykazania słuszności tezy opracowana została metodologia akwizycji oraz przetwarzania dwuwymiarowych obrazów ultrasonograficznych zintegrowana w multimodalnym systemie śledzenia i rejestracji w zabiegach biopsji gruboigłowej gruczołu sutkowego. Zakres prac koniecznych do opracowywania tego systemu obejmuje następujące zadania badawcze:

1. Przegląd aktualnej literatury naukowej z zakresu klasyfikacji i segmentacji struktur w obrazach ultrasonograficznych z wykorzystaniem metod

klasycznych i głębokich oraz z uwzględnieniem systemów nawigacji w zabiegach biopsji pod kontrolą ultrasonograficzną.

2. Opracowanie systemu rejestracji danych:

- konstrukcja stanowiska pomiarowego oraz integracja zapisywanych danych,
- kalibracja systemów śledzenia (optycznego i elektromagnetycznego),
- kalibracja pistoletu biopsyjnego,
- kalibracja obrazu ultrasonograficznego,
- wyznaczenie dokładności systemu rejestracji.

3. Opracowanie metod klasyfikacji i segmentacji zmiany ogniskowej w dwuwymiarowych obrazach USG:

- klasyfikacja obecności zmiany i typu zmiany ogniskowej w obrazie ultrasonograficznym z wykorzystaniem konwolucyjnych sieci neuronowych,
- filtracja obrazów z obecną zmianą ogniskową,
- segmentacja zmiany ogniskowej z wykorzystaniem komitetów klasyfikujących,
- przygotowanie zbioru obrazów wraz z obrysami eksperckimi,
- analiza dokładności przedstawionej metody.

4. Opracowanie algorytmów klasyfikacji i segmentacji wraz z wyznaczeniem trajektorii igły biopsyjnej w obrazach 2D:

- klasyfikacja obecności igły w obrazie ultrasonograficznym z wykorzystaniem konwolucyjnych sieci neuronowych,
- segmentacja semantyczna igły w obrazach USG,
- wyznaczenie trajektorii igły biopsyjnej,
- przygotowanie zbioru obrazów wraz z obrysami eksperckimi,
- analiza dokładności proponowanej metody.

Przedstawiony w niniejszej rozprawie system śledzenia i rejestracji był tworzony podczas prac w projekcie naukowo-badawczym MILESTONE realizowanym w granicę Narodowego Centrum Badań i Rozwoju STRATEGMED 2 (numer projektu STRATEGMED2/267398/4/NCBR/2015).

1.6 Układ rozprawy

Niniejsza rozprawa składa się z sześciu rozdziałów oraz załącznika. Rozdział pierwszy zawiera wprowadzenie w tematykę nowotworów gruczołu sutkowego oraz wybranych metod diagnostycznych stosowanych w tej chorobie: ultrasonografii oraz – wykonywanej pod jej kontrolą – biopsji gruboigłowej. Rozdział zakończono przedstawieniem tezy badawczej oraz celu pracy.

Rozdział drugi obejmuje przegląd aktualnego stanu wiedzy z zakresu przetwarzania obrazów ultrasonograficznych. Rozdział ten podzielono na części, w których kolejno zaprezentowano: techniki diagnostyki radiologicznej gruczołu sutkowego; metody segmentacji i klasyfikacji zmian ogniskowych w obrazie USG; algorytmy detekcji i wyznaczania trajektorii igły biopsyjnej w obrazie USG; porównanie optycznych oraz elektromagnetycznych systemów śledzenia.

W rozdziale trzecim opisano budowę i działanie autorskiego systemu rejestracji danych zwracając szczególną uwagę na kalibrację pistoletu biopsyjnego oraz obrazu ultrasonograficznego.

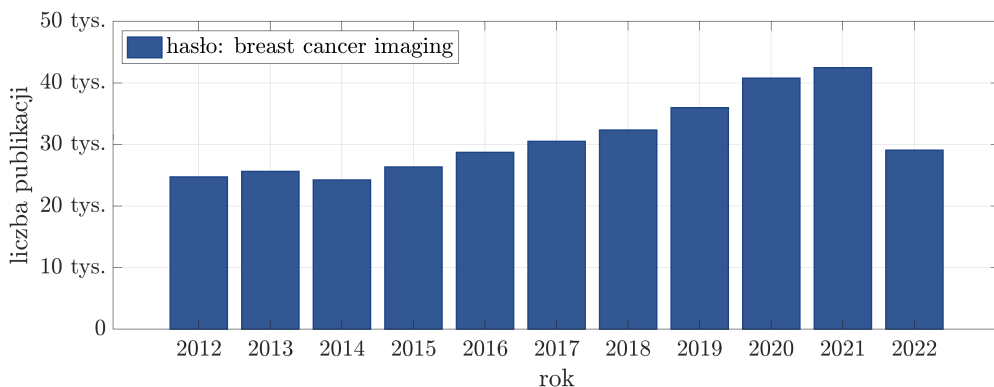
Rozdział czwarty przedstawia opis metodologii klasyfikacji i segmentacji obrazów ultrasonograficznych. Rozdział ten w pierwszej części zawiera ogólne informacje o architekturze konwolucyjnych sieci neuronowych oraz szczegółową prezentację autorskiej architektury klasyfikacji treści obrazów ultrasonograficznych. Dalej przedstawiono sposób podziału danych na zbiór treningowy, walidacyjny i testowy, augmentację zbiorów danych oraz zaprezentowano proces uczenia sieci. W kolejnym podrozdziale o segmentacji semantycznej igły biopsyjnej zaproponowaną autorską architekturą opartą o konwolucyjne sieci neuronowe przedstawiono analogiczne informacje o wykorzystanej sieci. Ostatnia część rozdziału zawiera opis segmentacji zmiany ogniskowej metodami klasycznymi, których wyniki agregowane są za pomocą komitetu klasyfikującego.

W rozdziale piątym przedstawiono wyniki eksperymentów weryfikujących dokładność poszczególnych elementów – zarówno skonstruowanego systemu rejestracji danych, jak również opisywanych algorytmów klasyfikacji i segmentacji poszczególnych struktur w obrazach ultrasonograficznych.

Ostatnim rozdziałem rozprawy jest podsumowanie zawierające propozycję praktycznego zastosowania systemu oraz plany dalszych badań wynikających z niniejszej pracy.

2. Istniejące rozwiązania

W związku z wysokimi wartościami zachorowalności i umieralności z powodu nowotworów gruczołu sutkowego i wielkich kosztów społecznych tej choroby, na przestrzeni lat opracowano wiele metod diagnozowania i leczenia nowotworów gruczołu sutkowego. Ma to odbicie w szeregu publikacji poświęconych tej chorobie (Rys. 2.1). W niniejszym rozdziale opisany został stan wiedzy z zakresu diagnostyki obrazowej – szczególnie obrazowania ultrasonograficznego – skupionej na ocenie zmian ogniskowych gruczołu sutkowego oraz zabiegu biopsji gruboigłowej wykonywanej pod kontrolą ultrasonograficzną. Dalsza część rozdziału poświęcona jest opisowi systemów śledzenia.

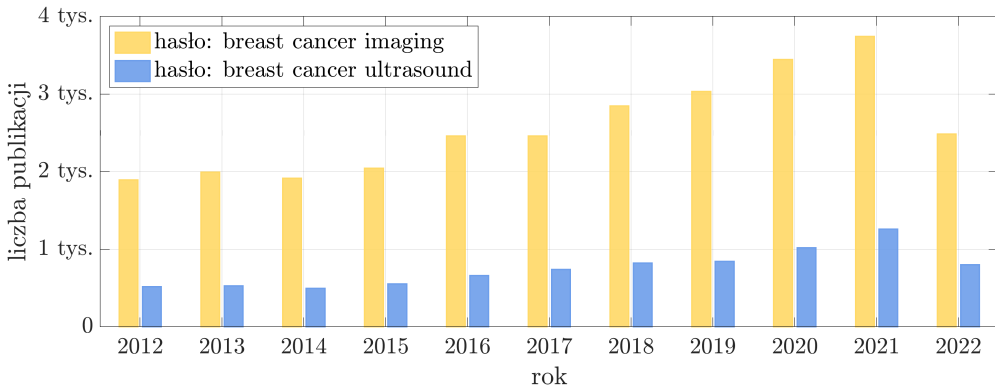


Rys. 2.1: Liczba publikacji poświęcona tematyce raka piersi. Stan na wrzesień 2022¹.

2.1 Diagnostyka radiologiczna gruczołu sutkowego

Nowotwór złośliwy gruczołu sutkowego stanowi najczęściej występujący typ nowotworu złośliwego u kobiet i najczęstszą przyczynę zgonów z powodu nowotworów złośliwych na całym świecie [41]. Określenie wymiarów zmiany ogniskowej oraz jej lokalizacja są ważnymi wskaźnikami stosowanymi w ocenie progresji.

¹ źródło: Dimensions AI | The most advanced scientific research database <https://app.dimensions.ai/discover/publication>



Rys. 2.2: Liczba publikacji z zakresu technik obrazowania zmian ogniskowych gruczołu sutkowego, na podstawie wyszukiwania hasłami *breast cancer imaging* i *breast cancer ultrasound*. Stan na wrzesień 2022¹.

sji klinicznej choroby [127]. Parametry te szacowane są z wykorzystaniem metod obrazowania medycznego, zarówno w celowanym procesie diagnostycznym, jak również w procedurach przesiewowych wykonywanych w kierunku wykrycia zmian nowotworowych. Na przestrzeni lat opracowanych zostało wiele metod analizy danych obrazowych, a liczba publikacji na ten temat rośnie w czasie (Rys. 2.2). Trend ten wskazuje na potrzebę dalszych badań poświęconych tematyce obrazowania zmian ogniskowych gruczołu sutkowego.

2.1.1 Mammografia

Podstawową metodą wczesnego wykrywania raka gruczołu sutkowego jest mammografia, w której wykonywane jest prześwietlenie promieniowaniem rentgenowskim. Otrzymany obraz może ujawnić obecność mikrozwapnień i małych guzów, które powinny zostać poddane dalszej ocenie diagnostycznej. Wadami tej metody jest inwazyjność zastosowanego promieniowania jonizującego oraz konieczność wykonania często bolesnej kompresji piersi do uzyskania obrazów o wysokiej jakości. Czułość badania mammograficznego jest zależna od gęstości tkanki gruczołu sutkowego, ponieważ przy wyższej gęstości zmiany ogniskowe mogą pozostać niewidoczne [15].

2.1.2 Ultrasonografia

Ultrasonografia piersi (ang. *breast ultrasound*) jest często stosowana jako uzupełniająca technika obrazowania mammograficznego [14, 22], a także jako

¹ źródło: Dimensions AI | The most advanced scientific research database <https://app.dimensions.ai/discover/publication>

samodzielne badanie medyczne [47]. Celem obrazowania USG jest detekcja [105, 197] i ocena zmian ogniskowych gruczołu sutkowego [172].

Technika obrazowania ultrasonograficznego pozwala zlokalizować oraz poddać ocenie wizualnej zmianę ogniskową w bezpiecznej dla pacjentki procedurze medycznej. Dwuwymiarowy obraz USG pozwala potwierdzić obecność zmiany ogniskowej, a na jego podstawie możliwe jest oszacowanie wymiarów tej zmiany. Ultrasonografia umożliwia również wstępne określenie złośliwości występującej zmiany [204], która następnie potwierdzana jest badaniem histopatologicznym [15, 47].

Podejście wykorzystujące ultrasonografię do oceny wizualnej zmiany ogniskowej stosowane jest nie tylko w przypadku gruczołu sutkowego, ale także innych wewnętrznych struktur ciała, takich jak na przykład ścięgna, mięśnie, naczynia krwionośne i narządy wewnętrzne. Obrazowanie USG stanowi również wsparcie w trakcie zabiegów o różnym stopniu złożoności [26, 135, 164] ze względu na możliwość obrazowania struktur w czasie rzeczywistym.

Wizualizacja gruczołu sutkowego za pomocą USG realizowana jest na podstawie oceny echogeniczności tkanek. Echogeniczność (ang. *echogenicity*) definiuje się jako poziom jasności obrazu, która odwzorowuje odbitą falę dźwiękową w uzależnieniu od gęstości obrazowanej tkanki [158]. Obszary odpowiadające niższym wartościom intensywności niż otaczające tło określane są jako hipoechogeniczne (hypoechogeniczne), a odpowiadające wyższym – hiperechogeniczne (hyperechogeniczne) [179]. Zmiana ogniskowa obrazowana za pomocą USG jest widoczna jako obszar o charakterystycznej echogeniczności różniującej tę zmianę od otaczających zdrowych tkanek. Zdolność zmiany ogniskowej do odbijania fal dźwiękowych umożliwia wizualizację tej zmiany w formie obszaru hipo- lub hiperechogenicznego [99].

W obrazach ultrasonograficznych często występującym utrudnieniem jest obecność szumów i cieni powodowanych echem. W przypadku tego obrazowania obecny jest multiplikatywny szum plamkowy o losowym charakterze [99]. Obecność ziarnistego szumu plamkowego jest charakterystyczna dla obrazów ultrasonograficznych [36] prowadząc do rozmywania krawędzi i drobnych szczegółów [59], co zmniejsza rozdzielczość kontrastu i utrudnia identyfikację zmienionych chorobowo tkanek [44].

Jednoznaczność granicy zmian ogniskowych jest również zaburzana tzw. spikulacjami (ang. *spiculation*), które są gwiazdzistymi zniekształceniami powodowanymi migracją komórek nowotworowych do otaczającej tkanki [69]. Efektem jest obecność cieni akustycznych [102]. Spikulacje częściej towarzyszą złośliwym zmianom nowotworowym niż łagodnym, co może być wskazówką w trakcie klasyfikacji typu zmiany ogniskowej [146]. Powyższe zaburzenia wpływają negatywnie na proces obiektywnego szacowania wymiarów zmiany lub określenia punktów pobrania wycinków tkanek do analizy histopatologicznej.

Na ocenę obrazów ultrasonograficznych ma wpływ zarówno użyta aparatura

tura, jak również doświadczenie lekarza radiologa, co przekłada się na poprawność detekcji zmiany ogniskowej w obrazie oraz wykonywanych obrysów [22]. Automatyzacja procesu detekcji, segmentacji i klasyfikacji zmian ogniskowych pozwala na zwiększenie obiektywności i powtarzalności oceny tych zmian poprzez komputerowy system wspomagania diagnostyki [47], odciążając diagnostę w żmudnym i czasochłonnym procesie manualnej oceny danych obrazowych [180].

2.2 Segmentacja i klasyfikacja zmian ogniskowych w obrazie USG

2.2.1 Segmentacja zmian ogniskowych

Na przestrzeni lat zaproponowano szereg algorytmów zarówno klasycznej, jak i semantycznej segmentacji w przetwarzaniu dwu- i trójwymiarowych obrazów ultrasonograficznych, co wraz z wynikami zostało zaprezentowane w tabelach 2.1 i 2.2 dla obrazów 2D oraz 2.3 i 2.4 dla danych 3D.

Tab. 2.1: Podsumowanie metod segmentacji zmian ogniskowych w obrazach ultrasonograficznych 2D.

Ref.	Rok	Przetwarzanie wstępne	Region zainteresowania	Finalna segmentacja
[193]	2021	–	wykrywanie granic	GG-Net z metodą global guidance block (GGB)
[17]	2020	skalowanie	–	selective kernel U-Net
[47]	2020	skalowanie	–	(1) FCN-AlexNet; (2) U-Net; (3) SegNet; (4) DeepLabV3+
[196]	2019	skalowanie	–	(1) FCN-AlexNet; (2) FCN-32s; (3) FCN-16s; (4) FCN-8s
[197]	2018	skalowanie	–	(1) patch-based LeNet; (2) U-Net; (3) wstępnie trenowany FCN-AlexNet
[43]	2018	dyfuzja anizotropowa	manualny wybór punktów konturu	aktywny kontur

[190]	2015	wygładzanie Gaussa; adaptacyjne nieliniowe mapowanie intensywności	punkty startowe oparte o anatomię; rekonstrukcja morfologiczna; metoda Otsu	optymalizacja wykresu maksymalnego przepływu z ograniczeniami przestrzennymi i częstotliwościowymi
[68]	2015	model całkowitej zmienności	grupowanie; maszyna wektorów nośnych (SVM) z wykorzystaniem funkcji intensywności, gradientu i kształtu	aktywny kontur
[90]	2015	filtracja medianowa	fragmentacja obrazu	metoda rozrostu obszaru
[137]	2013	wygładzanie z zachowaniem krawędzi metodami Lee, Frost, Kuan	metoda hierarchicznie rozmytych C-średnich	aktywny kontur uwzględniający rozszerzenie gradientu (GVF)
[76]	2012	–	AdaBoost; maszyna wektorów nośnych (SVM)	Random Walk
[67]	2004	sieć neuronowa typująca filtry na podstawie analizy tekstur	poziomy szarości	algorytm wododziałowy oparty o znaczniki

Tab. 2.2: Podsumowanie metod wykrywania zmian ogniskowych w obrazach ultrasonograficznych 2D – podsumowanie baz danych oraz miar z otrzymanymi wartościami. Objasnienia skrótów – strona xvi.

Ref.	Rok	Typ danych	Liczba obrazów	Miara	Wartość
[193]	2021	800 pacjentek	1412	<i>DI</i> <i>IoU</i> <i>ACC</i> <i>SEN</i> <i>PPV</i>	87,1±1,4 79,1±1,6 97,4±0,3 86,6±1,7 91,3±1,0
[17]	2020	678 zmian łagodnych, 204 zmiany złośliwe	893	<i>DI</i> ogólnie <i>DI</i> zm. łagodne <i>DI</i> zm. złośliwe	0,826 0,820 0,842
[47]	2020	różne źródła	3000	F_1 <i>IoU</i>	>0,90 >0,81 dla SegNet i DeepLabV3+

[196]	2019	Baza A: 60 zm. złośliwych, 246 zm. łagodnych; Baza B: 53 zm. złośliwe, 110 zm. łagodnych	306 163	<i>DI</i> zm. łagodne <i>DI</i> zm. złośliwe	0,7626 dla FCN-16s 0,5484 dla FCN-8s
[197]	2018	Baza A: 60 zm. złośliwych, 246 zm. łagodnych; Baza B: 53 zm. złośliwe, 110 zm. łagodnych	306 163	<i>DI</i> A <i>DI</i> B	0,91 0,89 dla FCN-AlexNet
[43]	2018	1 fantom tkankowy; 6 pacjentek	6 6	<i>DI</i> <i>HD</i> <i>DI</i> <i>HD</i>	78%–92% 6,26–4,89 64%–87% 14,18–6,86
[190]	2015	93 zm. łagodne, 91 zm. złośliwych	184	<i>TPR</i> <i>FPR</i> <i>SIR</i> <i>HD</i> <i>MAE</i>	91,23% 9,97% 83,73% 20,43% 4,93%
[68]	2015	46 zm. ogniskowych	46	<i>TP</i> <i>FP</i> <i>FN</i> <i>ARE</i>	85,01% 1,78% 14,99% 9,08
[90]	2015	20 zm. złośliwe, 10 zdrowych tkanek	30	<i>DI</i>	0,988±0,00147
[137]	2013	195 losowych kształtów fantomów; 48 pacjentek	195 48	<i>TP</i> <i>HD</i> <i>TP</i> <i>HD</i>	97,47%–99,89% 6,40–4,59 87,58%–92,06% 11,59–5,17
[76]	2012	80 zm. ogniskowych, 32 zdrowe tkanki	112	<i>ACC</i> <i>SEN</i> <i>SPE</i>	87,5% 88,8% 84,4%
[67]	2004	60 pacjentek	60	<i>PRE</i>	67,56%–88,80%

Segmentacja zmiany ogniskowej w metodach klasycznych obejmuje etap detekcji celem określenia obszaru zainteresowania (ang. *region of interest*, ROI), w którym przeprowadzana jest właściwa segmentacji obiektu (Tab. 2.1). Detekcja ROI realizowana jest w przypadku metod półautomatycznych poprzez wskazanie manualne punktów w obszarze zmiany [43, 102]. W przypadku metod automatycznych, jako punkty początkowe stosowane są struktury anatomiczne takie jak warstwy skóry, gruczołu sutkowego lub mięśni, widoczne

Tab. 2.3: Podsumowanie metod segmentacji zmian ogniskowych w obrazach ultrasonograficznych 3D.

Ref.	Rok	Przetwarzanie wstępne	Region zainteresowania	Finalna segmentacja
[203]	2021	–	sieć typu koder-dekoder	light-weight multi-scale network
[91]	2021	–	–	mask scoring R-CNN
[88]	2020	3D levelset; dyfuzja anizotropowa	adaptive distance regularized level set evolution (DRLSE)	rekonstrukcja wolumenu
[51]	2016	rekonstrukcja morfologiczna	detektor krawędzi Sobela	algorytm wododziałowy
[105]	2013	dyfuzja anizotropowa; filtr typu sigma	macierz Hessego oparta o detektor skupisk	modele regresji w przestrzeni cech; progowanie
[20]	2003	wyrównanie histogramu; filtracja 3D typu stick	automatyczne progowanie; morfologia	aktywny kontur metodą snake 3D

w obrazie USG [102, 190]. W innym podejściu parametrami początkowymi są poziomy szarości obrazu [67].

Detekcja regionu zainteresowania realizowana jest również bardziej złożonymi metodami opartymi o klasyfikację pikseli. W tym celu wykorzystywane są klasyfikatory AdaBoost [76], maszyna wektorów nośnych (ang. *support vector machine*, SVM) [68, 76] czy metoda hierarchicznie rozmytych C-średnich [137]. Klasyfikacja wartości intensywności umożliwia wskazanie spójnych obszarów. Wynikiem detekcji jest obszar częściowo odpowiadający zmianie ogniskowej, co pozwala na wykorzystanie wskazanych fragmentów jako punktów początkowych w procesie segmentacji [68, 76, 137, 180]. W podejściu opartym o klasyczne metody segmentacji w przypadku przetwarzania dwuwymiarowych obrazów bardzo często stosowana jest segmentacja krawędziowa za pomocą parametrycznego modelu aktywnego konturu – węża (ang. *snake*) [43, 68, 99, 102, 183] lub aktywnego konturu uwzględniającego rozszerzenie zasięgu gradientu [137, 183].

Model aktywnego konturu umożliwia wyznaczenie obrysu obiektu w zasumionych obrazach dwuwymiarowych poprzez iteracyjną ewolucję ciągłych konturów pod wpływem oddziaływania tzw. sił wewnętrznych i zewnętrznych w procesie optymalizacji energii konturu [78]. Wykorzystanie metody aktyw-

Tab. 2.4: Podsumowanie metod wykrywania zmian ogniskowych w obrazach ultrasonograficznych 3D – podsumowanie baz danych oraz miar z otrzymanymi wartościami.

Ref.	Rok	Typ danych	Liczba obrazów	Miara	Wartość
[203]	2021	107 pacjentów	107	<i>DI</i>	0,778±0,145
				<i>IoU</i>	0,650±0,170
[91]	2021	70 pacjentów	–	<i>DI</i>	85,0±10,4
				<i>JI</i>	75,2±14,4
[88]	2020	58 pacjentów, 15 fantomów	70 15	<i>DI</i>	0,86±0,03
				overlap ratio	0,91±0,04
					dla zestawu B
[51]	2016	21 przypadków	21	overlap ratio	85,7%
				<i>IoU</i>	70,3%–80,6%
				segmentation error	10,2%–5,1%
[105]	2013	159 pacjentów	173	<i>SEN</i>	100%
				false positive per pass	17,4
[20]	2003	8 pacjentek	8	match rate	90,2%–99,6%

nego konturu pozwala skutecznie segmentować zmiany ogniskowe w obrazach ultrasonograficznych cechujących się wysokim poziomem zaszumienia [36, 99, 120]. Zaletą modelu aktywnego konturu jest zdolność do segmentacji obiektów o dowolnym kształcie, co często ma miejsce w przypadku zmian ogniskowych gruczołu sutkowego, ale metoda ta jest wrażliwa na dobór parametrów oraz początkowe położenie modelu. Brak standaryzacji i duża zmienność obrazów ultrasonograficznych utrudnia wskazanie właściwego zestawu parametrów.

Innymi stosowanymi narzędziami w przypadku klasycznych algorytmów przetwarzania są metody obszarowe takie jak algorytm wododziałowy wykorzystujący amplitudę gradientu [67, 180, 183] oraz metoda rozrostu obszaru bazująca na wartości intensywności pikseli [90, 183].

Detekcja i segmentacja metodami klasycznymi poprzedzana jest filtracją obrazów. Obrazy ultradźwiękowe są często wstępnie przetwarzane za pomocą wygładzania Gaussa [102, 190] umożliwiającego usunięcie szumu addytywnego. Szum multiplikatywny odfiltrowywany jest za pomocą dyfuzji anizotropowej [43, 102]. Odszumianie obrazów ultrasonograficznych z zachowaniem krawędzi realizowane jest również filtracją medianową [90, 180] oraz metodą średnich nielokalnych (ang. *non-local means*, NLM) [183].

W ostatnich latach segmentacja zmian ogniskowych w dwuwymiarowych obrazach ultrasonograficznych realizowana jest z wykorzystaniem metod semantycznych [17, 47, 193, 196, 197, 205], w których każdy z pikseli poddawany jest klasyfikacji do jednej z klas wskazywanych za pomocą etykiet.

Najczęściej stosowanym podejściem w segmentacji zmian ogniskowych z wykorzystaniem dwuwymiarowych obrazów USG, jest wykorzystanie sieci neuro-

nowej Alexnet typu *fully convolutional network* (FCN) [17, 47, 196]. Zastosowanie FCN znacząco poprawia dokładność trenowanych modeli poprzez transfer wstępnie wytrenowanych wag klasyfikatorów, które są dostrajane w kolejnym etapie uczenia do stawianego segmentacji zadania [148]. Inną często opisywaną architekturą jest również U-Net [17, 47, 197], która została zaprojektowana do segmentacji struktur w obrazach biomedycznych [138]. Architektura ta, oparta o FCN, została zmodyfikowana poprzez dodanie dużej liczby kanałów w operacji transponowanego splotu uwzględniającego próbkowanie mapy cech w górę, co umożliwia propagację informacji kontekstowej do warstw cechujących się wyższą rozdzielczością. Podejście to pozwala na realizację treningu mniejszymi bazami obrazowymi przy zachowaniu efektywności segmentacji [148].

Zbiory obrazów USG z wykonanymi obrysami eksperckimi cechują się relatywnie małym rozmiarem (Tab. 2.2), stąd w tych zbiorach wykorzystywane są architektury wstępnie trenowane, a następnie dostrajane do problemu segmentacji zmiany ogniskowej.

Ze względu na brak wyraźnych granic pomiędzy zmianami a otaczającą tkanką [44], analiza kontekstowa obrazu w klasyfikacji poszczególnych pikseli umożliwia poprawę efektywności działania automatycznych metod segmentacji. W porównaniu do metod klasycznych, powszechnym jest pomijanie etapu detekcji zmiany – analizie poddawany jest cały obraz diagnostyczny. Filtracja obrazów poprzedzająca wykorzystanie sieci głębokich również nie jest powszechnie stosowana, ale wykonywane jest skalowanie obrazów celem dopasowania obrazu do wymogów architektury [17, 47, 196, 197].

W przypadku przetwarzania trójwymiarowych obrazów USG na przestrzeni lat obserwuje się podobne odejście od metod klasycznych [20, 51, 88, 105] ku segmentacji semantycznej [91, 203]. W przypadku segmentacji metodami klasycznymi, podobnie jak w segmentacji obrazów 2D, występują trzy etapy: przetwarzanie wstępne (obejmujące filtrację), detekcja zmiany oraz jej segmentacja (Tab. 2.3). Filtracja realizowana jest na przykład za pomocą dyfuzji anizotropowej [88, 105], jak również metodami wyrównania histogramu czy filtracji 3D typu *stick* [20]. Następnie na podstawie treści obrazu diagnostycznego określany jest region zainteresowania, który wyznacza parametry początkowe metod segmentacji aktywnym konturem 3D [20] albo algorytmem wododziałowym [137]. Ostatnim krokiem przetwarzania jest rekonstrukcja trójwymiarowego modelu zmiany ogniskowej na podstawie maski binarnej [88].

W przypadku segmentacji semantycznej stosowane podejścia uwzględniają przetwarzanie poszczególnych warstw obrazu za pomocą metod typu autoenkoder [203] albo przeprowadzana jest analiza małych regionów obrazu z wykorzystaniem CNN [91]. Pomijany jest etap filtracji w przetwarzaniu wstępnym trójwymiarowych obrazów ultrasonograficznych.

Zdecydowana większość automatycznych metod segmentacji, zarówno w przypadku dwu-, jak i trójwymiarowych obrazów USG, wykorzystuje ob-

razy ciała człowieka. Raportowane zbiory danych w przetwarzaniu metodami klasycznymi obrazów 2D zawierają od 6 [43] do 184 obrazów [190], gdzie mediana liczebności zbiorów przyjmuje wartość 48, a w segmentacji semantycznej od 163 [197] do 3000 obrazów [47] z medianą liczebności zbiorów równą 306. W raportowanych zbiorach bardzo często dokonywany jest podział na zmiany o charakterze łagodnym lub złośliwym oparty o wyniki histopatologiczne [17, 190, 196, 197].

Prace wykorzystujące dane sztuczne (fantomowe) obejmują pojedyncze przykłady [43, 137]. Tkanki ludzkie charakteryzują się większą heterogenicznością oraz licznymi artefaktami, dlatego porównanie metod detekcji pomiędzy zbiorami klinicznymi i fantomowymi wskazuje na niższe wyniki w przypadku danych rzeczywistych (Tab. 2.2 i 2.4). Na przestrzeni lat zostało przygotowanych i upublicznonych kilka baz dwuwymiarowych obrazów ultrasonograficznych, które mogą zostać wykorzystane do procesu treningu metod opartych o uczenie maszynowe oraz pozwolić na ich obiektywne porównanie z metodami opisywanymi w literaturze. Przykładem publicznych źródeł danych z obrysami eksperckimi jest zbiór udostępniony przez szpital Baheya Hospital for Early Detection & Treatment of Women's Cancer (Kair, Egipt) zawierający 779 obrazów USG wykonanych u 561 pacjentek w wieku od 25 do 75 lat za pomocą dwóch aparatów ultrasonograficznych [34] oraz zbiór 163 obrazów zmian łagodnych i złośliwych (odpowiednio 110 i 53 obrazy) pochodzących od 142 pacjentek wykonanych w UDIAT Diagnostic Centre of the Parc Taulí Corporation (Sabadell, Hiszpania) [196].

2.2.2 Klasyfikacja złośliwości zmian ogniskowych

Obrazy ultrasonograficzne mogą stanowić również podstawę do klasyfikacji zmian ogniskowych gruczołu sutkowego w celu określenia charakteru zmiany – łagodnego lub złośliwego [172]. Typ zmiany często powiązany jest z jej cechami morfologicznymi [30]. Zmiany łagodne zwykle są owalnego kształtu oraz mogą osiągać wyższe wartości intensywności. Wizualnie zmiany te w obrazie ultrasonograficznym prezentują się jako struktury hiperechogeniczne, bezechowe lub izoechogeniczne [99]. W przypadku zmian złośliwych często występuje nieregularny kształt i spikulacje [146]. Zmiany złośliwe często są widoczne w obrazach ultrasonograficznych jako obiekty hipoechogeniczne [99]. Torbiele będące zmianami łagodnymi również stanowią przykład obiektów bezechowych lub hipoechogenicznych w obrazach USG [179], co wskazuje, że nie każda zmiana o określonym opisie jest zmianą złośliwą [80]. Dokładne badanie ultrasonograficzne i biopsja mogą pomóc w odróżnieniu ich od nowotworów złośliwych [80]. Automatyczna klasyfikacja zmian ogniskowych na podstawie obrazu USG może stanowić wstępne wnioskowanie do planowania procesu diagnostyczno-terapeutycznego w czasie oczekiwania na wynik analizy histo-

patologicznej [176].

W procesie przetwarzania obrazów ultrasonograficznych do klasyfikacji zmiany, kolejno następują etapy odpowiadające przetwarzaniu wstępnemu i detekcji zmiany ogniskowej, a następnie ekstrakcji cech, które stanowią podstawę do klasyfikacji [99, 102]. Przykładowe rozwiązania problemu klasyfikacji wraz z wyznaczonymi wynikami przedstawiono w tabelach 2.5 i 2.6.

Maszyna wektorów nośnych (ang. *support vector machine*, SVM) jest najczęściej stosowanym klasyfikatorem w ocenie kategorii zmiany ogniskowej na podstawie dwuwymiarowego obrazu USG [30, 94, 99, 102, 199] w oparciu o różne zestawy cech opisujących zmianę. Najprostsze podejście zakłada podział obrazu na fragmenty i ocenę każdego z nich [94] lub wprowadzenie maski wysegmentowanej metodą aktywnego konturu [102]. Aktywny kontur może stanowić również podstawę do wskazania regionu, dla którego szacowane są metryki teksturalne, morfologiczne [99, 152, 176] i histogramowe [99, 134]. Wektor cech zredukowany jest za pomocą analizy głównych składowych (ang. *principal component analysis*, PCA) celem wskazania istotnych statystycznie cech różniących klasy [99]. Analiza tekstur, dla których konstruowana jest na przykład macierz współwystępowania poziomu szarości (ang. *gray-level cooccurrence matrix*, GLCM) albo ocena kształtu, są często stosowanymi deskryptorami klasyfikacji zmian ogniskowych [30, 152].

Innym sposobem klasyfikacji zmiany ogniskowej jest ekstrakcja cech z histogramu intensywności obrazu za pomocą konwolucyjnych sieci neuronowych (ang. *convolutional neural network*, CNN) [134]. Architektury oparte o CNN także realizują zadanie klasyfikacji typu zmiany ogniskowej na podstawie uogólnionych map cech, ekstrahowanych w kolejnych warstwach spłotowych [27, 87, 172]. Podejście to pozwala na analizę również kontekstu zmiany ogniskowej. Ekstrakcja cech głębokich i wykorzystanie metod transfer learning [27, 87, 204] czyli wykorzystanie wstępnie wytrenowanych modeli [5, 147], których przykład stanowi architektura VGG19, również pozwala na określenie charakteru zmiany [18, 28].

Algorytmy automatycznej klasyfikacji trenowane są z wykorzystaniem danych klinicznych (Tab. 2.6). Przedstawiane w literaturze zbiory danych, dla których określony jest typ obrazowanej zmiany ogniskowej, przyjmują wartości z zakresu od 78 [99] do 895 [172] elementów z medianą równą 189. Na przestrzeni lat obserwowana jest tendencja wzrostowa w liczebności wykorzystywanych w procesie uczenia obrazów, co jest skutkiem publikacji kolejnych baz obrazowych.

W przygotowaniu klasyfikatorów wykorzystywane są bazy danych, które również stosowane są w segmentacji zmian ogniskowych gruczołu sutkowego, a ich przykładem jest baza 779 obrazów USG wykonanych u 561 pacjentek w wieku od 25 do 75 lat za pomocą dwóch aparatów ultrasonograficznych [34] oraz zbiór 163 obrazów zmian łagodnych i złośliwych (odpowiednio 110 i 53 ob-

razy) pochodzących od 142 pacjentek wykonanych w UDIAT Diagnostic Centre of the Parc Taulí Corporation (Sabadell, Hiszpania) [196]. Przykładem bazy zawierającej etykiety kategoryzujące obrazowane zmiany ogniskowe, ale bez obrysu eksperckiego, jest publiczna baza danych udostępniona przez Rodriguesa i wsp. w 2017 roku [136]. Zbiór zawiera 250 obrazów łagodnych i złośliwych zmian ogniskowych, które zostały zdiagnozowane u 50 pacjentek.

2.3 Detekcja i segmentacja igły w obrazie USG

Interwencja z wykorzystaniem igły jest stosowana w wielu medycznych zabiegach diagnostycznych i terapeutycznych, takich jak wstrzykiwanie leków, biopsja czy podanie znieczulenia miejscowego [31]. Ultrasonograficzna kontrola toru igły może być przydatna zwłaszcza u pacjentów otyłych [71]. Ultrasonografia jest stosowana też w zabiegach terapeutycznych, takich jak odbarczanie odmy opłucnowej [26] lub paracenteza jamy brzusznej [21]. Niski koszt, nieinwazyjność i możliwości obrazowania położenia igły w czasie rzeczywistym sprawiają, że ultrasonografia jest szeroko stosowana do uzyskiwania obrazów śródzabiegowych [61].

Diagnostyka i terapia zmian ogniskowych piersi obejmuje różne techniki obrazowania medycznego, wśród których szczególnie ważna jest biopsja gruboigłowa wykonywana pod kontrolą USG. W trakcie zabiegu pierś jest nakłuta pistoletem biopsyjnym wyposażonym w dwuczęściową igłę punkcyjną o rozmiarze 18–14 G. Następnie lekarz radiolog pobiera od 3 do 5 próbek tkanek, które poddawane są ocenie histopatologicznej i molekularnej [135, 164]. Ostateczna diagnoza jest zatem w dużej mierze zależna od dokładności nakłucia badanej zmiany ogniskowej. Trafność diagnozy na podstawie poprawnie wykonanej biopsji sięga 95,5% [39]. Biopsja gruboigłowa pod kontrolą ultrasonograficzną znajduje również zastosowanie w diagnostyce zmian ogniskowych innych narządów np. płuc, nerek, wątroby, prostaty lub węzłów chłonnych [150].

2.3.1 Detekcja igły

Widoczność igły na obrazie USG jest związana z lokalizacją głowicy ultrasonograficznej względem niej. Igła jest dobrze widoczna na obrazie USG tylko wtedy, gdy jest prawidłowo ustawiona w stosunku do wiązki ultradźwiękowej. W przypadku, kiedy wiązka jest prostopadła do osi igły, występuje silne odbicie lustrzane, którego efektem jest wzmocnienie obrazu igły [8]. Odwrotne zjawisko w wizualizacji igły występuje w przypadkach, kiedy igła nie znajduje się bezpośrednio w płaszczyźnie obrazowania lub gdy otaczające tkanki zawierają długie, liniowe obiekty (np. granice kości, powięzi) [31]. Lokalizacja położenia igły w obrazie USG jest też utrudniona wówczas, gdy orientacja igły nie odpowiada kątowi wiązki USG. W takiej sytuacji tylko nieznaczny fragment igły

Tab. 2.5: Podsumowanie metod klasyfikacji zmian ogniskowych w obrazach ultrasonograficznych.

Ref.	Rok	Przetwarzanie wstępne	Ekstrakcja cech	Finalna klasyfikacja
[172]	2021	–	–	1) losowe lasy decyzyjne (RF) 2) CNN 3) AutoML
[140]	2020	–	hybrydowy zestaw cech	1) algorytm k-najbliższych sąsiadów (KNN) 2) drzewa decyzyjne (DT) 3) ensemble
[134]	2020	poprawa kontrastu	histogram intensywności skali szarości	CNN
[18]	2018	macierz Grama	liniowa analiza dyskryminacyjna	wstępnie trenowana architektura VGG19
[199]	2016	–	bramkowa maszyna Boltzmanna (PGBM); ograniczona maszyna Boltzmanna (RBM)	maszyna wektorów nośnych (SVM)
[152]	2016	–	457 cechy tekstury i kształtu	połączenie FCM i BPANN
[30]	2016	–	macierz współwystępowania poziomu szarości (GLCM); analiza morfologiczna metryki teksturalne,	maszyna wektorów nośnych (SVM)
[99]	2015	aktywny kontur	morfologiczne i histogramowe; PCA	maszyna wektorów nośnych (SVM)
[102]	2012	wygładzenie Gaussa; dyfuzja anizotropowa	manualny wybór punktów; aktywny kontur	maszyna wektorów nośnych (SVM)
[94]	2010	maksymalna entropia rozmyta	kratownica	maszyna wektorów nośnych (kernelized SVM) oparty na deskryptorach tekstur

jest widoczny na obrazie i jej automatyczna segmentacja jest trudniejsza ze względu na np. pobliskie jasne, liniowe struktury anatomiczne [4], przypominające igłę i tym samym uniemożliwiające jej poprawną lokalizację [55, 66]. Obrazowi igły na obrazie USG często towarzyszy artefakt echa (występowanie

Tab. 2.6: Podsumowanie metod klasyfikacji zmian ogniskowych w obrazach ultrasonograficznych – podsumowanie baz danych oraz miar z otrzymanymi wartościami.

Ref.	Rok	Typ danych	Liczba obrazów	Miara	Wartość
[172]	2021	różne źródła	895	<i>ACC</i>	0,90, 0,91, 0,86
				<i>SEN</i>	0,71, 0,82, 0,84
				<i>SPE</i>	1,00, 0,96, 0,88
				<i>F₁</i>	0,83, 0,87, 0,83
					odpowiednio dla RF, CNN i AutoML
[140]	2020	262 zm. złośliwe, 485 zm. łagodnych	747	<i>ACC</i>	95,36% 95,83% 96,60%
				<i>SEN</i>	92,86% 92,96% 94,34%
				<i>SPE</i>	96,57% 97,24% 97,70%
				<i>PPV</i>	92,86% 94,29% 95,24%
				<i>NPV</i>	96,57% 96,57% 97,25%
					odpowiednio dla KNN, DT i ensemble
[134]	2020	100 zm. złośliwych, 100 zm. łagodnych	200	<i>ACC</i>	95%
[18]	2018	52 zm. złośliwe, 48 zm. łagodnych	200	<i>AUC</i>	0,847
				<i>ACC</i>	0,80
				<i>SEN</i>	0,808
				<i>SPE</i>	0,792
[199]	2016	121 pacjentów: 135 zm. łagodnych, 92 zm. złośliwych	227 SWE	<i>ACC</i>	93,4%
				<i>SEN</i>	88,6%
				<i>SPE</i>	97,1%
[152]	2016	88 zm. łagodnych, 90 zm. złośliwych	178	<i>ACC</i>	95,862±1,002
				<i>SEN</i>	95,139±1,552
				<i>SPE</i>	96,575±1,227
				<i>AUC</i>	95,857±1,061
				<i>MCC</i>	91,732±2,281
[30]	2016	46 zm. złośliwe, 64 zm. łagodnych	100	<i>ACC</i>	98,2%
				<i>SPE</i>	98,4%
				<i>SEN</i>	97,8%
[99]	2015	52 zm. złośliwe, 26 zm. łagodnych	78	<i>ACC</i>	95,7%
				<i>SEN</i>	93,8%
				<i>SPE</i>	100,0%
[102]	2012	40 zm. ogniskowych, 60 zdrowych tkanek	100	<i>ROC</i>	0,88
				<i>SEN</i>	92,7%
				<i>SPE</i>	90,3%
[94]	2010	60 zm. złośliwych, 52 zm. łagodne	112	<i>PRE</i>	82,33%
				<i>SEN</i>	83,81%

zwielokrotnionych obrazów igły) [66].

Precyzyjne wykrycie igły podczas zabiegu biopsji gruboigłowej przekłada się na jakość oceny histopatologicznej i zmniejsza ryzyko uszkodzenia tkanek i nerwów otaczających bioptowaną zmianę ogniskową [39]. Metody szacowania trajektorii igły można podzielić na dwie kategorie: metody sprzętowe oraz programowe.

W metodach sprzętowych pozycja igły jest określana przez urządzenia śledzące oraz czujniki umieszczone na powierzchni pistoletu lub igły. Systemy śledzenia wykorzystują m.in. techniki optyczne oraz elektromagnetyczne [42] bazujące na kalibracji integrującej elementy systemu [86]. Zewnętrzne śledzenie igły można również przeprowadzić za pomocą np. systemu SonixGPS (Ultra-sonix Inc., Richmond, BC, Kanada), w którym położenie igły jest wyznaczane w odniesieniu do obrazu ultrasonograficznego za pomocą elektromagnetycznego systemu pozycjonowania [92]. Metody sprzętowe są wrażliwe na błędy spowodowane elastycznością igły, z wyjątkiem przypadku umieszczenia czujnika bezpośrednio do jej końcówki [189].

Metody programowe detekcji igły bazują na szerokim zakresie algorytmów przetwarzania obrazów, zarówno dwu- jak i trójwymiarowych, począwszy od klasycznych metod segmentacji aż do najnowszych metod bazujących na uczeniu maszynowym. Algorytmy stosowane w przypadku obrazów 2D przedstawione zostały w tabeli 2.7, natomiast oparte na obrazach 3D w tabeli 2.8. Techniki bazujące na danych obrazowych w większości przypadków przyjmują schemat przetwarzania obejmujący trzy etapy: przetwarzanie wstępne, wyznaczenie regionu zainteresowania w oparciu o segmentację igły i określenie na jego podstawie trajektorii igły.

W etapie przetwarzania wstępnego najczęściej stosowane są metody umożliwiające wyodrębnienie jasnych, liniowych obiektów. W tym celu powszechnie używana jest metoda binaryzacji z dobieraną empirycznie lub adaptacyjnie wartością progą [177] lub metoda Otsu [1, 79]. Wartość progowa jest silnie zależna od obrazu, a segmentacja pikseli igły przy użyciu tej metody może powodować nieciągłości w strukturze igły na obrazie USG, dlatego ze względu na te ograniczenia często detekcja igły realizowana jest z użyciem transformacji top-hat [50], filtracji metodą Gabora [2, 25, 79, 129] czy bankiem filtrów Log-Gabora [55]. Filtracja metodą Gabora lub Log-Gabora pozwala na precyzyjną detekcję igły o określonej grubości i wprowadzanej pod określonym kątem.

Kolejnymi technikami umożliwiającymi detekcję jasnych, liniowych obiektów są transformaty Hougha [25, 181] i Radona [2, 31, 32], wykorzystywane również do wstępnego wykrywania położenia igły w obrazie oraz zawężenie obszaru detekcji w kolejnych etapach przetwarzania. W przypadku obrazów trójwymiarowych etap przetwarzania wstępnego często zakłada wykorzystanie macierzy Hessego [201, 202], która opisuje lokalne zmiany drugiego rzędu wartości intensywności woksela. Częstym zabiegiem w przypadku obrazów 3D jest również

ich podział na przekroje (ang. *slice*) lub mniejsze fragmenty (ang. *patch*), które są następnie poddawane przetwarzaniu [128, 194].

2.3.2 Segmentacja igły

Etap segmentacji czyli wyznaczenia obszaru odpowiadającego igle biopsyjnej coraz częściej jest realizowany z wykorzystywaniem technik uczenia maszynowego [4, 50, 89, 111, 128, 194]. Segmentacja igły realizowana jest z wykorzystaniem algorytmów boostingu, których przykładem jest AdaBoost, gdzie duża grupa słabych klasyfikatorów stanowi jeden mocny [55]. Ograniczeniem metody AdaBoost jest długi czas treningu bez możliwości wykorzystania mocy obliczeniowej procesora graficznego. Metody oparte o konwolucyjne sieci neuronowe, niepodlegające temu ograniczeniu, spotykane są zarówno w propozycjach oryginalnych architektur [4, 50] jak i adaptacjach istniejących podejść [89, 128, 194]. Stosowane jest również podejście uwzględniające w architekturze sieci informację czasową (CNN-LSTM, [112]), wprowadzając tym samym zależność pomiędzy kolejnymi ramkami i uodparniając algorytm na przypadkowe wykrycia innych liniowych struktur. Efektywnością cechują się również metody oparte o niedeterministyczny algorytm RANdom SAMple Consensus (RANSAC), który na podstawie zbioru punktów otrzymanych w etapie przetwarzania wstępnego, iteracyjnie wykrywa elementy odstające [79, 201]. Alternatywnym podejściem stosowanym celem wskazania istotnych elementów jest metoda wykorzystująca analizę głównych składowych [202].

Igła, będąca na obrazie długim i cienkim obiektem zajmującym średnio mniej niż jeden procent pikseli całego obrazu (Tab. 5.7), stanowi trudny obiekt do segmentacji [111], szczególnie przy ostrych kątach wprowadzenia igły [29]. Z tego powodu podejścia mające na celu jej precyzyjne wskazanie często wymagają dodatkowego przetwarzania obrazu do wyznaczenia trajektorii igły.

W tym celu wykorzystywane są zarówno proste metody wybierające obiekt o największym konturze spośród wysegmentowanych [89], jak również bardziej złożone oparte o transformatę Radona [29, 55, 182], filtry Gabora [25, 79, 181], bank filtrów log-Gabor [111] czy filtry Kalmana [1, 2, 202]. Ostatni etap przetwarzania danych trójwymiarowych opiera się o algorytm RANSAC [128] lub technikę PCA [4, 201] celem eliminacji wokseli nie należących do linii leżącej w trajektorii igły biopsyjnej.

Określenie współrzędnych końcówki igły realizowane jest różnymi metodami, wśród nich lokalizującymi skrajny punkt wyniku segmentacji [10, 181], wykorzystującymi ruchome okno analizujące intensywność obrazu [32] lub wykrywającymi największy spadek wartości intensywności osi igły [129]. W przypadku podejść wykorzystujących sekwencje obrazowe, współrzędne końcówki wyznaczane są ze średniej wartości sąsiadujących ramek [110]. Predykcja położenia końcówki igły realizowana jest również z wykorzystaniem bardziej złożo-

Tab. 2.7: Podsumowanie metod detekcji igły w obrazach ultrasonograficznych 2D.

Ref.	Rok	Przetwarzanie wstępne	Region zainteresowania	Detekcja igły
[89]	2020	obcinanie	adaptacja architektury LinkNet	największy kontur obiektu
[79]	2020	filtr Gabora; metoda Otsu	RANSAC	filtr Gabora
[1]	2019	filtr medianowy; metoda Otsu	krawędziowanie Canny	filtr Kalmana
[50]	2019	przeskalowanie; interpolacja dwuliniowa	CNN	–
[110]	2019	wzmocnienie oparte o pozycję czasową końcówki igły	CNN	CNN
[25]	2018	transformata Hougha	histogram gradientów (HoG)	filtr Gabora
[111]	2018	transformacja Top-Hat	konwolucyjna sieć neuronowa (FCN)	sieć neuronowa splotowa oparta na regionie (R-CNN); bank filtrów 2D Log-Gabor
[10]	2018	–	filtr Gabora	maszyna wektorów nośnych (SVM)
[181]	2018	transformata Hougha	filtr shock	filtr Gabora
[24]	2015	wygładzenie Gaussa	histogram gradientów (HoG)	kernelized weighted c-means (KWCM)
[55]	2015	bank filtrów log-Gabor 2D	AdaBoost	transformata Radona; filtracja 2. pochodną Gaussa

nych podejść opartych o uczenie głębokie, których przykładem może być detekcja końcówki igły za pomocą algorytmu opartego o architekturę YOLO [109] lub CNN-LSTM [112].

W oparciu o wyznaczoną oś igły, jej końcówka może również zostać wska-

Tab. 2.8: Podsumowanie metod detekcji igły w obrazach ultrasonograficznych 3D.

Ref.	Rok	Przetwarzanie wstępne	Region zainteresowania	Detekcja igły
[85]	2021	filtr Kalmana	naive Bayesian	ważony centroid
[194]	2021	Slice-based UNet	Direction-fused UNet; 3D Pyramid-UNet	–
[4]	2019	–	CNN	PCA
[128]	2018	patch classification; konwolucyjna sieć neuronowa (FCN)	ShareCNN; IndepCNN	RANSAC
[11]	2017	cechy czasowe	SVM	ważona transformata Hougha
[202]	2017	macierz Hessego	PCA; transformata Hougha	filtr Kalmana
[201]	2014	macierz Hessego filtracja; filtr liniowy Frangięgo	RANSAC	PCA; znormalizowana korelacja krzyżowa; filtr Kalmana
[129]	2014	adaptacyjna korekcja histogramu z ograniczonym kontrastem (CLAHE)	transformacje Gabora; RANSAC	zejście gradientowe

zana przy użyciu podejścia probabilistycznego do analizy obszaru otaczającego oś igły [29], przy użyciu funkcji fazy obrazu opartych na projekcji [53] lub za pomocą ważonego środka ciężkości na podstawie maski igły [85].

Porównanie wykorzystania metod bazujących na dwu- i trójwymiarowym obrazie ultrasonograficznym wskazuje, że szerzej opisane są techniki oparte na obrazach 2D, szczególnie w przypadku biopsji gruczołu sutkowego. Wynika to z częstszego wykorzystania głowic 2D w zabiegach biopsji oraz z łatwiejszej interpretacji położenia igły w obrazie dwuwymiarowym, choć w dalszym ciągu nie jest to prostym zadaniem.

Znacząca liczba prac opisujących automatyczne metody segmentacji igły zarówno w przypadku przetwarzania obrazów 2D jak i 3D, opiera się o dane fantomowe. W przypadku fantomów zwierzęcych najczęściej wykorzystywane są tkanki wieprzowe [4, 24, 29, 55, 111, 112, 181, 194], wołowe [2, 29, 31, 32, 109, 111, 112] i drobiowe [2, 4, 24, 55, 109, 112, 129], dla których zbierane są dane obrazowe o licznosciach rzędu kilku tysięcy obrazów (Tab. 2.9

i Tab. 2.10). Bazy z obrazami skanowanych fantomów przeważają zarówno pod względem powszechności jak i liczności nad bazami zawierającymi dane kliniczne. Tkanki ludzkie cechuje większa heterogeniczność z wieloma artefaktami, dlatego ważne jest dostarczanie metod detekcji igły biopsyjnej opartych o dane kliniczne, nawet jeśli ich liczba jest mniejsza. Niestety badania wykorzystujące ludzkie obrazy *in vivo* są zdecydowanie rzadsze i nie przetwarzają więcej niż tysiąca obrazów zebranych podczas zabiegów obrazowania od kilku do kilkunastu pacjentów. Przykładem prac opartych o dwuwymiarowe ultrasonograficzne obrazy kliniczne są metody Lee i wsp. [89] oraz Hatt i wsp. [55]. Wykorzystana liczba obrazów 2D w bazie kształtuje się od 557 [55] do 996 [89], gdzie dane te zebrane zostały podczas zabiegów wykonanych u kilku pacjentów (odpowiednio 6 i 8). W przypadku metod opartych o trójwymiarowe obrazowanie wykorzystywano od kilkunastu do około 150 wolumenów [4, 194], które zostały pozyskane w grupie kilkunastu pacjentów.

Opisane w literaturze automatyczne metody segmentacji igły i wyznaczania jej trajektorii skoncentrowane są na przetwarzaniu odpowiednio przygotowanych danych obrazowych, gdzie a priori odrzucone są obrazy, w których igła nie występuje [50], a klasyfikacja przeprowadzana jest jedynie na obrazach z widoczną igłą [201, 202]. W przypadku prac opartych na danych klinicznych przetwarzane były obrazy, które zostały ręcznie obrysowane przez eksperta [55, 89, 182].

Podczas zabiegu biopsji gruboigłowej dowolnego narządu tkanki są obrazowane nie tylko w trakcie wkłucia, ale również przed, jak i po usunięciu igły, w celu zlokalizowania i zminimalizowania ryzyka ewentualnych krwawień [89]. Istotne zatem wydają się prace nad automatyczną metodą klasyfikacji obecności igły w obrazie USG, aby późniejsza segmentacja realizowana była tylko na właściwych obrazach.

2.4 Systemy śledzenia

Systemy śledzenia wspomagające zabiegi chirurgiczne (ang. *Image-Guided Surgery*) rozwijane są od ponad 20 lat, podczas których systematycznie wzrasta ich dostępność oraz zakres wykorzystania. Są to urządzenia umożliwiające lokalizację położenia obiektów w przestrzeni trójwymiarowej i śledzenie ich w czasie rzeczywistym. Ze względu na ich zasadę działania, możliwe jest wskazanie kilku kategorii systemów: optyczne, elektromagnetyczne, ultradźwiękowe, mechaniczne oraz systemy bazujące na czujnikach inercyjnych lub żyroskopowych [132]. W zastosowaniach medycznych, ze względu na największą dokładność działania, stosowane są głównie systemy optyczne oraz elektromagnetyczne. W dalszej części rozprawy opisywane będą tylko te dwa rodzaje systemów śledzenia.

Tab. 2.9: Podsumowanie metod wykrywania igły w obrazach ultrasonograficznych 2D – podsumowanie baz danych oraz miar z otrzymanymi wartościami.

Ref.	Rok	Typ danych	Liczba obrazów	Miara	Wartość
[89]	2020	8 pacjentów	996	IoU DI $\Delta\theta_{rms}$ $\Delta\lambda_{rms}$	41,01% 56,65% 13,3° 9,5 pikseli
[79]	2020	2 fantomy wodne; 3 fantomy żelatynowe	8491	$RMSE$ kąta α $RMSE$ końcówki igły	2,29° 1,17 mm
[1]	2019	1 fantom tkankowy	ok. 160	$RMSE$	0,598 mm
[50]	2019	3 fantomy wodne	5958	$RMSE$ MAE	0,62×0,74 mm 6,08×7,62 pikseli 0,50×0,51 mm 4,90×5,24 pikseli
[110]	2019	1 fantom tkankowy	12000 500	błąd lokalizacji kończówki igły wskaźnik sukcesu	0,55±0,07mm 94%
[25]	2018	1 fantom tkankowy	1088	$\Delta\theta$	2° – 8°
[111]	2018	2 fantomy (tkanki wieprzowe i wołowe)	2900	$\Delta\lambda$ $\Delta\theta$	0,38±0,1 mm 0,82°±0,4°
[10]	2018	1 fantom tkankowy	60 filmów	wskaźnik sukcesu $RMSE$ kąta α $RMSE$ końcówki igły	100,00% 1,28° 0,82 mm
[181]	2018	8 fantomów (tkanki wieprzowe)	1440	odległość euklidesowa	17,60±18,44 pikseli
[24]	2015	2 fantomy (tkanki wieprzowe i drobiowe)	442	SEN SPE	84,6% 99,0%
[55]	2015	1 fantom (tkanki wieprzowe i drobiowe); 6 pacjentów	45 577	$\Delta\lambda$ $NLSR$ $\Delta\lambda$ $NLSR$	0,48 mm 86,2% 0,19 mm 99,80%

Tab. 2.10: Podsumowanie metod wykrywania igły w obrazach ultrasonograficznych 3D – podsumowanie baz danych oraz miar z otrzymanymi wartościami.

Ref.	Rok	Typ danych	Liczba obrazów	Miara	Wartość
[85]	2021	62 wklucia (różne tkanki)	—	błąd położenia końcówki igły	<1 mm, 1,5±0,9 mm 1,7±0,8 mm
[194]	2021	8 fantomów (serce wieprzowe)	92	<i>DI</i> <i>HD</i>	70,5%±9,2% 7,3±4,1 wokseli
		18 pacjentów	18	czas predykcji <i>DI</i> <i>HD</i>	4,1 s 66,0%±8.3% 8,2±2.9 wokseli
[4]	2019	17 pacjentów	149	czas predykcji czas detekcji pozycja orientacja	1,4 s 190 ms 1,59 mm 1,74°
[128]	2018	1 fantom (tkanki drobiowe i wieprzowe)	ok. 1000	<i>SEN</i> <i>PRE</i> <i>SPE</i> <i>SEN</i> <i>PRE</i> <i>SPE</i>	89,6±4,2 79,8±5,5 99,97±0,0001 87,9±4,2 83,0±3,7 99,99±0,00001
[202]	2017	20 fantomów	411800	<i>ACC</i> kąt α kąt β błąd końcówki czas detekcji	0,4±0,1 mm 0,0°±0,0° 0,1°±1,2° 0,6±0,3 mm 120,7±12,4 s
[11]	2017	20 zestawów	—	<i>ACC</i> $\Delta\theta$ Δp	81%±4% 2,12°±1,79° 1,69±1,54 mm
[201]	2014	55 fantomów (serce owcy)	788480	<i>ACC</i> <i>RMSE</i> igły <i>RMSE</i> kąta α <i>RMSE</i> kąta β czas detekcji	1,8 mm 1,6 mm 8° 5° 0,1 s
[129]	2014	1 fantom (tkanka drobiowa)	32	wskaźnik sukcesu błąd pozycji ϵ_p błąd orientacji ϵ_v	100% 0,65×0,9 mm, 2,2° i 3,5° odpowiednio dla igły 17G i 22G

Lokalizacja obiektów w przestrzeni za pomocą systemów śledzenia możliwa jest poprzez wykorzystanie znaczników (markerów) kompatybilnych z danego typu systemem. System śledzi bezpośrednio położenie znaczników. Określenie

położenia innego obiektu (np. narzędzia medycznego) wymaga umocowania znacznika do tego obiektu, a następnie przeprowadzenia procedury kalibracji [86, 169], która dla różnych obiektów przebiega w różny sposób. Precyzja wykonania kalibracji bezpośrednio przekłada się na dokładność działania systemu. Efektem poprawnie przeprowadzonej kalibracji jest wartość translacji i rotacji pomiędzy pewnym punktem obiektu a znacznikiem oraz informacja o dokładności pomiaru.

Systemy optyczne wykorzystują stereowizję, czyli technikę obrazową, za pomocą której możliwa jest rekonstrukcja scen trójwymiarowych. Rekonstrukcja sceny, a tym samym lokalizacja znacznika, realizowana jest za pomocą obrazów pochodzących z dwóch obiektywów kamery. Systemy optyczne charakteryzują się wysoką dokładnością śledzenia rzędu dziesiątych części milimetra. Zaletą systemu optycznego jest również rozległe pole widzenia, co pozwala na śledzenie położenia znaczników w relatywnie dużym obszarze oraz śledzenie niemal nieograniczonej liczby markerów, pod warunkiem zróżnicowania ich geometrii. Wadą systemu jest konieczność utrzymania znacznika bezpośrednio w polu widzenia kamery [117, 155].

Działanie systemów elektromagnetycznych oparte jest na wykorzystaniu pola elektromagnetycznego generowanego przez element systemu śledzenia, którym jest generator pola elektromagnetycznego (ang. *field generator*, FG). Pole elektromagnetyczne zmienia impedancję elektryczną elementów znacznika, a jej pomiar pozwala na lokalizację znacznika względem położenia generatora pola. Wykorzystywane pole elektromagnetyczne cechuje się niskim natężeniem, co pozwala na bezpieczne przenikanie przez tkanki ustrojowe. Zjawisko to umożliwia śledzenie położenia i orientacji znacznika zarówno na zewnątrz, jak i wewnątrz ciała ludzkiego. System elektromagnetyczny nie jest ograniczony koniecznością zachowania widoczności pomiędzy kamerą a markerem, jednak dokładność systemów elektromagnetycznych jest niższa w porównaniu do systemów optycznych [195]. Zaletami systemu elektromagnetycznego są niewielki rozmiar znaczników oraz możliwość ich przesłonięcia w trakcie badania. Wadą tego rozwiązania jest obecność przewodu łączącego marker z pozostałymi elementami systemu, ograniczona liczba znaczników, które można wykorzystać podczas zabiegu, a także wysoka wrażliwość systemu na elementy zakłócające pole elektromagnetyczne, takie jak przykładowo metale ferromagnetyczne, urządzenia elektroniczne oraz implanty ortopedyczne. Wytwarzane pole elektromagnetyczne wyklucza również użycie tego systemu u osób ze sztucznym rozrusznikiem serca ze względu na ryzyko zaburzenia jego pracy [116, 155].

Część eksperymentalna niniejszej rozprawy została zrealizowana z wykorzystaniem optycznego systemu śledzenia NDI Polaris Vicra oraz systemu magnetycznego NDI Aurora. Poniżej opisane są obydwie urządzenia.

2.4.1 Optyczny system śledzenia Polaris

System Polaris firmy Northern Digital Inc. (NDI) stanowi przykład optycznego systemu śledzenia (ang. *optical tracking system*, OTS), który wyznacza pozycję i orientację śledzonych znaczników w zdefiniowanej przestrzeni za pomocą trójwymiarowego układu współrzędnych. Elementem kluczowym optycznego systemu nawigacji jest kamera stereoskopowa, która zarówno emituje światło podczerwone o długości fali 800 nm, jak i nagrywa obrazy dwoma detektorami podczerwieni (Rys. 2.3).

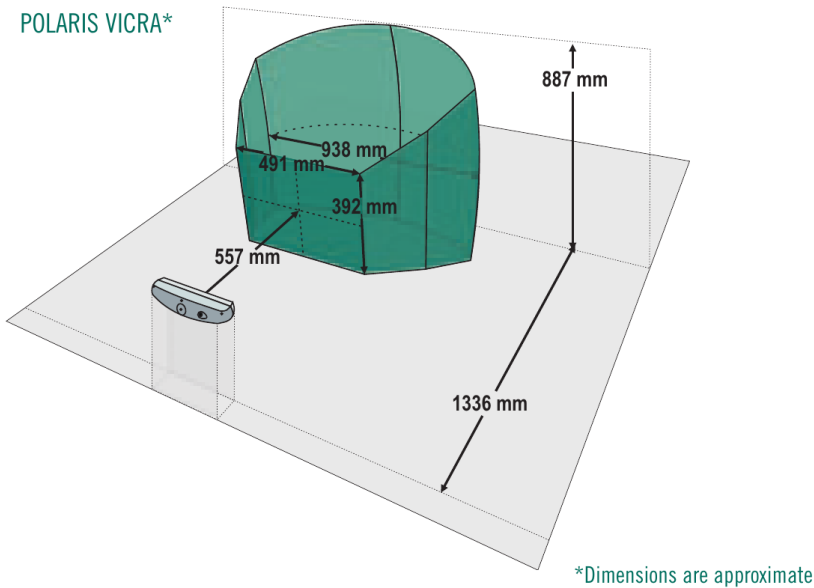


Rys. 2.3: Kamera stereoskopowa Polaris Vicra¹.

System Polaris działa śledząc pozycje znaczników w kształcie krzyża. Każdy ze znaczników składa się z umocowanych sztywno w określonej pozycji 4 kul o średnicy ok. 10 mm każda (Rys. 3.7). Znaczniki dzielą się na dwie grupy: pasywne i aktywne. Pierwsze pokryte są metalizowaną substancją o określonych właściwościach optycznych, umożliwiającą odbijanie światła z zakresu widzialnego i podczerwonego. Znaczniki aktywne są rzadziej stosowane, zamiast refleksyjnych powierzchni są w nich źródła światła podczerwonego. Lokalizacja markerów w przestrzeni trójwymiarowej realizowana jest na podstawie stereoskopowych obrazów różnicowych sceny bez oświetlenia i z oświetleniem o długości fali 800 nm. Efektem są wzmocnione obrazy stereoskopowe jasnych kół na czarnym tle, z pewnym przesunięciem uzależnionym od odległości znacznika od obiektywów. Obrazy można złożyć i na podstawie wielkości przesunięcia wyznaczyć wartość współrzędnych znacznika w przestrzeni trójwymiarowej [49].

Deklarowana przez producenta w nocie katalogowej dokładność śledzenia znaczników wynosi 0,25 mm [117] i została udokumentowana pracą [185]. Na dokładność systemu wpływa liczba sfer tworzących znacznik (im jest ich więcej, tym dokładność systemu jest wyższa [185]) oraz odległość od kamery. Pole widzenia kamery stereoskopowej optycznego systemu śledzenia przyjmuje kształt ściętego ostrosłupa z minimalnym zakresem śledzenia 557 mm i maksymalnym 1336 mm (Rys. 2.4).

¹ źródło: Northern Digital Inc. <https://www.ndigital.com/>



Rys. 2.4: Pole widzenia kamery stereoskopowej optycznego systemu śledzenia Polaris Vicra firmy NDI (źródło: [118]).

2.4.2 Elektromagnetyczny system śledzenia

System śledzenia Aurora firmy Northern Digital Inc. (NDI) stanowi przykład elektromagnetycznego systemu śledzenia (ang. *electromagnetic tracking system*, EMTS), który stosowany jest do wyznaczania położenia i orientacji znaczników w czasie rzeczywistym względem zdefiniowanego polem elektromagnetycznym trójwymiarowego układu współrzędnych. Kompletny system składa się z przenośnego generatora pola elektromagnetycznego (ang. *field generator*, FG), jednostki sterującej (ang. *system control unit*, SCU), portu czujników (ang. *sensor interface unit*, SIU), kompatybilnych czujników oraz komputera z dedykowanym oprogramowaniem (Rys. 2.5). Elementem kluczowym ETMS jest generator zmiennego pola elektromagnetycznego, które stanowi podstawę do określenia rotacji i translacji śledzonego znacznika [116].

Działanie systemu elektromagnetycznego Aurora oparte jest o cewki generujące silne zmienne pole elektromagnetyczne. Znacznik zawiera dwie cewki elektryczne, w których następuje indukcja napięcia proporcjonalna do natężenia pola, które na tę cewkę oddziałuje. Mierząc indukowane napięcie każdej cewki, możliwe jest wskazanie jej orientacji i położenia w przestrzeni pomiarowej [195] względem generatora. Za pomocą EMTS można wskazać trzy pozycje (x, y, z) oraz dwa lub trzy kąty orientacji [141] dla maksymalnie 8 znaczników

w tym samym czasie.

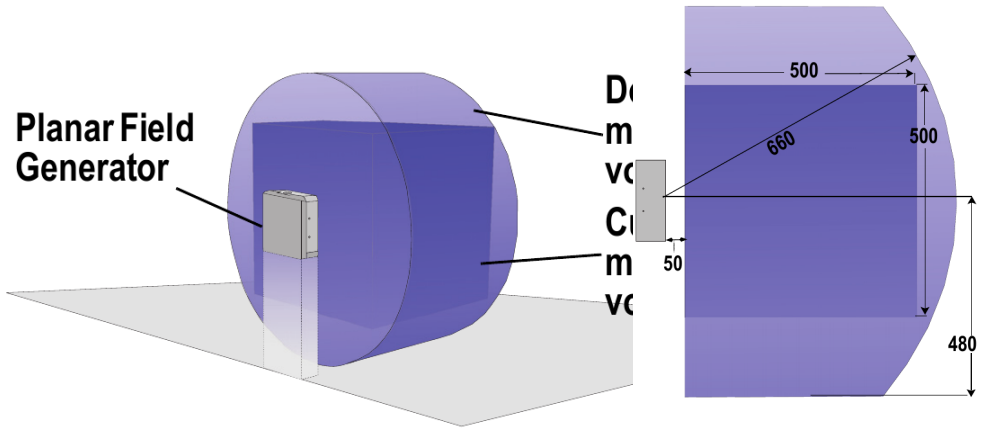


Rys. 2.5: System elektromagnetyczny Aurora firmy NDI (źródło: [116]).

Przestrzenny obszar śledzenia (przestrzeń pomiarowa) może przyjmować dwa kształty: sześcianu albo kopuły (Rys. 2.6). Sześcienne przestrzeń pomiarowa ma objętość roboczą o wymiarach $500 \times 500 \times 500$ mm. Dokładność śledzenia deklarowana przez producenta w tym obszarze wynosi 0,48 mm. W przypadku śledzenia w obszarze kopuły obszar roboczy jest większy względem sześciennego, co widoczne jest na rysunku 2.7. Dokładność śledzenia dla tego przypadku jest niższa i wynosi 0,70 mm [116].

2.5 Podsumowanie

W niniejszym rozdziale opisano stan wiedzy z zakresu przetwarzania i analizy obrazów USG oraz działania systemów nawigacji stosowanych w zabiegach medycznych. Przedstawiony został charakter prezentowanych w literaturze prac, które poruszają zagadnienia filtracji obrazów USG, poddają pod dyskusję metody detekcji oraz segmentacji struktur – zarówno anatomicznych jak i narzędzi chirurgicznych – w obrazach ultrasonograficznych, oraz dokonują analizy metod klasyfikacji zmian ogniskowych.



Rys. 2.6: Kształt obszaru śledzenia systemu Aurora (źródło: [116]).

Rys. 2.7: Rzut boczny dwóch typów obszarów śledzenia systemu Aurora. Wartości wyrażone w mm (źródło: [116]).

W niniejszym rozdziale poddano analizie dane, które są prezentowane w obecnych pracach oraz zwrócono uwagę na ogólnodostępne bazy obrazów oraz ich licznosc. W przypadku detekcji i segmentacji zmian ogniskowych na obrazach gruczołu sutkowego, proces uczenia i testowania technik przetwarzania realizowany jest najczęściej w oparciu o dane kliniczne, a sporadycznie o dane fantomowe. Odwrotna zależność obserwowana jest w ocenie narzędzi dedykowanych detekcji obecności igły. W tym obszarze powszechne jest wykorzystanie danych fantomowych.

Przegląd istniejących rozwiązań uwzględnia również opis systemów śledzenia stosowanych w zabiegach wspieranych nawigacją obrazową ze wskazaniem ich działania wraz z deklarowanymi przez producentów dokładnościami.

3. System rejestracji danych

W niniejszym rozdziale opisano elementy składowe systemu rejestracji danych i przedstawiono kolejne etapy ich kalibracji. Dalsza część rozdziału poświęcona jest opisowi integracji optycznego i elektromagnetycznego systemu śledzenia z obrazem ultrasonograficznym.

3.1 Konstrukcja stanowiska pomiarowego

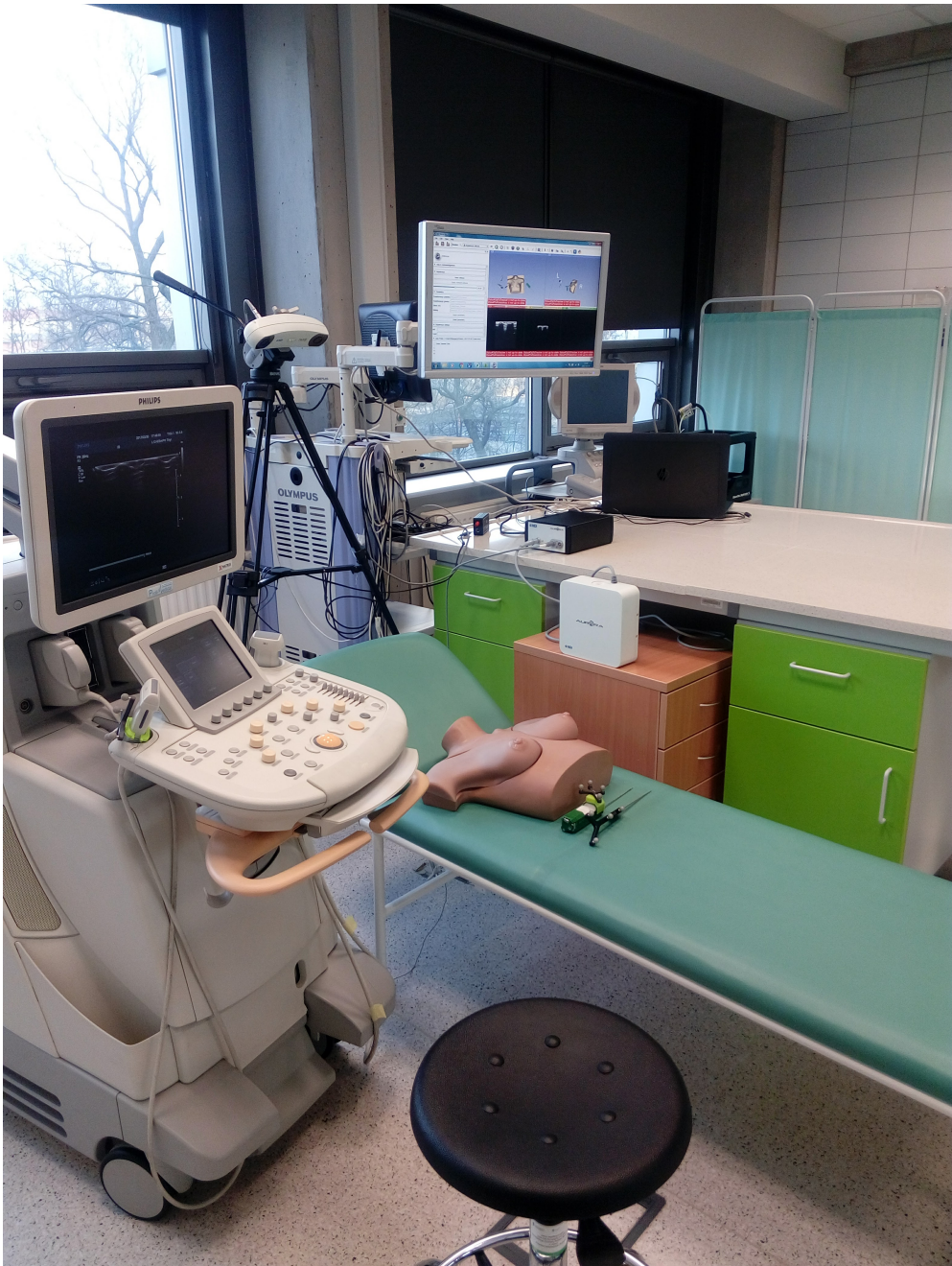
Stanowisko kliniczne wykorzystywane przez lekarza radiologa podczas zabiegu biopsji gruboigłowej realizowanej pod kontrolą ultrasonograficzną zawiera następujące elementy:

- aparat ultrasonograficzny,
- liniowa głowica ultrasonograficzna,
- pistolet biopsyjny.

Komputerowe śledzenie pistoletu biopsyjnego i głowicy ultrasonograficznej rozszerza możliwości interwencji igłowych względem biopsji pod kontrolą samego obrazowania ultrasonograficznego.

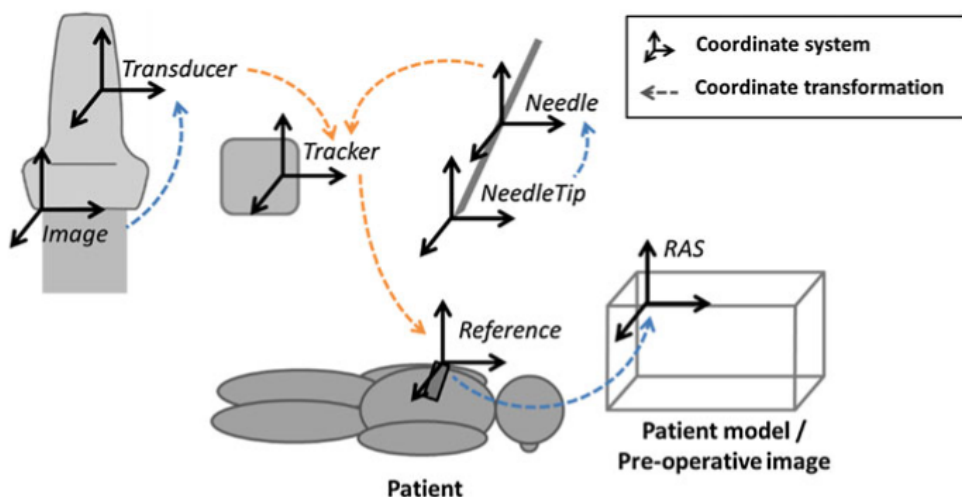
W celu umożliwienia akwizycji różnych modalności danych podczas zabiegu biopsji gruboigłowej, stanowisko kliniczne zostało rozszerzone o elementy śledzące położenie w przestrzeni narzędzi zabiegowych oraz głowicy ultrasonograficznej (Rys. 3.1). W zaproponowanym systemie rejestracji danych pochodzących z zabiegu biopsji, stanowisko kliniczne zostało rozbudowane o następujące elementy do stanowiska pomiarowego:

- system śledzenia,
- kompatybilne markery (znaczniki),
- konwerter sygnału video (ang. *frame grabber*),
- komputer z monitorem oraz oprogramowaniem.



Rys. 3.1: Stanowisko pomiarowe z opisywanym systemem rejestracji danych.

Zastosowanie systemu śledzenia umożliwia rejestrację w przestrzeni trójwymiarowej położenia głowicy ultrasonograficznej w oparciu o kompatybilny



Rys. 3.2: Układy współrzędnych obiektów i transformacje pomiędzy obiektami a znacznikami zapisywane za pomocą systemu śledzenia podczas biopsji gruboigłowej pod kontrolą ultrasonograficzną (źródło: [169]).

znacznik umieszczony na jej powierzchni. Obraz pochodzący z monitora ultrasonografu przekazywany jest złączem USB 2.0 komputera za pomocą konwertera obrazu, który jest podłączony do złącza standardu S-Video lub VGA, w zależności od wyposażenia aparatu USG. Zastosowany w niniejszej pracy protokół kalibracji głowicy umożliwia wyznaczenie przekształcenia afinicznego obrazu USG względem znacznika umieszczonego nieruchomo na powierzchni głowicy (Rys. 3.10), natomiast procedura kalibracji wskaźnika pozwala na wyznaczenie końcówki igły pistoletu biopsyjnego, które to elementy są śledzone z użyciem prezentowanego systemu rejestracji danych.

Monitorowanie i akwizycja położenia głowicy liniowej oraz pistoletu biopsyjnego jest efektem zastosowania dwóch systemów śledzenia:

- optycznego systemu śledzenia (ang. *optical tracking system*, OTS) Polaris Vicra firmy Northern Digital Inc. (NDI) [117],
- elektromagnetycznego systemu śledzenia (ang. *electromagnetic tracking system*, EMTS) Aurora firmy NDI [116].

Niniejsze stanowisko pomiarowe zostało opracowane we współpracy z lekarzami radiologami, celem zapewnienia im ergonomii i swobody podczas wykonywania zabiegu. Zdjęcie systemu rejestracji danych przedstawiono na rysunku 3.1.

Statyw z kamerą stereoskopową został umieszczony za głową pacjentki zgodnie z polem widzenia systemu optycznego, co zostało przedstawione na

rysunku 2.4. Kamerę skierowano względem pacjentki tak, aby w środkowej części monitorowanego obszaru znajdowała się jej klatka piersiowa, w celu monitorowania pozycji pistoletu biopsyjnego. Ze względu na ograniczony obszar śledzenia EMTS, generator pola elektromagnetycznego umieszczono po lewej stronie pacjentki. Lekarz wraz z aparatem ultrasonograficznym jest po prawej stronie pacjentki. Umieszczenie generatora pola po przeciwnej stronie niweluje kolizję ręki lekarza z elementami systemu rejestracji oraz pozwala na obrazowanie interesującego obszaru (Rys. 2.6).

System optyczny cechuje się wyższą dokładnością w porównaniu do systemu elektromagnetycznego, co zostało przedstawione w tabeli 3.1. W obu przypadkach rzeczywista dokładność śledzenia narzędzi uzależniona jest od dokładności wykonywanych kroków protokołu kalibracji, przedstawionego dalej w podrozdziale 3.3.

Zaletą EMTS są znacznie mniejsze wymiary znaczników (Rys. 3.6), co znacząco wpływa na komfort pracy lekarza i nie ogranicza jego ruchów. Każdy z markerów systemu elektromagnetycznego jest połączony delikatnym przewodem z interfejsem, przez co wymagane jest zachowanie ostrożności podczas użytkowania.

OTS jest narzędziem pracującym w świetle podczerwonym, którego obicie przez śledzony marker pozwala na określenie jego położenia w przestrzeni trójwymiarowej. Umieszczenie jakiegokolwiek obiektu (przedmiotu lub na przykład ręki operatora) pomiędzy kamerą a znacznikiem lub odwrócenie znacznika tyłem względem kamery uniemożliwia jego lokalizację. W przypadku EMTS zasłonięcie markera ręką nie wpływa na zapis pozycji śledzonego czujnika, co w mniejszym stopniu oddziałuje na płynność ruchów lekarza wykonującego zabieg. Oba systemy cechują się ograniczonym rozmiarem śledzonej przestrzeni. W przypadku OTS obszar ten jest większy.

Istotnym ograniczeniem systemów EMTS jest zakaz wykorzystywania ich w przypadku pacjentów z wszczepionym kardiostymulatorem lub paramagnetycznymi implantami. W przypadku rozrusznika serca, system elektromagnetyczny może zakłócić jego pracę, natomiast w przypadku implantów, mogą wystąpić artefakty w pozyskiwanych danych, z powodu obecności metalowych elementów. System optyczny nie ma ograniczeń związanych ze stanem zdrowia pacjenta. Najważniejsze cechy różniące system optyczny Polaris i elektromagnetyczny Aurora przedstawiono w tabeli 3.1.

Wady i zalety każdego z systemów śledzenia zdecydowały o wykorzystaniu systemu EMTS w monitorowaniu położenia głowicy oraz systemu OTS do śledzenia pistoletu biopsyjnego. Lekarz radiolog przeprowadzając obrazowanie gruczołu sutkowego obraca głowicą ultrasonograficzną w różnych kierunkach, co nie pozwala na ciągle monitorowanie jej kamerą stereoskopową. Głowica połączona jest przewodem z aparatem ultrasonograficznym, stąd monitorowanie jej położenia za pomocą znacznika wyposażonego w przewód zasilający nie sta-

Tab. 3.1: Zestawienie systemów śledzenia.

	System OTS (NDI Polaris)	System EMTS (NDI Aurora)
Dokładność lokalizacji	0,25 mm	0,70 mm
Obszar śledzenia	1336×887×938 mm	500×500×500 mm
Markery	kształt krzyża o wymiarach ok. 5 cm	walce (0,8×2 mm)
Maksymalna liczba markerów	nieograniczona	8
Aparatura	stereokamera	generator pola, port markerów, jednostka kontrolna systemu
Ograniczenia	utrzymanie widoczności znacznika w polu widzenia kamery	błędy działania w otoczeniu metalowych elementów i urządzeń elektronicznych w odległości 1 m; nie można stosować u osób z rozrusznikiem serca, implantami
Komfort użytkowania	duże znaczniki utrudniają korzystanie	przewody znaczników są bardzo wrażliwe na urwanie

nowi znaczącego ograniczenia dla radiologa. Zastosowanie natomiast czujnika EMTS (ograniczonego przewodem) do pistoletu biopsyjnego ograniczałoby zespół medyczny ze względu na konieczność przekazania przez lekarza narzędzia pomiędzy kolejnymi biopsjami pielęgniarsce zabezpieczającej materiał biopsyjny oraz przygotowującej pistolet do kolejnej iniekcji. Z drugiej strony, podczas pobierania wycinków, lekarz wprowadza igłę z zachowaniem określonego położenia pistoletu, przez co możliwe jest wcześniejsze odpowiednie ustawienie kamery optycznej, która może śledzić znacznik przez cały czas procedury biopsji.

3.2 Stanowisko pomiarowe – integracja akwizycji danych

Systemy nawigacji stosowane w warunkach klinicznych celem wsparcia zabiegów medycznych są wdrażane w różnorodnych warunkach sprzętowych ga-

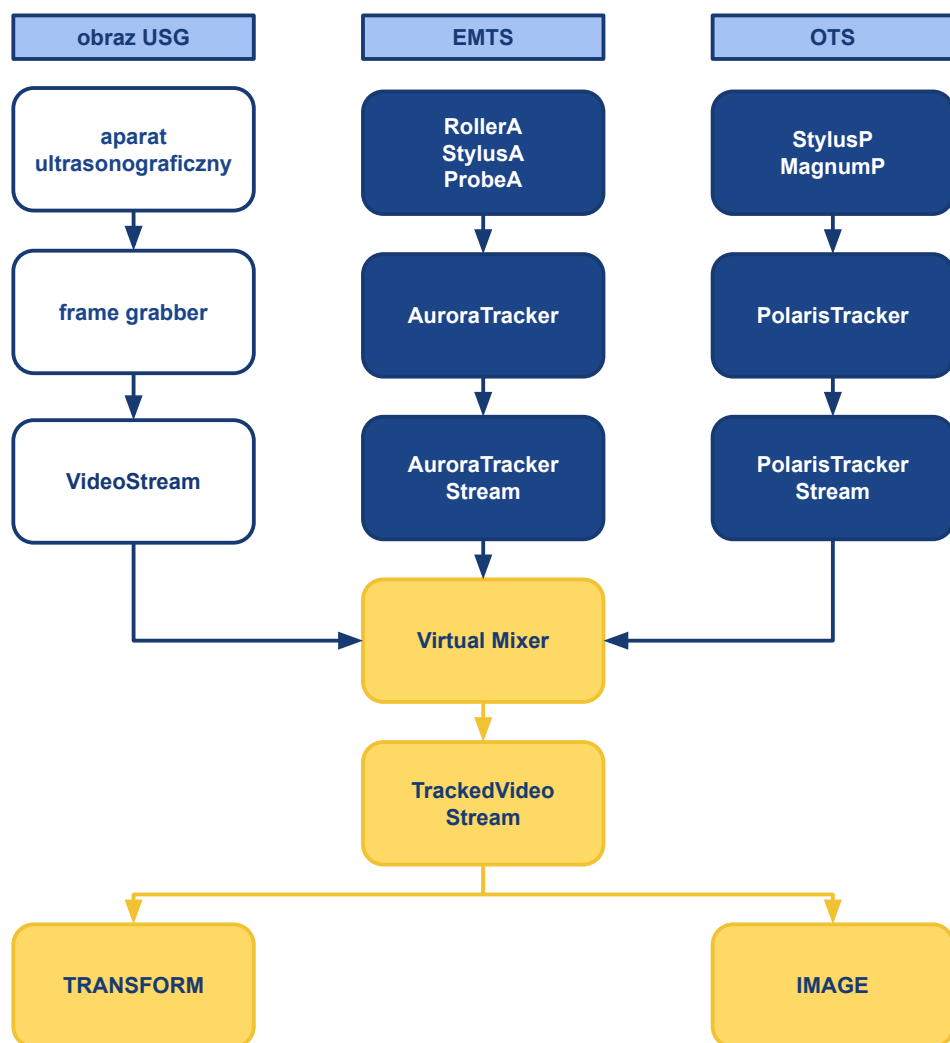
binetu zabiegowego. Do akwizycji danych (pochodzących z różnych źródeł), ich przetwarzania oraz wizualizacji zastosowano bibliotekę PLUS Toolkit (ang. *Public software Library for UltraSound imaging research*). Biblioteka PLUS jest pakietem oprogramowania umożliwiającym rejestrację dwuwymiarowych obrazów USG wraz informacją o jej położeniu w przestrzeni, kalibrację wskaźnika, kalibrację czasową i przestrzenną głowicy ultrasonograficznej, a także przetwarzanie i wizualizację zapisanych danych w czasie rzeczywistym [86] za pomocą zewnętrznego oprogramowania takiego jak 3D Slicer, z którym komunikacja jest realizowana poprzez protokół OpenIGTLink [165].

Jednym z narzędzi biblioteki PLUS jest aplikacja PlusServer umożliwiająca uruchamianie poszczególnych modułów tej biblioteki i przesyłanie danych zbieranych za pomocą systemów śledzenia (i innych urządzeń) do dowolnej liczby aplikacji klienckich [86]. Typowe zastosowanie programu PlusServer to akwizycja obrazów ultrasonograficznych wraz z ich lokalizacją oraz ich przesyłanie do programów wizualizacyjnych. Przesyłanie danych w czasie rzeczywistym realizowane jest poprzez protokół OpenIGTLink. Jest to szeroko stosowany otwarty protokół komunikacyjny oparty o gniazda TCP/IP zaprojektowany w celu obsługi komunikacji pomiędzy urządzeniami medycznymi wspomagającymi nawigację obrazową w trakcie zabiegów [165].

Klient połączony z aplikacją PlusServer ma możliwość wysyłania komend do urządzeń akwizycji danych za pomocą protokołu OpenIGTLink, na przykład w celu rozpoczęcia i zakończenia zapisu danych, rekonstrukcji wolumenu z dwuwymiarowych obrazów ultrasonograficznych, przetwarzania obrazów albo dodawania nowych funkcjonalności. Możliwość ta została wykorzystana w implementacji oprogramowania w środowisku Python 2.7.11, w celu automatycznego rozpoczęcia i zakończenia zapisu danych wykorzystującego unikatowy identyfikator pacjenta.

Do skonfigurowania aplikacji PlusServer konieczne jest zdefiniowanie urządzeń wejściowych, protokołu przetwarzania oraz danych wyjściowych. Wszystkie te elementy podawane są w plikach konfiguracyjnych, w których umieszcza się także informację o systemach śledzenia, wykorzystywanych czujnikach czy zastosowanym konwerterze sygnału video. W tym samym pliku definiowany jest też sygnał wyjściowy, którym są połączone strumienie video obrazu ultrasonograficznego i dane o położeniu śledzonych narzędzi. Poprawny plik konfiguracyjny musi uwzględniać otrzymane macierze rotacji i translacji wyznaczone podczas kalibracji [169]. Elementy pliku konfiguracyjnego pozwalającego na akwizycję danych za pomocą dwóch systemów przedstawiono na rysunku 3.3, natomiast fragment przykładowego pliku konfiguracyjnego przedstawiono na grafice 3.4.

Dane wyjściowe zapisywane są w formacie MetaImage (pliki o rozszerzeniu MHA) zdefiniowanym przez Insight Segmentation and Registration Toolkit. Format MetaImage zawiera metadane niezbędne do analizy obrazów ultra-



Rys. 3.3: Elementy pliku konfiguracyjnego aplikacji PlusServer.

sonograficznych, takie jak rozdzielczość, offset, kierunki osi obrazu, znacznik czasowy czy status poprawności zebranych danych.

3.3 Zintegrowany system śledzenia narzędzi i obrazu USG

W typowym diagnostycznym badaniu ultrasonograficznym nie korzysta się z żadnych zewnętrznych układów współrzędnych, a lokalizacja obrazu może być jedynie subiektywnie opisywana poprzez adnotacje umieszczane na tym obrazie. Na podstawie samego obrazu nie ma możliwości odtworzenia jego po-

```

1 <PlusConfiguration version="2.1" PlusRevision="Plus-2.2.0.4177 - Win64">
2 <DataCollection StartupDelaySec="1">
3 <DeviceSet Name="aurora_polaris 25 05 2018" Description="aurora_polaris 25 05 2018" />
4 <Device
5 MeasurementVolumeNumber="2"
6 Id="Tracker"
7 Type="AuroraTracker"
8 SerialPort="4"
9 BaudRate="115200"
10 AcquisitionRate="40"
11 LocalTimeOffsetSec="0"
12 ToolReferenceFrame="Tracker">
13 <DataSources>
14 <DataSource Type="Tool" Id="RollerA" PortName="0" BufferSize="150" />
15 <DataSource Type="Tool" Id="ProbeConvexA" PortName="1" BufferSize="150" />
16 <DataSource Type="Tool" Id="StylusA" PortName="3" BufferSize="150" />
17 </DataSources>
18 <OutputChannels>
19 <OutputChannel Id="TrackerStream">
20 <DataSource Id="RollerA" />
21 <DataSource Id="ProbeConvexA" />
22 <DataSource Id="StylusA" />
23 </OutputChannel>
24 </OutputChannels>
25 </Device>
26 <Device
27 Id="Polaristracker"
28 Type="PolarisTracker"
29 SerialPort="3"
30 BaudRate="115200"
31 AcquisitionRate="40"
32 LocalTimeOffsetSec="0"
33 ToolReferenceFrame="Polaristracker"
34 MeasurementVolumeNumber="0">
35 <DataSources>
36 <DataSource Type="Tool" Id="MagnumP" RomFile="8700338.rom" BufferSize="150" />
37 <DataSource Type="Tool" Id="ProbelinearP" RomFile="8700339.rom" BufferSize="150" />
38 <DataSource Type="Tool" Id="PhantomP" RomFile="8700449.rom" BufferSize="150" />
39 <DataSource Type="Tool" Id="StylusP" RomFile="8700340.rom" BufferSize="150" />
40 </DataSources>
41 </OutputChannels>

```

Rys. 3.4: Plik konfiguracyjny aplikacji PlusServer.

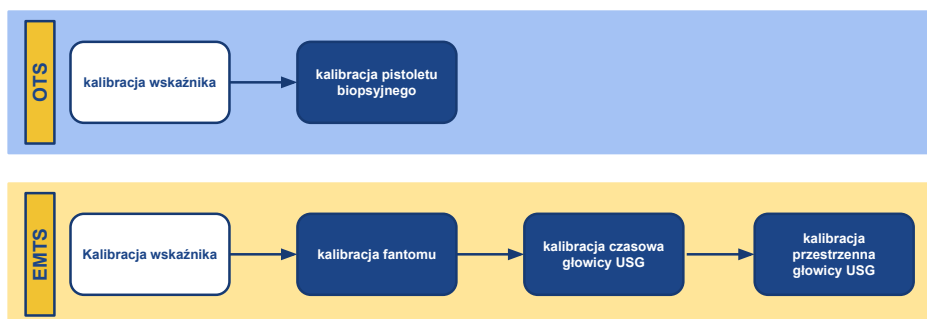
łożenia w przestrzeni ani ustalenia zależności przestrzennej pomiędzy kolejnymi zachowanymi ujęciami.

Wykorzystanie systemu śledzenia w badaniu USG pozwala na przypisanie każdemu obrazowi położenia w stałym zewnętrznym układzie odniesienia. Dzięki temu możliwe jest wykonanie rekonstrukcji trójwymiarowego wolumenu z sekwencji obrazów [107], przez co można na przykład wyznaczać objętości narzędzi lub zmian ogniskowych [103].

Monitorowanie położenia narzędzi wykorzystywanych w zabiegu biopsji gruczołowej za pomocą systemów śledzenia wymaga przeprowadzenia czterech etapów kalibracji (Rys. 3.5). Proces ten uwzględnia kalibrację wskaźnika i pistoletu biopsyjnego oraz czasowo-przestrzenną kalibrację liniowej głowicy ultrasonograficznej. Protokół kalibracyjny jest identyczny dla obu wykorzystywanych systemów śledzenia.

3.3.1 Działanie systemu śledzenia

Położenie znacznika w przestrzeni definiowane jest za pomocą macierzy translacji i rotacji w przestrzeni trójwymiarowej. Wartości macierzy wyznaczone są w oparciu o detekcję znacznika przez system śledzenia [170]. W systemie optycznym za detekcję markera odpowiedzialne są dwa obiektywy kamery stereowizyjnej, a w systemie elektromagnetycznym detekcja opiera się o pomiar zmian impedancji markera w polu elektromagnetycznym. W obu przypadkach określana jest lokalizacja śledzonego przez dany system znacznika. Aby możliwe było śledzenie innego obiektu, który nie jest znacznikiem (np. pi-



Rys. 3.5: Etapy kalibracji końcówki igły pistoletu biopsyjnego oraz kalibracji przestrzennej i czasowej głowicy biopsyjnej za pomocą systemu odpowiednio optycznego i elektromagnetycznego.



Rys. 3.6: Porównanie znaczników systemów elektromagnetycznego Aurora (lewa strona) i optycznego Polaris (prawa strona).

stoletu biopsyjnego lub głowicy ultrasonograficznej), niezbędnym krokiem jest sztywne umocowanie znacznika do jego powierzchni, a następnie wyznaczenie przekształcenia geometrycznego pomiędzy znacznikiem a lokalnym układem współrzędnych tego obiektu. Wszystkie przekształcenia geometryczne pozwala-

jące na umieszczenie elementów w jednym, globalnym układzie współrzędnych, są transformacjami izometrycznymi, przechowywanymi w macierzy translacji i rotacji o wymiarach 4×4 [131].

Przyjmując oznaczenie, zgodnie z którym P^X stanowi współrzędne punktu P w układzie współrzędnych X , natomiast P^Y to współrzędne punktu P w układzie Y , możliwe jest przeniesienie P^Y do układu X poprzez zastosowanie macierzy transformacji M_Y^X , zgodnie z zależnością:

$$P^X = M_Y^X \cdot P^Y. \quad (3.1)$$

Gdy dany jest obiekt z przypisanym do niego lokalnym układem współrzędnych O , wówczas, na podstawie informacji o zależności między układem współrzędnych O a układem współrzędnych znacznika L oraz o położeniu znacznika w globalnym układzie współrzędnych G , możliwe jest wyznaczenie przekształcenia umieszczającego ten obiekt w globalnym układzie współrzędnych systemu śledzenia G .

Przyjmując, że znane są współrzędne punktu P w lokalnym układzie współrzędnych obiektu O , przekształcenia geometryczne transformujące wartości współrzędnych punktu P^O do globalnego układu współrzędnych G są realizowane zgodnie z zapisem:

$$P^G = M_L^G \cdot M_O^L \cdot P^O, \quad (3.2)$$

gdzie M_O^L jest stałą macierzą kalibracji obiektu, a M_L^G jest zmieniającą się w czasie macierzą transformacji, którą generuje bezpośrednio system śledzenia (innymi słowy: jest to położenie znacznika L w globalnym układzie współrzędnych).

Wyznaczanie macierzy transformacji obarczone jest każdorazowo błędem ze względu na ograniczoną dokładność lokalizacji znaczników oraz błędy wynikające z kolejnych etapów kalibracji (wyznaczania kolejnych macierzy transformacji). Błąd kalibracji jednego etapu skutkuje narastającymi niedokładnościami w kolejnych.

3.3.2 Kalibracja wskaźnika w systemach śledzenia

Wskaźnikiem (ang. *stylus*) nazywane jest sztywne szpiczaste narzędzie z nieruchomo zamocowanym znacznikiem odpowiedniego systemu śledzenia: optycznego lub elektromagnetycznego. Za pomocą końcówki wskaźnika (ang. *tip*) – niezależnie od typu systemu śledzenia – wskazywane są punkty charakterystyczne biorące udział w procesie kalibracji. Oba systemy wyposażone są przez producentów w standardowe wskaźniki. Wskaźnik systemu optycznego został przedstawiony na rysunku 3.7, natomiast wskaźnik systemu elektromagnetycznego na rysunku 3.8. Wskaźnik optyczny zbudowany jest z czterech kul o śred-

Tab. 3.2:

[117]]Współrzędne sfer wskaźnika w lokalnym układzie współrzędnych [mm] [117].

nazwa	x	y	z
Marker A	0,00	0,00	0,00
Marker B	0,00	0,00	50,00
Marker C	0,00	25,00	100,00
Marker D	0,00	-25,00	135,00

nicy ok. 1 cm rozmieszczonych niesymetrycznie na ramionach krzyża. Związany z nim układ współrzędnych pokrywa się z punktem mocowania jednej z kul do krzyża (Tab. 3.2). Wskaźnik elektromagnetyczny ma swój lokalny układ współrzędnych wewnątrz cewek.

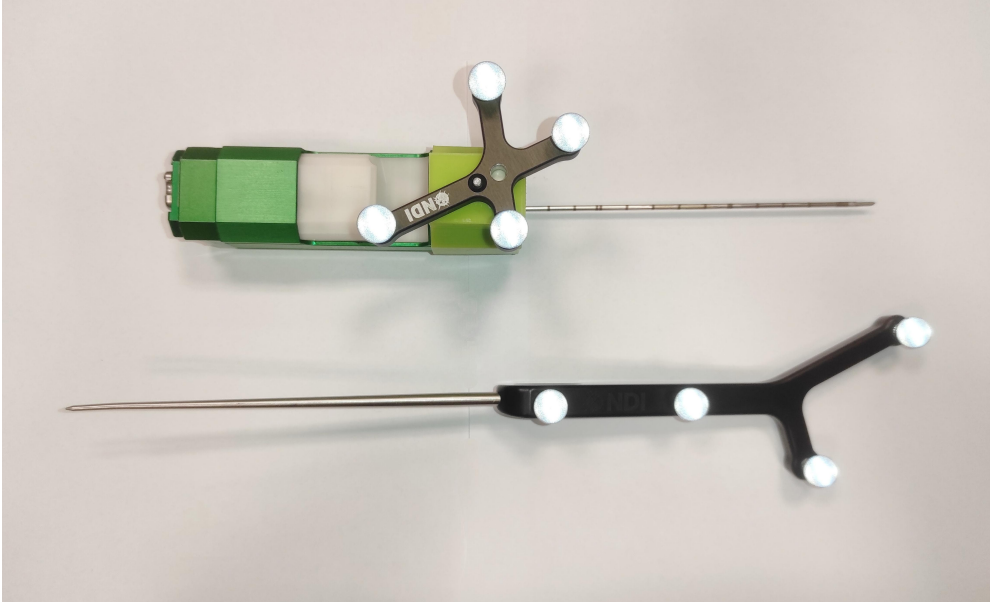
Kalibracja wskaźnika w odpowiadającym mu systemie śledzenia wykonywana jest w celu określenia stałej macierzy transformacji pomiędzy układem współrzędnych śledzonego znacznika S a położeniem jego końcówki P w układzie K . Położenie to (P^K) nie ulega zmianie podczas poruszania wskaźnika. Obliczenie transformacji końcówki wskaźnika do znacznika wskaźnika zwykle wykonywane jest na podstawie kalibracji, w trakcie której końcówka jest ustabilizowana, a cały wskaźnik wykonuje ruchy wodzące po sferze [169]. Śledzony znacznik obracany jest wokół końcówki przez kilka sekund, w trakcie których zapisywana jest wartość współrzędnych znacznika. Na tej podstawie obliczana jest macierz transformacji M_K^S minimalizująca przemieszczenie końcówki znacznika i umożliwiająca przeniesienie początku układu współrzędnych końcówki K do układu znacznika S .

Wynikiem kalibracji jest macierz transformacji końcówki do lokalnego układu współrzędnych wskaźnika M_K^S , natomiast finalne położenie końcówki wskaźnika w globalnym układzie współrzędnych G wyznaczone jest zgodnie z zależnością:

$$P^G = M_S^G \cdot M_K^S \cdot P^K. \quad (3.3)$$

3.3.3 Kalibracja pistoletu biopsyjnego

Pistolet biopsyjny Bard Magnum MG1522 – wykorzystywany w niniejszej pracy – wyposażony jest w wymienną igłę biopsyjną mocowaną za pomocą dwóch elementów blokujących. Na powierzchni pistoletu, poprzez nakładkę zaprojektowaną własnoręcznie z użyciem oprogramowania graficznego Blender i wydrukowaną z tworzywa ABS drukarką 3D Makerbot, umieszczony został znacznik optyczny zbudowany z 4 kul, których współrzędne zostały zebrane



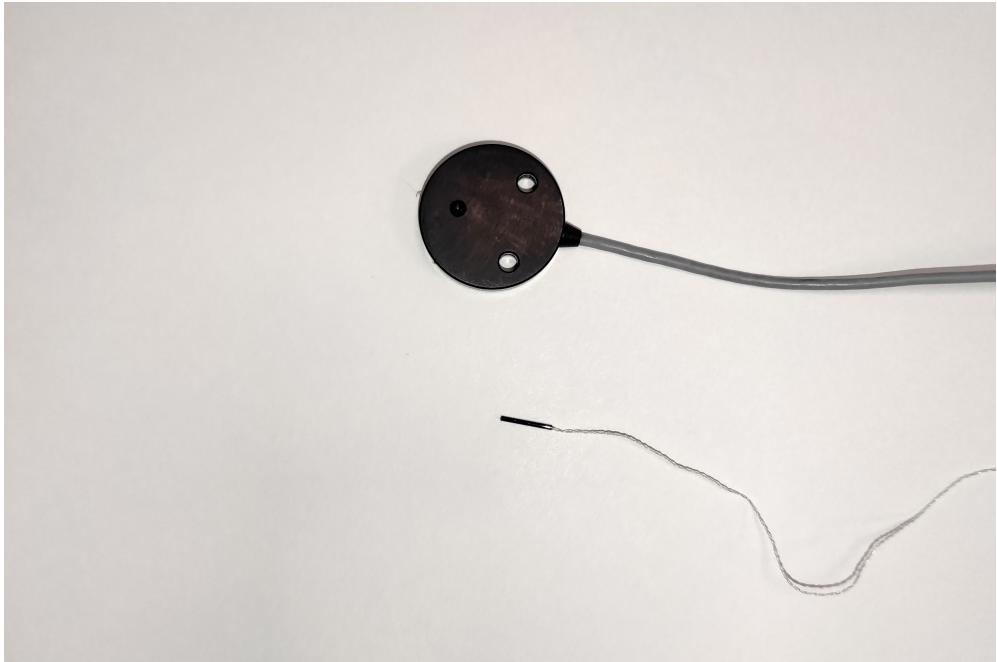
Rys. 3.7: Znaczniki optycznego systemu Polaris.

w tabeli 3.3. Kompletne narzędzie przypomina wskaźnik (Rys. 3.9), zatem możliwa jest jego kalibracja zgodnie z protokołem kalibracji wskaźnika opisanym w poprzednim podrozdziale. Proces kalibracji zakłada akwizycję szeregu lokalizacji pozwalających na transformację końcówki igły P znajdującej się w układzie współrzędnych K_M do lokalnego układu współrzędnych znacznika umieszczonego na powierzchni pistoletu M . Następnie układ przekształcany jest w celu umieszczenia końcówki igły w globalnym układzie współrzędnych G kamery systemu Polaris. Kolejne przekształcenia prowadzące do wyznaczenia położenia końcówki igły w globalnym układzie współrzędnych są następujące:

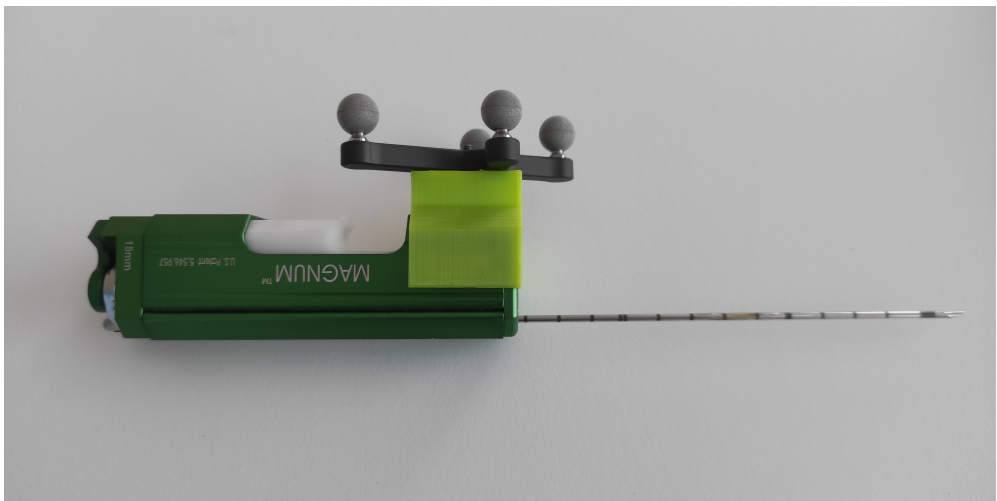
$$P^G = M_M^G \cdot M_{K_M}^M \cdot P^{K_M}. \quad (3.4)$$

Eksperymenty wykazały, że powyższe podejście obarczone jest dużymi błędami wyznaczenia $M_{K_M}^M$ ze względu na elastyczność igły biopsyjnej, która pod wpływem obciążenia pistoletem podczas zataczania ruchów po sferze w znacznym stopniu się wygina. Z tego powodu zmodyfikowano protokół kalibracji w celu eliminacji wpływu niedokładności detekcji końcówki igły biopsyjnej na dokładność działania całego systemu.

Modyfikacja protokołu kalibracji pistoletu biopsyjnego uwzględnia wykorzystanie uprzednio skalibrowanego wskaźnika, który traktowany jest jako narzędzie referencyjne. Końcówka wskaźnika umieszczona jest dokładnie na końcówce igły biopsyjnej. Następnie wykonywane są ruchy wskaźnikiem po sferze



Rys. 3.8: Znaczniki optycznego systemu Aurora.



Rys. 3.9: Pistolet Magnum z igłą przeznaczony do zabiegu biopsji gruboigłowej z umieszczonym znacznikiem optycznym.

bez przemieszczenia końcówki wskaźnika. Ponieważ układ współrzędnych obracającego się wskaźnika jest układem referencyjnym dla znacznika na pistolecie biopsyjnym, to można przyjąć, że – względem układu współrzędnych kamery

Tab. 3.3: Współrzędne sfer markera umieszczonego na powierzchni pistoletu biopsyjnego w lokalnym układzie współrzędnych [mm] (źródło: [117]).

nazwa	x	y	z
Marker A	0,00	0,00	0,00
Marker B	0,00	25,28	16,15
Marker C	0,00	0,00	65,00
Marker D	0,00	-18,65	29,61

– to pozycja znacznika na pistolecie zmienia swoje położenie. W efekcie możliwe jest wyznaczenie macierzy rotacji i translacji pomiędzy układem lokalnym końcówki igły a znacznikiem umieszczonym na powierzchni pistoletu biopsyjnego. Wynikiem są macierze transformacji pozwalające na kolejne przekształcenia zgodnie z zapisem 3.4. Błąd tak wykonanej kalibracji igły jest znacząco mniejszy niż błąd w klasycznej procedurze kalibracji nieangażującej dodatkowe narzędzie [25] (Rozdział 5.1.3).

3.3.4 Kalibracja czasowa

Głowica ultrasonograficzna skanująca narządy i tkanki podczas badania oraz system śledzenia pozycji tej głowicy przekazują do komputera informacje za pomocą różnych protokołów komunikacji. Obraz USG przekazywany jest wyjściem analogowym video aparatu USG do urządzenia konwertującego, zwanego też frame grabberem. Konwerter podłączony do gniazda USB komputera przekazuje dane z częstotliwością 30 Hz.

Położenie głowicy określane jest za pomocą znacznika śledzonego systemem elektromagnetycznym, skąd informacja przekazywana jest do innego portu USB komputera z częstotliwością do 40 Hz [116]. Każde urządzenie jest źródłem pewnych opóźnień w propagacji sygnału.

Niezależna transmisja danych obciążona jest ryzykiem występowania rozbieżności czasowych pomiędzy danymi obrazowymi a informacją o ich położeniu przestrzennym, pomimo akwizycji danych w tym samym czasie. Brak synchronizacji czasowej niekorzystnie wpływa na dokładność kalibracji przestrzennej obrazu ultrasonograficznego, podczas której konieczne jest jej przemieszczanie [86].

Celem kalibracji czasowej jest wyznaczenie wartości przesunięcia czasowego pomiędzy strumieniami danych przekazywanymi z systemów śledzenia oraz z konwertera sygnału video do komputera zapisującego dane. Każde z wykorzystywanych urządzeń ma własne znaczniki czasowe oparte na działaniu zegara sprzętowego lub znacznik jest nadawany przez komputer podczas zapisu. Zależ-

ność pomiędzy tymi dwoma systemami określana jest poprzez analizę sygnału obrazowego oraz współrzędnych położenia głowicy liniowej.

Protokół kalibracji zakłada umieszczenie głowicy czołem w dół w pojemniku z wodą, napełnioną do poziomu pozwalającego na obserwację w obrazie USG dna naczynia przy minimalnym zanurzeniu głowicy. Następnie, wykonywane są powolne ruchy głowicą w osi góra–dół, przy czym dno naczynia powinno pozostawać w polu widzenia głowicy. Zapis położenia znacznika umocowanego na powierzchni głowicy oraz obrazów ultrasonograficznych dna naczynia – widocznego jako jasny, jednorodny pas – realizowany jest w ciągu ok. 10 s [169].

Dalsze przetwarzanie zebranych danych obejmuje segmentację struktury odpowiadającej granicy dna naczynia z wodą. Następnie wyznaczany jest punkt środkowy struktury wraz ze znacznikiem czasowym. Drugi etap analizy danych składa się z wyznaczenia głównej osi ruchu i określenia głównej pozycji znacznika umieszczonego na głowicy względem tej osi. Na podstawie wartości minimalnego błędu średniokwadratowego pomiędzy analizowanym sygnałem obrazowym a sygnałem odpowiadającym położeniu znacznika określana jest stała wartość przesunięcia czasowego [168]. Etap kalibracji czasowej jest niezbędny w celu minimalizacji błędów przestrzennych, które są obecne podczas ruchu śledzonych narzędzi.

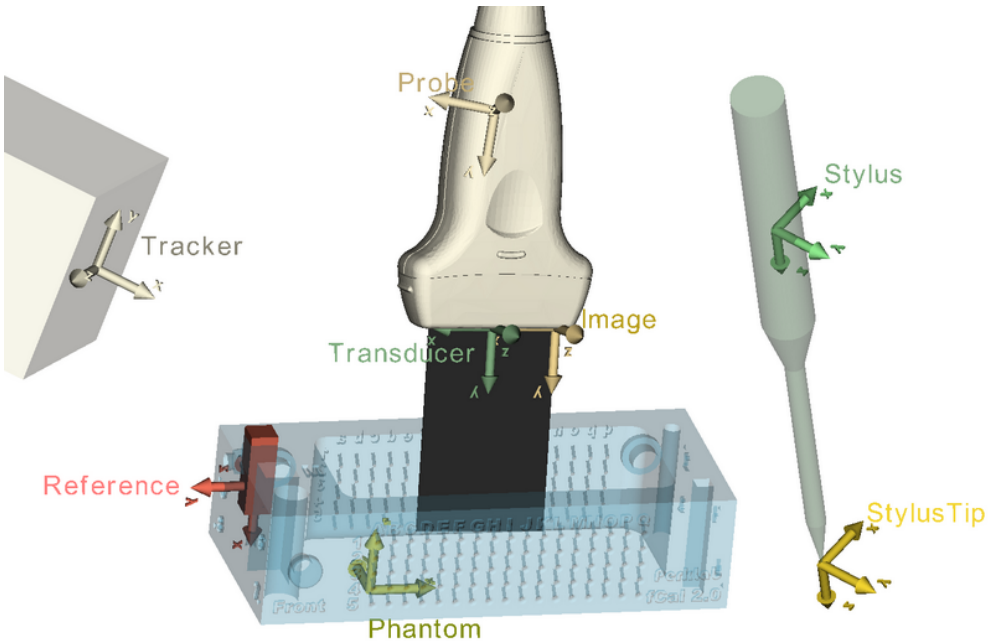
3.3.5 Kalibracja przestrzenna obrazu ultrasonograficznego

Podczas badania ultrasonograficznego za pomocą głowicy USG obrazowane są wybrane przez lekarza tkanki i narządy. Wykorzystanie systemu śledzenia umożliwi określenie pozycji znacznika umieszczonego sztywno na powierzchni głowicy USG. Koniecznym zatem staje się określenie zależności geometrycznej pomiędzy śledzonym znacznikiem a płaszczyzną obrazowania głowicy. W tym celu wykonywana jest kalibracja przestrzenna dwuwymiarowego obrazu ultrasonograficznego. Wynikiem kalibracji jest wartość macierzy translacji i rotacji M_I^P , która pozwala na transformację układu współrzędnych obrazu I do układu współrzędnych znacznika P (Rys. 3.10).

Protokół kalibracji przestrzennej opisywany w literaturze najczęściej oparty jest o wykorzystanie fantomu żyłkowego o znanej geometrii [19, 63, 64]. Przykład takiego fantomu przedstawiono na rysunku 3.11. Procedura kalibracji przestrzennej może też opierać się o konieczność wskazania określonych punktów za pomocą skalibrowanego wskaźnika zgodnie ze schematem przedstawionym na rysunku 3.12 [169, 178].

Określenie zależności pomiędzy obrazem ultrasonograficznym a znacznikiem głowicy USG wymaga umieszczenia tej głowicy czołem w dół w pojemniku wypełnionym płynem umożliwiającym swobodną propagację fal dźwiękowych

¹ PLUS toolkit - PLUS applications user manual <http://perk-software.cs.queensu.ca>

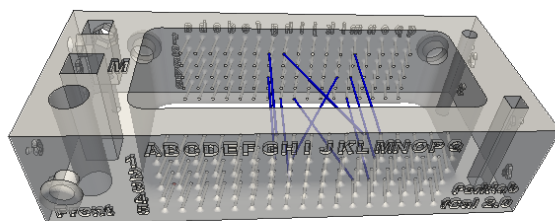


Rys. 3.10: Markery umieszczone na śledzonych narzędziach¹.

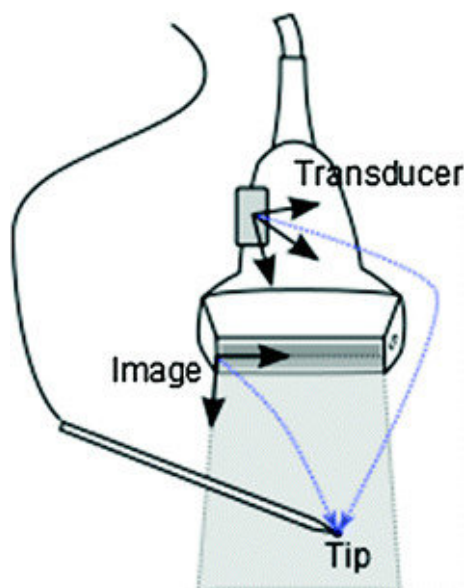
z prędkością $1540 \frac{\text{m}}{\text{s}}$ [58]. Prędkość ta jest typową wartością prędkości propagacji fal ultradźwiękowych w tkance miękkiej, która została zaimplementowana w większości aparatów ultrasonograficznych. Przykładem substancji spełniających to kryterium jest czysta woda w temperaturze 50°C [58]. W razie stosowania płynu o innej prędkości propagacji fal, należy zastosować przesunięcie segmentowanych punktów w kierunku czoła głowicy na podstawie współczynnika $\frac{v_{\text{woda}}}{v_{\text{rzeczywista}}}$ [64, 93], ewentualnie ustalić w aparacie USG wartość prędkości o wartości $1498 \frac{\text{m}}{\text{s}}$, która to wartość odpowiada prędkości fal ultradźwiękowych w wodzie w temperaturze pokojowej [58]. Poprawne określenie prędkości fali ultradźwiękowej jest wymagane do właściwego zrealizowania kalibracji przestrzennej.

Następny etap związany jest z ustawieniem wzmocnienia i określeniem zakresu segmentacji dynamicznych elementów widocznych w śledzonym obrazie USG. Istotne jest dobranie parametrów, dla których nie są widoczne takie obiekty jak dno lub pęcherzyki gazu, a obiekty odpowiadające żyłkom fantomu stanowią jasne, kontrastujące z tłem plamki o niewielkiej średnicy. To właśnie na podstawie żyłek określane są kąty i odcinki pomiędzy określonymi punktami, których wartości są mierzalne. Współrzędne żyłek są zdefiniowane w pliku konfiguracyjnym w rzeczywistym układzie współrzędnych [19].

¹ PLUS toolkit - printable 3D models catalog <http://perk-software.cs.queensu.ca>



Rys. 3.11: Fantom żyłkowy służący do przestrzennej kalibracji dwuwymiarowego obrazu USG ¹.



Rys. 3.12: Przestrzenna kalibracja obrazu USG z wykorzystaniem końcówki czujnika (źródło: [169]).

Detekcja żyłek widocznych na obrazie USG – a tym samym poprawne wyznaczenie zależności pomiędzy obrazem a markerem – możliwa jest podczas obrazowania wszystkich żyłek fantomu (Rys. 3.13). Na podstawie wysegmentowanych punktów fantomu i macierzy zależności pomiędzy fantomem a jego znacznikiem referencyjnym, wyznaczana jest macierz rotacji i translacji obrazu do znacznika na powierzchni głowicy.

Dokładne określenie transformacji obrazu do znacznika na głowicy USG jest utrudnione z powodu grubości wiązki ultradźwiękowej. Główną przyczynę niedokładności stanowi lokalizacja punktów 3D w oparciu o obraz ultrasono-



Rys. 3.13: Przekrój poprzeczny przez żyłki fantomu w obrazowaniu USG.

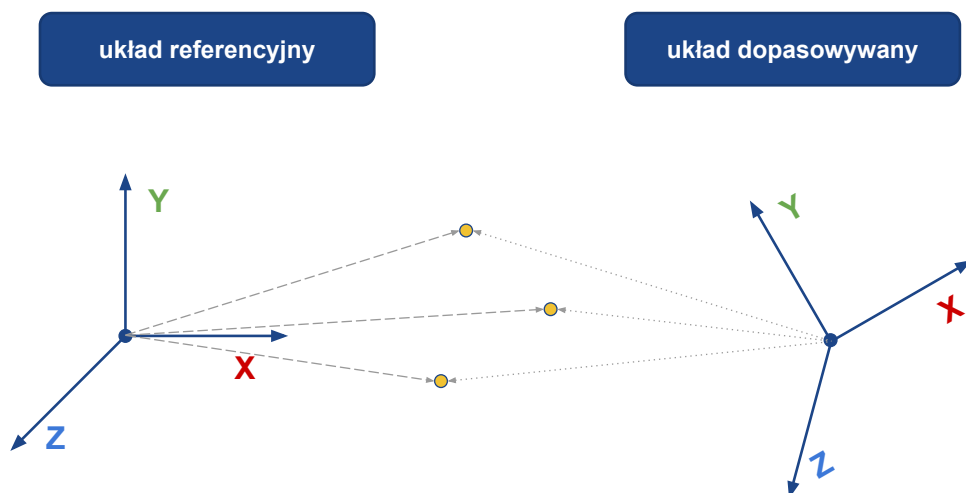
graficzny powodująca pojawianie się na obrazie obiektów oddalonych o kilka milimetrów względem płaszczyzny obrazowania oraz rozmycie granic obrazowanych obiektów [169].

Stała macierz rotacji i translacji M_I^L stanowiąca zależność pomiędzy śledzonym znacznikiem L a płaszczyzną obrazowania I w układzie współrzędnych O wyznaczana jest protokołem kalibracji przestrzennej z wykorzystaniem fantomu żyłkowego o znanej geometrii definiowanej za pomocą żyłek polimerowych rozpiętych na wydrukowanym z tworzywa ABS szkielecie, którego model został przedstawiony na rysunku 3.11 [19]. Na obrazowanym ultrasonograficznie fantomie – a dokładniej: jego przekroju – uwidocznione są jasne struktury odpowiadające żyłkom, które silnie odcinają się od ciemnego tła wody. Segmentacja jasnych obiektów pozwala na określenie ich położenia, a przez to wyznaczenie płaszczyzny obrazu I . Przemieszczenie położenia śledzonych punktów względem znacznika pozwala na wskazanie przestrzennego położenia obrazu USG względem tego znacznika, zgodnie z zależnością:

$$P^G = M_L^G \cdot M_O^L \cdot I^O, \quad (3.5)$$

gdzie za globalny układ współrzędnych przyjmuje się układ współrzędnych definiowany przez generator pola elektromagnetycznego systemu EMTS.

System śledzenia rejestruje znaczniki i podaje ich położenie wyrażając wartości w milimetrach. Wymiary obrazów domyślnie wyrażane są w pikselach, które odpowiadają wierszom i kolumnom obrazów. Przy zachowaniu stałej głębokości wiązki obrazującej, w trakcie kalibracji możliwe jest wyznaczenie zależności wyrażającej wymiary obrazów również w milimetrach.



Rys. 3.14: Rejestracja tożsamyh punktów w dwóch układach współrzędnych.

3.3.6 Kalibracja układów współrzędnych systemów śledzenia położenia

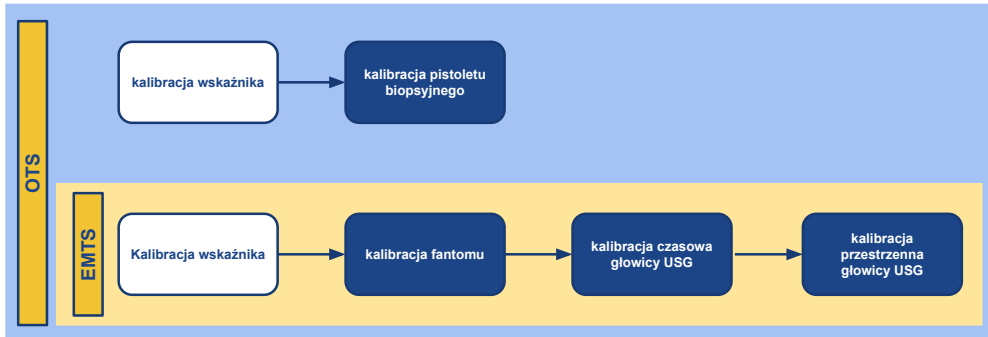
Wykorzystane systemy śledzenia: optyczny Polaris i elektromagnetyczny Aurora definiują własne globalne układy współrzędnych. Kalibracja układów współrzędnych tych systemów również jest związana z wyznaczeniem stałej macierzy transformacji pomiędzy globalnym układem współrzędnych systemu elektromagnetycznego, a globalnym układem współrzędnych systemu optycznego. W tym celu wyznaczane są przynajmniej trzy te same punkty w przestrzeni (z zachowaniem kolejności ich akwizycji), za pomocą śledzonych wskaźników natywnych dla każdego z systemów [170], celem znalezienia związku pomiędzy dwoma układami współrzędnych (Rys. 3.14). Wskazywane punkty powinny być nieruchome podczas rejestracji. Następnie na podstawie par punktów wyznaczana jest, z wykorzystaniem metody najmniejszych kwadratów, macierz rotacji i translacji [62].

Podejście to pozwala na określenie stałego przekształcenia

$$OTS^G = M_L^G \cdot M_O^L \cdot EMTS^O, \quad (3.6)$$

gdzie za globalny układ współrzędnych przyjmuje się układ współrzędnych definiowany przez generator pola elektromagnetycznego systemu EMTS. Odległości wyrażane są w milimetrach.

Zależność przekształcająca układ globalny jednego z systemów do układu drugiego systemu pozwala na wyznaczanie położenia każdego znacznika względem układu współrzędnych dowolnego systemu śledzenia. Kolejne kroki kalibracji przedstawiono na rysunku 3.15.



Rys. 3.15: Etapy kalibracji końcówki igły pistoletu biopsyjnego oraz kalibracji przestrzennej i czasowej głowicy biopsyjnej za pomocą optycznego i elektromagnetycznego systemu śledzenia.

3.4 Podsumowanie

W niniejszym rozdziale opisano konstrukcję systemu rejestracji danych, zasadę działania optycznego oraz elektromagnetycznego systemu śledzenia oraz integrację ich działania. Przedstawiony został sposób kalibracji wszystkich elementów systemu niezbędnych do prawidłowego śledzenia obrazu ultrasonograficznego i trajektorii igły biopsyjnej.

Pierwsza część rozdziału poświęcona została inżynierskiej konstrukcji stanowiska pomiarowego, które zostało wypracowane w konsultacji z personelem medycznym szpitala Narodowego Instytutu Onkologii im. Marii Curie-Skłodowskiej – Oddziału w Gliwicach, celem rozszerzenia stanowiska klinicznego o elementy, które w minimalnym stopniu ingerują w swobodę wykonywania zabiegu biopsji gruczołowej. Opisano zasadę działania obu systemów śledzenia, przedstawiono ich wady i zalety, jak również wpływ na pracę zespołu medycznego.

Druga część rozdziału przedstawia protokół kalibracji stanowiska pomiarowego. Kalibracja i śledzenie obrazu ultrasonograficznego realizowane jest za pomocą systemu elektromagnetycznego. Kalibracja i śledzenie pistoletu biopsyjnego monitorowane są za pomocą systemu optycznego. Przedstawiona została zmodyfikowana metoda kalibracji pistoletu biopsyjnego oraz kalibracja systemów śledzenia.

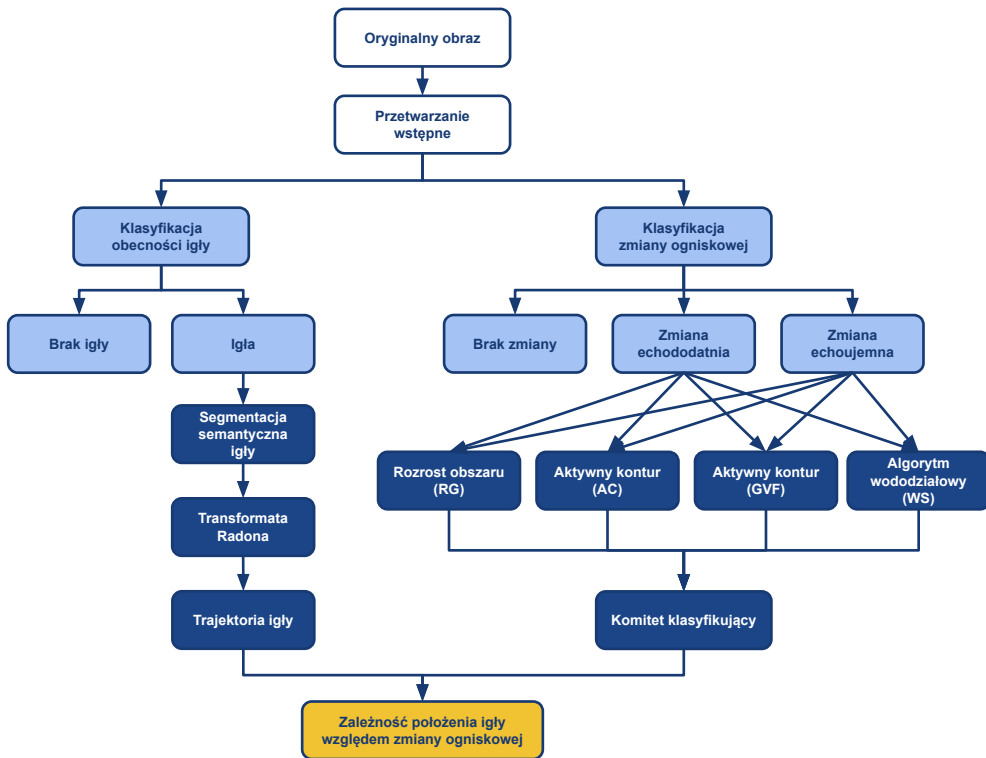
4. Klasyfikacja i segmentacja struktur w obrazach USG

W rozdziale tym zawarto opis przetwarzania dwuwymiarowych obrazów ultrasonograficznych metodami klasycznymi oraz konwolucyjnymi (splotowymi) sieciami neuronowymi (ang. *convolutional neural network*, CNN). Zarówno metody klasyczne jak i głębokie wykorzystywane są z powodzeniem w wielu algorytmach przetwarzania obrazów medycznych. W trakcie realizacji niniejszej rozprawy zaimplementowano dwie sieci konwolucyjne, których architekturę i działanie przygotowano do realizacji następujących zadań: sieć pierwsza dokonuje klasyfikacji obrazów pod kątem (i) widoczności igły biopsyjnej oraz (ii) typu zmiany ogniskowej. Druga sieć realizuje segmentację semantyczną wykrytej igły. Za pomocą komitetu klasyfikującego (ang. *ensemble*) realizowana jest klasyczna segmentacja zmiany ogniskowej. Ogólny schemat organizacji przetwarzania przedstawiono na schemacie 4.1.

W pierwszej kolejności zostaną przedstawione informacje dotyczące przetwarzania wstępnego obrazów, następnie ogólne i teoretyczne dotyczące sieci konwolucyjnych. W dalszych podrozdziałach opisane są szczegółowo architektury sieci, sposób ich treningu, walidacji i testowania, a także przedstawiono etapy przygotowania danych źródłowych. W kolejnych przedstawiono propozycję segmentacji zmiany ogniskowej. Zaprojektowanie, zaimplementowanie, nauczanie i przetestowanie proponowanych metod oraz przygotowanie zbiorów danych stanowi w ocenie autorki rozprawy jej główne dokonanie naukowe.

4.1 Przetwarzanie wstępne

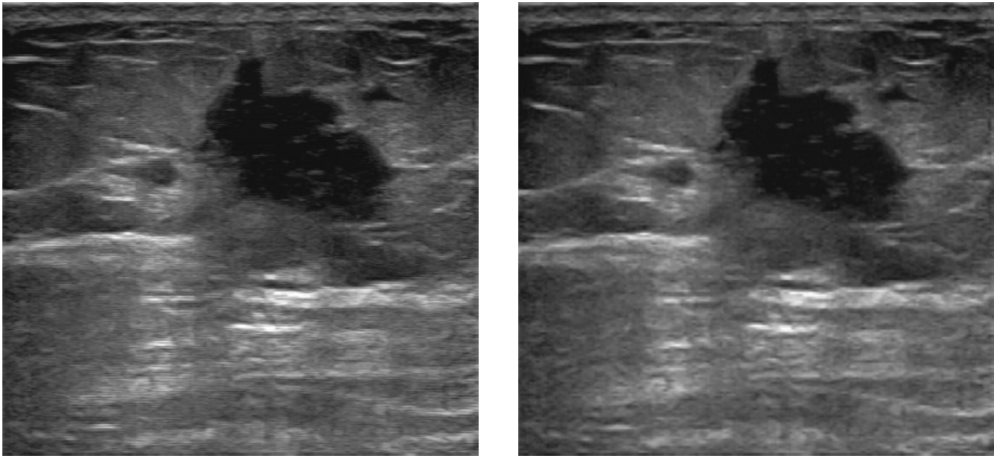
Akwizycja obrazu ultrasonograficznego pochodzącego z aparatu ultrasonograficznego obarczona jest dodatkowymi informacjami na marginesach w postaci danych osobowych pacjenta, informacji dot. sposobu realizowania obrazowania czy też czarnej ramki. Informacje te są istotne z powodów administracyjnych, niemniej jednak z pozycji przetwarzania danych obrazowych stanowią artefakty, oraz zmniejszają istotny z punktu widzenia klasyfikacji i segmentacji obraz po konwersji do wskazanych wymiarów. Etap przetwarzania wstępnego



Rys. 4.1: Schemat proponowanego całościowego przetwarzania obrazów ultrasonograficznych.

zakłada więc wycięcie istotnego diagnostycznie fragmentu. Przechwytywany sygnał ultrasonograficzny obarczony jest zjawiskiem przeplotu (ang. *interlacing*), ze względu na akwizycję sygnału ultrasonograficznego za pomocą złącza standardu S-Video [122]. Przeplot widoczny w obrazach USG usuwany jest za pomocą metody interpolacji liniowej, natomiast wartość pikseli jest zaokrąglana z wykorzystaniem metody najbliższego sąsiada (Rys. 4.2) [115].

Po usunięciu przeplotu wykonana została konwersja danych obrazowych z przestrzeni kolorów RGB do skali szarości oraz zmieniony został rozmiar obrazów zgodnie z wymaganiami proponowanej architektury sieci konwulcyjnej (256×256) przy użyciu interpolacji metodą najbliższego sąsiada. Wymiary obrazu przekazywanego na wejściu do proponowanych architektur zostały wybrane na podstawie analizy udziału procentowego maski igły oraz zmian ogniskowych w całym obrazie. Wstępnie przetworzony obraz jest oznaczony jako $I(x, y)$.



(a) Obraz oryginalny.

(b) Obraz po usunięciu przeplotu.

Rys. 4.2: Przykładowy obraz ultrasonograficzny zmiany ogniskowej poddawany przetwarzaniu.

4.2 Konwolucyjne sieci neuronowe

Konwolucyjne sieci neuronowe (ang. *convolutional neural network*, CNN) są szeroko stosowanym narzędziem umożliwiającym zaawansowaną analizę i przetwarzanie obrazów [3]. CNN działają poprzez kolejne ekstrakcje złożonych cech wyższego rzędu charakterystycznych dla sąsiadujących regionów pikseli w danej warstwie, a następnie ich uogólnienie w formie obrazów w warstwie następnej. Efektem kolejnych uogólnień jest reprezentacja analizowanego zbioru poprzez zestaw map cech charakterystycznych dla obrazu źródłowego.

Architektura danej CNN zależy od zadania, jakie ma być przez nią realizowane. W przypadku sieci klasyfikującej charakterystyczne są następujące warstwy: wejściowa, warstwy ukryte odpowiedzialne za ekstrakcję cech (to jest warstwa konwolucyjna z funkcją aktywacji oraz warstwa głosująca), a także warstwa wyjściowa, za pomocą której szacowane jest prawdopodobieństwo przynależności obrazu do danej klasy [12, 48].

W przypadku sieci CNN segmentującej również stosowane są warstwy ukryte: konwolucyjna i głosująca (łączenia) celem ekstrakcji cech (podobnie jak w przypadku sieci klasyfikującej [48]). Prawdopodobieństwo przynależności piksela do segmentowanego obiektu lub tła obliczane jest przez warstwę konwolucyjną z funkcją aktywacji typu *softmax* i warstwę klasyfikacji pikseli. Zasadniczą różnicą między sieciami CNN klasyfikującymi a segmentującymi jest to, że w przypadku tych drugich występuje dodatkowa warstwa ukryta, czyli warstwa transponowanej konwolucji [45, 57], za pomocą której mapy cech są próbkowane w górę ku coraz wyższym rozdzielczościom (ang. *up-sampling*) aż

do wygenerowania finalnej maski wynikowej.

Zarówno w przypadku klasyfikacji, jak i segmentacji, stosuje się mechanizmy normalizacji i techniki regularyzacji celem uniknięcia nadmiernego dopasowania trenowanego modelu do danych w zbiorze treningowym. Rolę regularyzatora pełni warstwa normalizacji; stosowany jest także regularyzator typu *dropout*. Warstwa łączenia, poza dokonywaniem ekstrakcji cech, również pełni funkcję regularyzatora [147].

W literaturze naukowej opisano różne sieci CNN, których istotnym elementem jest sposób przesyłu informacji między poszczególnymi warstwami. W szczególności wyróżnia się następujące typy topologii tych sieci.

Architektury jednogałęziowe (ang. *single-branch*), których przykładami są sieć AlexNet [83] czy rodzina sieci VGG [151], cechują się szeregowym połączeniem kolejnych warstw, to znaczy dane wyjściowe warstwy n -tej stanowią dane wejściowe warstwy $n+1$.

W przypadku architektury wielogałęziowej (ang. *multi-branch*) dane wejściowe warstwy n -tej mogą stanowić wejście do kilku kolejnych, równoległych warstw. Podejście to umożliwia zrównoleglenie ekstrakcji różnych zestawów cech zbioru danych. Przykładem wykorzystującym tę architekturę jest algorytm przetwarzania segmentacji igły w obrazie 3D w oparciu o równoczesne próbkowanie obrazu w trzech różnych kierunkach [194].

Wyróżniana jest również architektura oparta o acykliczny graf skierowany (ang. *directed acyclic graph*, DAG). W podejściu tym możliwe jest przekazanie danych wyjściowych warstwy n -tej do innych, niekoniecznie bezpośrednio następujących po sobie warstw, do tego warstwa n -ta może przyjmować na swoje wejście dane z więcej niż jednej warstwy. Przykładami tego typu topologii są sieć GoogLeNet [162], rodzina sieci ResNet [56] lub InceptionResNetV2 [161].

W tabeli 4.1 przedstawiono kilka popularniejszych CNN, których wstępny trening realizowany był obrazami z bazy ImageNet [139]. Głębokość sieci jest to liczba kolejnych warstw na ścieżce od warstwy wejściowej do warstwy wyjściowej. Reprezentacja obrazu w warstwie wejściowej różni się wymiarami szerokości i wysokości, ale wstępny trening wszystkich tych sieci przeprowadzany był obrazami w przestrzeni barw RGB, to znaczy liczba kanałów obrazów wejściowych równa jest 3. Wymiary obrazów wejściowych są dla każdej sieci niezmiennie i wynoszą od 224 do 299 pikseli szerokości i wysokości. Każdą z wymienionych w tabeli sieci można doszkolić do bardziej specyficznych zastosowań zgodnie z zasadami transfer learningu opisywanego w publikacjach poświęconych uczeniu maszynowemu (np. [83, 147, 161]).

Oryginalne architektury splotowych sieci neuronowych [134, 172], a także rozwiązania oparte na wstępnie wytrenowanych modelach są powszechnie stosowane w analizie obrazów ultrasonograficznych raka piersi [5, 18]. Pierwszą historycznie szeroko stosowaną, wstępnie trenowaną architekturą jest AlexNet, która składa się z 8 warstw ukrytych oraz warstwy maxpooling, która zmniejsza

Tab. 4.1: Wybrane pretrenowane sieci CNN i ich parametry.

Ref	Sieć	Głębokość	Warstwy	Połączenia	Rozmiar [MB]	Parametry [mln]	Wymiary obrazu wejściowego
[83]	AlexNet	8	25	–	227	61	227×227×3
[151]	VGG16	16	41	–	515	138	224×224×3
[151]	VGG19	19	47	–	535	144	224×224×3
[70]	SqueezeNet	18	68	75×2	4,5	1,24	227×227×3
[56]	ResNet-18	18	72	79×2	44	11,7	224×224×3
[162]	GoogLeNet	22	144	170×2	27	7	224×224×3
[56]	ResNet-50	50	177	192×2	96	25,6	224×224×3
[162]	Inception-v3	48	316	350×2	89	23,9	299×299×3
[56]	ResNet-101	101	347	379×2	167	44,6	224×224×3
[65]	DenseNet-201	201	709	806×2	77	20	224×224×3
[161]	Inception-ResNet-v2	164	825	922×2	209	55,9	299×299×3

sza nadmiarowość wyodrębnionych cech. Ostatnie trzy warstwy architektury AlexNet to warstwy w pełni połączone (ang. *fully connected layer, FC*). Ten przykład architektury wykorzystuje jako funkcję aktywacji rektyfikowaną jednostkę liniową (ang. *rectified linear unit, ReLU*). Podejście to pozwala skrócić czas szkolenia i poprawia efektywność klasyfikacji [83, 139].

Architektura VGG16 jest przykładem jednołązowej sieci zbudowanej z 16 warstw ukrytych, która została opracowana przez Visual Geometry Group (VGG). Cechą charakterystyczną rodziny VGG jest zastosowanie małego pola receptywnego o wymiarach 3×3, jak również wykorzystanie funkcji aktywacji ReLU po każdej z warstw ukrytych. Efektem jest zwiększenie efektywności klasyfikacji modelem opartym o VGG16 w stosunku do architektury AlexNet [151, 162]. Innym przykładem pochodzącym z rodziny VGG jest architektura VGG19, której struktura jest zgodna z VGG16. Różnica polega na liczbie warstw ukrytych, która w tym przypadku wynosi 19 [151].

Architektura SqueezeNet jest splotową siecią neuronową, w której zmniejszeniu ulega liczba parametrów za pomocą modułów ogniowych (ang. *fire modules*). Każdy z modułów składa się ze ściskanych warstw konwolucji (squeeze convolution layers), które cechują się filtrami o wymiarach 1×1. Podejście to skutkuje kompresją liczby parametrów i rozmiaru sieci, co powoduje wzrost mobilności rozwiązania opartego o architekturę SqueezeNet [70].

Sieć GoogLeNet stanowi przykład ukierunkowanego grafu acyklicznego, w którym wynik n -tej warstwy jest przekazany do innych, niekoniecznie bezpośrednio następujących kolejno po sobie warstw, a n -ta warstwa może przyj-

mować mapy cech jako dane wejściowe z więcej niż jednej warstwy. Architektura 22-warstwowa wykorzystuje podejście modułu inicjacji (ang. *Inception module*), który wykrywając cechy w różnych skalach oraz używając splotów z różnymi filtrami, pozwala zmniejszyć koszt obliczeniowy uczenia się szerokiej architektury [72, 162].

Architektura ResNet101 należy do grupy określanej mianem sieciami rezydualnymi (ang. *residual networks*), w których przypadku zakłada się wykorzystanie zwiększonej liczby ułożonych sekwencyjnie warstw ukrytych. Takie podejście pozwala zwiększyć dokładność modelu kosztem wydłużenia czasu przetwarzania [56].

Sieć InceptionResNetV2 jest przykładem architektury, która łączy moduły inicjacji z podejściem sieci rezydualnej. Wykorzystanie połączeń rezydualnych pozwala wyeliminować problem degradacji spowodowanej głębokością sieci oraz skrócić czas szkolenia [161].

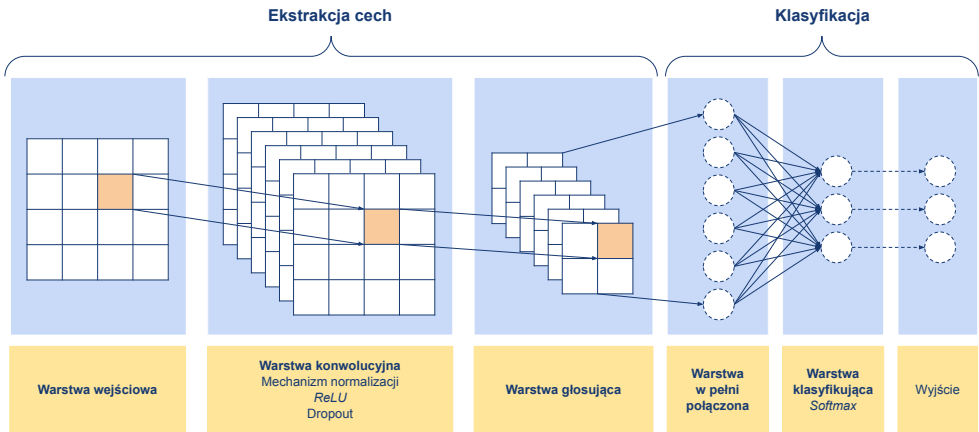
Każda z wymienionych sieci została wstępnie przeszkolona przy użyciu bazy danych ImageNet, która jest bazą zawierającą ponad 14 milionów zdjęć z adnotacjami widocznych obiektów. ImageNet to największa i najpopularniejsza baza danych wykorzystywana w rozpoznawaniu obrazów [139].

4.3 Elementy architektury sieci CNN

Reprezentacja obrazu wejściowego generowana jest za pomocą pierwszej warstwy — warstwy wejściowej, poddawana jest następnie operacji splotu za pomocą konwolucyjnych sieci neuronowych celem ekstrakcji złożonych i uogólnionych cech wyższego rzędu, które stanowią podstawę do klasyfikacji treści obrazu (Rys. 4.3). Ekstrakcja cech realizowana jest również za pomocą warstw głosujących. Pomiedzy tymi dwoma kategoriami cech, zastosowanie znajduj mechanizmy normalizacji oraz regularyzatory, których zadaniem jest eliminacja nadmiarowej informacji celem wykluczenia zjawiska przeuczenia [12, 48]. Druga cze architektury powiazana jest ścisle z realizowanym zadaniem. W przypadku architektury klasyfikujcej kluczowa jest warstwa w pelni polaczona z funkcj aktywacji typu *softmax*, celem wyznaczenia prawdopodobieństwa przynależności obrazu do zdefiniowanych klas. W przypadku architektury, której celem jest segmentacja, klasyfikacja rozpoznajca obiekt lub tlo, realizowana jest na poziomie pikseli a nie obrazów, a nastepnie wymagane jest zastosowanie warstw transponowanej konwolucji, celem wyznaczenia maski binarnej o wymiarach obrazu podawanego na wejściu [45, 57].

4.3.1 Warstwa wejściowa

Pierwsza warstwa sieci, określanej mianem warstwy wejściowej (ang. *input layer*), odpowiada reprezentacji obrazu wejściowego I . Kazdy piksel obrazu



Rys. 4.3: Schematyczne przedstawienie elementów architektury konwolucyjnej sieci neuronowej. Pogrubione nazwy wskazują niezbędne elementy architektury klasyfikującej, zaznaczone kursywą funkcje aktywacji, pozostałe to mechanizmy regularyzacji.

przekazywany jest osobnemu neuronowi wejściowemu. Liczba kanałów c obrazu uzależniona jest od jego typu. W przypadku obrazów kolorowych w przestrzeni barw RGB, tak jak w obrazach z bazy ImageNet [139] wartość parametru c jest równa 3, natomiast dla obrazów ultrasonograficznych nagrywanych w skali szarości $c = 1$.

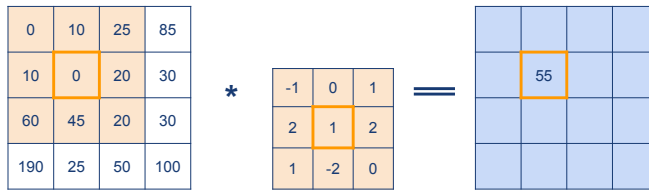
4.3.2 Warstwa konwolucyjna

Idea konwolucyjnych sieci neuronowych bazuje na operacji splotu (ang. *convolution*). Dwuwymiarowy obraz w skali szarości reprezentowany jest za pomocą macierzy I , której elementy $I(i, j)$ stanowią wartość odpowiadającą jasności piksela o współrzędnych (i, j) . Zbiór wag w warstwie konwolucyjnej określany jest mianem jądra (ang. *kernel*, K) lub maską filtra. Operacja splotu dwuwymiarowego (Rys. 4.4) definiowana jest wzorem:

$$J(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i - m, j - n)K(m, n) \quad (4.1)$$

gdzie para (i, j) to współrzędne piksela w analizowanym obrazie I oraz obrazie wynikowym J , natomiast para (m, n) odpowiada współrzędnym wag jądra K [48].

Wynikiem J konwolucji danych wejściowych i jądra jest mapa cech (ang. *feature map*), będąca zbiorem cech wyższego rzędu, wyodrębnionym z analizowanego obrazu. W przypadku pierwszej operacji splotu danymi wejściowymi jest oryginalny obraz, natomiast na wejście kolejnych warstw konwolucji podawane



Rys. 4.4: Przykład operacji splotu filtrem 3×3 . Wynik mapowania dla określonego regionu został zaznaczony kolorem pomarańczowym.

są trójwymiarowe mapy cech obliczone w warstwie poprzedzającej. Obrazy wielokanałowe traktowane są jako obrazy trójwymiarowe, gdzie liczba kanałów c stanowi trzeci wymiar, a ich konwolucja przeprowadzana jest zgodnie z zależnością:

$$J(i, j) = (I * K)(i, j) = \sum_m \sum_n \sum_c I(i - m, j - n, c) K(m, n, c) \quad (4.2)$$

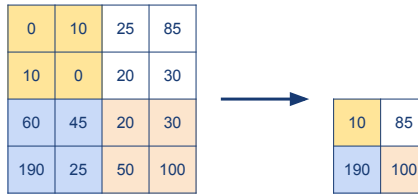
gdzie: c jest trzecią współrzędną mapy cech w przypadku obliczeń dla kolejnych warstw konwolucyjnych lub numerem kanału obrazu kolorowego. Efektem przetwarzania są nowe mapy cech danej warstwy [45, 123].

Wykorzystanie operacji splotu w CNN w porównaniu do operacji macierzowych stosowanych w sieciach klasycznych jest korzystne ze względu na lokalność połączeń, powtarzalne wartości wag połączeń jądra przesuwającego się po analizowanym obrazie oraz ekwiwariancję względem tego przesunięcia [48].

4.3.3 Warstwa głośująca

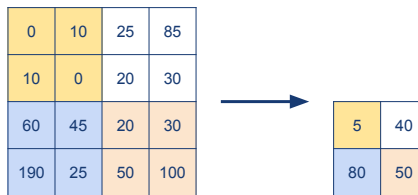
Warstwy konwolucyjne prowadzą do nadmiarowości (redundancji) danych ze względu na wielokrotne wykorzystywanie tych samych pikseli w trakcie wykonywania operacji przeplotu. Nadmiarowa informacja nie poprawia skuteczności trenowanej sieci, wręcz przeciwnie, może skutkować nadmiernym dopasowaniem modelu do danych w zbiorze treningowym (przeuczeniem, ang. *overfitting*). Jedną ze strategii pozwalającą na eliminację negatywnego wpływu nadmiarowości danych na trenowany model jest zastosowanie warstw głośujących.

Warstwa głośująca (ang. *pooling layer*), nazywana również warstwą łączenia lub redukującą, realizuje progresywną redukcję przestrzenną obrazu lub warstwy, skutkując tym samym zmniejszeniem obszaru, który poddawany jest operacji splotu w kolejnych warstwach konwolucyjnych. Efektem jest wyodrębnienie inwariantnych cech, które są bardziej złożone ze względu na uogólnienie cech charakterystycznych dla sąsiadujących grup pikseli. Redukcja danych eliminuje również wpływ szumów oraz niewielkich zmian lub obiektów na efektywność sieci [113, 123].



Rys. 4.5: Przykład operacji *maxpooling* z filtrem 2×2 i krokiem 2. Wyniki mapowania określonego regionu zostały zaznaczone odpowiadającym temu regionowi kolorem.

Idea redukcji oparta jest o podział obrazu redukowanego na regiony za pomocą maski filtracji z zadaniem krokiem. Następnie dla każdego z regionów wyznaczana jest wartość, która jest przypisywana pikselowi w nowo powstałym, zredukowanym obrazie. Stosowanymi najczęściej metodami łączenia jest wybór największej wartości (ang. *maxpooling layer*, Rys. 4.5) lub wartości średniej (ang. *averagepooling layer*, Rys. 4.6) z danego regionu.



Rys. 4.6: Przykład operacji *averagepooling* z filtrem 2×2 i krokiem 2. Wyniki mapowania określonego regionu zostały zaznaczone odpowiadającym temu regionowi kolorem.

Zmniejszenie wymiarów obrazu skutkuje zmniejszeniem liczby parametrów poddawanych treningowi, a w efekcie uproszczeniem modelu. Pozwala to na skrócenie czasu trwania zarówno treningu, jak również późniejszej klasyfikacji (segmentacji) danych za pomocą ostatecznie wytrenowanego modelu. Warstwy głośujące zwiększają zatem skuteczność sieci. Redukcja oparta o warstwy głośujące powoduje jednak utratę części informacji, co może utrudnić wskazanie małych obiektów. Istotny jest zatem właściwy dobór wymiarów regionu poddawane łączeniu. Najczęściej stosowanym rozmiarem filtra jest 2×2 z krokiem 2, który powoduje dwukrotne zmniejszenie obrazu, czyli czterokrotną redukcję danych. W obrazach wielokanałowych każdy kanał obrazu redukowany jest indywidualnie [113].

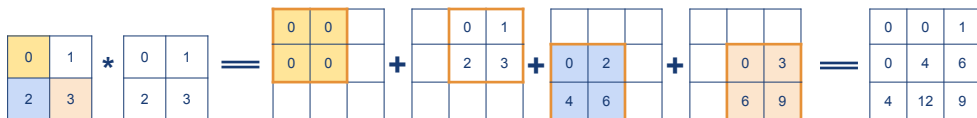
4.3.4 Warstwa transponowanej konwolucji

Operacje splotu i łączenia realizowane w warstwach konwolucyjnej i łączenia celem ekstrakcji cech ogólnych prowadzą do redukcji wymiarów wynikowych obrazów. W algorytmach segmentacji kluczowy element stanowi wykorzystanie map cech do uzyskania binarnego obrazu wynikowego o wymiarach odpowiadających reprezentacji obrazu wejściowego I . Utworzenie wyjściowej mapy cech o wyższej rozdzielczości możliwe jest poprzez zastosowanie warstwy transponowanej konwolucji (ang. *transposed convolutional layer*).

Operacja splotu stosowana w warstwie konwolucyjnej na podstawie obrazu wejściowego wyodrębnia mapę cech, której wymiary są mniejsze lub takie same w porównaniu do danych wyjściowych. Transponowany splot działa w przeciwnym kierunku do splotu: na wejście podawana jest mapa cech o mniejszych wymiarach (niższej rozdzielczości), a wynikiem działania jest mapę o rozdzielczości wyższej. Operacja próbkowania w górę (ang. *up-sampling*) realizowana jest zgodnie ze wzorem:

$$f_{\uparrow}(i, j) = \begin{cases} f(\frac{i}{N}, \frac{j}{M}) & \text{jeżeli } N \bmod i = 0 \vee M \bmod j = 0 \\ 0 & \text{w przeciwnym wypadku.} \end{cases} \quad (4.3)$$

Podobnie jak w przypadku warstwy konwolucyjnej, również w warstwie transponowanej konwolucji przeprowadzana jest operacja splotu z użyciem zdefiniowanego jądra K , które stanowi zbiór uczonych w procesie treningu wag przesuwanych po wejściowej mapie cech zgodnie z zadaniem krokiem s (ang. *stride*). Efektem przetwarzania jest nowa, powiększona mapa cech [147].



Rys. 4.7: Przykład operacji próbkowania w górę filtrem 2×2 . Wynik mapowania dla określonego regionu został zaznaczony odpowiadającymi kolorami w mapie cech regionom.

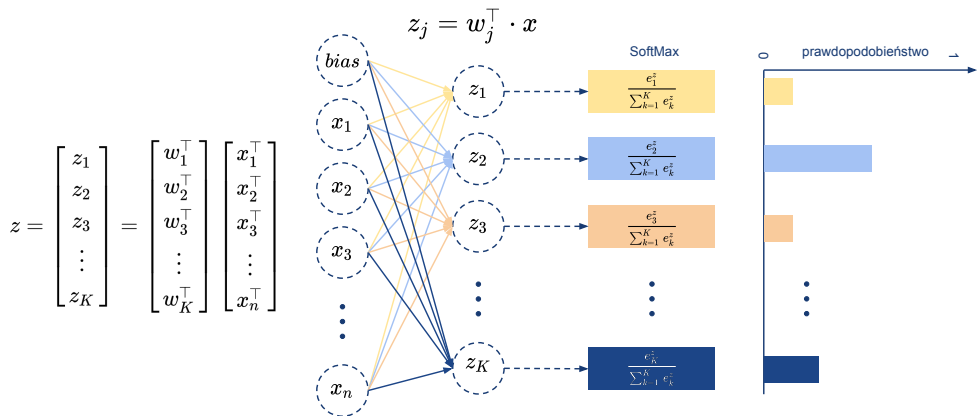
Operacja transponowanej konwolucji w segmentacji stosowana jest w celu wygenerowania maski wynikowej o wymiarach zgodnych z obrazem wejściowym I . Wykorzystanie operacji splotu z użyciem wyznaczonych w procesie treningu wartości wag jądra umożliwia eliminację artefaktów przyjmujących formę szachownicy w obrazie wynikowym, które są widoczne w przypadku zastosowania interpolacji, jak również cechuje się większą skutecznością [148].

4.3.5 Warstwa w pełni połączona

Warstwa w pełni połączona (ang. *fully connected layer*, *FC*) nazywana również warstwą gęstą (ang. *dense layer*) stanowi najczęściej przedostatnią warstwę w klasycznej, konwolucyjnej sieci neuronowej. Często stosowana jest ona również w niekonwolucyjnych sieciach neuronowych. Warstwa *FC* mnoży dane wejściowe x przez trenowaną macierz wag w oraz dodaje wektor odchylenia b (ang. *bias vector*).

$$z_j = w_j^\top \cdot x + b \quad (4.4)$$

Zadaniem warstwy gęstej jest określenie prawdopodobieństwa przynależności analizowanego obrazu wejściowego do klas [57]. Każda ukryta jednostka jest połączona z każdym neuronem warstwy gęstej i dla każdego z tych połączeń jest określona waga połączenia (Rys. 4.8). Wartość wagi połączenia wyznaczana jest podczas procesu treningu. Warstwy konwolucyjna, normalizacyjna i głosująca mają na celu ekstrakcję cech z analizowanych danych, natomiast warstwa w pełni połączona uogólnia te cechy na przestrzeń wyjściową. Na tej podstawie określone jest prawdopodobieństwo z jakim analizowane dane przynależą do klasy. Liczba argumentów na wyjściu z sieci neuronowej opowiada liczbie klas zbioru danych [45].



Rys. 4.8: Szacowanie prawdopodobieństwa przynależności do klasy w architekturze sieci neuronowej (na podstawie: [142]).

4.3.6 Warstwa klasyfikująca

Ostatnią warstwę architektury sieci klasyfikującej stanowi warstwa klasyfikująca (ang. *classification output layer*), która przyjmuje na wejściu dane przekształcone funkcją aktywacji typu *softmax*. W przypadku klasyfikacji wie-

loklasowej, gdzie każdej z klas odpowiada jedna jednostka ukryta warstwy klasyfikującej, stosowana jest funkcja straty (ang. *loss function*) szacowana na podstawie entropii krzyżowej (ang. *cross-entropy*, CE) definiowanej wzorem:

$$CE = -\frac{1}{N} \sum_{n=1}^N \sum_{i=1}^K p_{ni} \ln q_{ni} \quad (4.5)$$

gdzie N stanowi liczbę obserwacji, K liczbę klas, natomiast p_{ni} jest rzeczywistym rozkładem prawdopodobieństwa przynależności n -tej obserwacji do i -tej klasy a q_{ni} szacowanym prawdopodobieństwem przynależności n -tej obserwacji do i -tej klasy. Wartość q_{ni} jest wartością wyznaczoną za pomocą funkcji aktywacji typu *softmax* ($q_{ni} = \sigma(j)$) [48].

W przypadku niezbalansowanego zbioru danych w poszczególnych klasach stosuje się ważoną entropię krzyżową (ang. *weighted_cross_entropy*, CE_w) [60]. Zmodyfikowana funkcja straty pozwala na zmniejszenie wpływu nierównomiernej liczebności próbek w poszczególnych klasach poprzez określenie wartości wagi straty w_i dla każdej z klas zgodnie ze zmienioną zależnością [12]:

$$CE_w = -\frac{1}{N} \sum_{n=1}^N \sum_{i=1}^K w_i p_{ni} \ln q_{ni}. \quad (4.6)$$

Waga w_i dla danej klasy jest dopełnieniem do jedności ilorazu liczby obserwacji N_i dla danej klasy w zbiorze treningowym i liczby wszystkich obserwacji N w tym zbiorze:

$$w_i = 1 - \frac{N_i}{N}. \quad (4.7)$$

4.3.7 Warstwa klasyfikacji pikseli

Ostatnią warstwę architektury sieci przeznaczonej do segmentacji jest warstwa klasyfikacji pikseli (ang. *pixel classification layer*), za pomocą której wskazywana jest etykieta kategorii piksela lub woksela obrazu przetwarzanego sieciami CNN. Podobnie jak w warstwie klasyfikującej, również w tym przypadku obliczana jest funkcja straty szacująca skuteczność trenowanego modelu względem wartości oczekiwanej. W segmentacji z użyciem CNN jako funkcja straty stosowana jest m.in. entropia krzyżowa (wzór 4.5) oraz ważona entropia krzyżowa (wzór 4.6) [60], które zostały opisane wcześniej. W przypadku obrazów powszechny problem stanowi brak równowagi w liczebności pikseli należących do poszczególnych klas, ponieważ piksele stanowiące segmentowany obiekt zazwyczaj mają znacznie mniejszy procentowy udział w całym obrazie niż piksele będące tłem. Z powodu braku zbalansowanych klas często zamiast entropii krzyżowej jako funkcji straty stosowane są metryki i ich modyfikacje oparte o miary oceny jakości segmentacji [73], szczególnie bazujące na indeksie

Dice'a [159].

Funkcja straty wyznaczana w oparciu o uogólniony indeks Dice'a (ang. *Generalised Dice loss*, DL), umożliwia zbalansowanie zbioru danych za pomocą kontroli udziału każdej z klas (obiekt, tło) poprzez przypisanie im pewnych wag w_k [159]. Funkcja straty DL definiowana jest wzorem [73, 159]:

$$DL = 1 - K \frac{\sum_{k=1}^K \left(w_k \sum_{n=1}^N S_{kn} GT_{kn} \right)}{\sum_{k=1}^K \left(w_k \sum_{n=1}^N (S_{kn} + GT_{kn}) \right)} \quad (4.8)$$

gdzie GT jest obrysem eksperckim (ang. *ground truth*), S stanowi wynik segmentacji, K oznacza liczbę klas, natomiast N odpowiada liczbie elementów wzdłuż pierwszych dwóch wymiarów S . Waga w_k kontroluje udział każdej z klas, zgodnie z zależnością:

$$w_k = \left(\sum_{n=1}^N GT_{kn} \right)^{-2} \quad (4.9)$$

Ważenie funkcji straty pozwala na zmniejszenie wpływu nierównomiernej liczebności pikseli w poszczególnych klasach, czyli większych regionów obrazu, jakim zazwyczaj jest tło, na wartość indeksu Dice'a. Zabieg ten wspiera proces treningu sieci w segmentacji również mniejszych obiektów [159], co jest istotne ze względu na częste trudności z ich segmentacją na skutek redukcji informacji za pomocą warstwy łączenia [113].

4.3.8 Funkcja aktywacji typu ReLU

Rektyfikowana jednostka liniowa (ang. *rectified linear unit*, ReLU) stosowana jest celem wykonania operacji progowania każdego elementu wejścia zgodnie z następującą definicją:

$$f(x) = \max(0, x) = \begin{cases} x & \text{dla } x \geq 0 \\ 0 & \text{dla } x < 0 \end{cases} \quad (4.10)$$

Rektyfikowana jednostka liniowa jest często wykorzystywana w sieciach CNN jako funkcja aktywacji celem wprowadzenia nieliniowości odpowiedzi warstwy splotowej [57, 123]. Zastosowanie funkcji aktywacji typu ReLU aktywuje część ukrytych jednostek, które mają niezerowe wyjście oraz pozwala na rozwiązanie problemu zanikającego gradientu występującego w przypadku strategii uczących opartych o metody gradientowe dla wielu warstw. Jednostka ReLU umożliwia również skrócenie procesu uczenia, ze względu na przyspieszenie zbieżności stochastycznej metody gradientu [114].

4.3.9 Funkcja aktywacji typu *softmax*

Funkcja aktywacji typu *softmax* pozwala wyznaczyć dla każdej jednostki ukrytej należącej do ostatniej warstwy gęstej, prawdopodobieństwo przynależności analizowanego obrazu wejściowego do każdej z rozważanych klas (Rys. 4.8). Warstwa *softmax* wykorzystująca znormalizowaną funkcję wykładniczą definiowana jest następująco:

$$\sigma(j) = \frac{e^{w_j x}}{\sum_{k=1}^K e^{w_k x}} = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}, \quad (4.11)$$

gdzie $z_j = w_j x$ jest znormalizowaną wartością wejścia. Warstwa *softmax* jest rodzajem funkcji aktywacji, dla której istnieje zależność $\sum_{j=1}^K \sigma(j) = 1$. Wartość sumy wynosząca 1 pozwala na traktowanie wartości wyjściowych funkcji aktywacji typu *softmax* jako prawdopodobieństwa przynależności do j -klasy, co umożliwia zastosowanie jej w klasyfikatorach [48].

4.3.10 Mechanizm normalizacji

Mechanizm normalizacji (ang. *batch normalization*, BN) stosowany jest w przypadku małych podzbiorów zbioru treningowego (ang. *mini-batch*). Mechanizm normalizacji, umieszczony pomiędzy warstwą splotową i funkcją aktywacji (np. ReLU), umożliwia przyspieszenie procesu uczenia oraz zmniejszenie wrażliwości treningu na losowe wartości inicjalizacji wag sieci poprzez normalizację aktywacji poprzedniej warstwy [123]. Zastosowanie mechanizmu BN umożliwia również zmniejszenie ryzyka nadmiernego dopasowania modelu do zbioru treningowego [147]. Element x_i podzbioru treningowego poddawany jest normalizacji za pomocą wartości średniej μ_B i wariancji σ_B^2 wyznaczonej dla tego podzbioru przed normalizacją. Znormalizowana wartość aktywacji obliczana jest zatem zgodnie ze wzorem:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}, \quad (4.12)$$

gdzie ϵ stanowi stałą poprawiającą stabilność w przypadku niewielkiej wartości wariancji.

Znormalizowana wartość aktywacji \hat{x}_i modyfikowana jest za pomocą współczynników skali γ i translacji β celem eliminacji wpływu zerowej wartości średniej oraz jednostkowej wariancji, które nie stanowią wystarczającej wartości progowej pobudzenia, zgodnie z zależnością [72]:

$$y_i = \gamma \hat{x}_i + \beta. \quad (4.13)$$

4.3.11 Regularyzator typu *dropout*

Regularyzator typu *dropout* jest techniką regularyzacji opuszczeń stosowaną w sieciach CNN. Podobnie jak w warstwach łączenia i mechanizmie normalizacji, użycie regularyzatora pozwala na uniknięcie nadmiernego dopasowania trenowanego modelu do danych w zbiorze treningowym. Podczas treningu, regularyzator typu *dropout* zeruje wagi ukrytych jednostek (neuronów) z danym prawdopodobieństwem p , co pozwala na ekstrakcję cech ze zbioru treningowego za pomocą pozostałych jednostek i eliminację zjawiska przeuczenia (ang. *overfitting*) oraz na przyspieszenie procesu nauki [83, 157].

Regularyzator typu *dropout* stanowi również skuteczny sposób łączenia kilku sieci neuronowych, gdy dochodzi do losowego wybierania kilku ukrytych jednostek dla każdego z testowanych przypadków. W ten sposób powstają różne architektury dla każdego z przypadków. Technika ta stanowi ekstremalny przypadek uśredniania modelu [147].

4.4 Sieć CNN klasyfikująca obecność igły

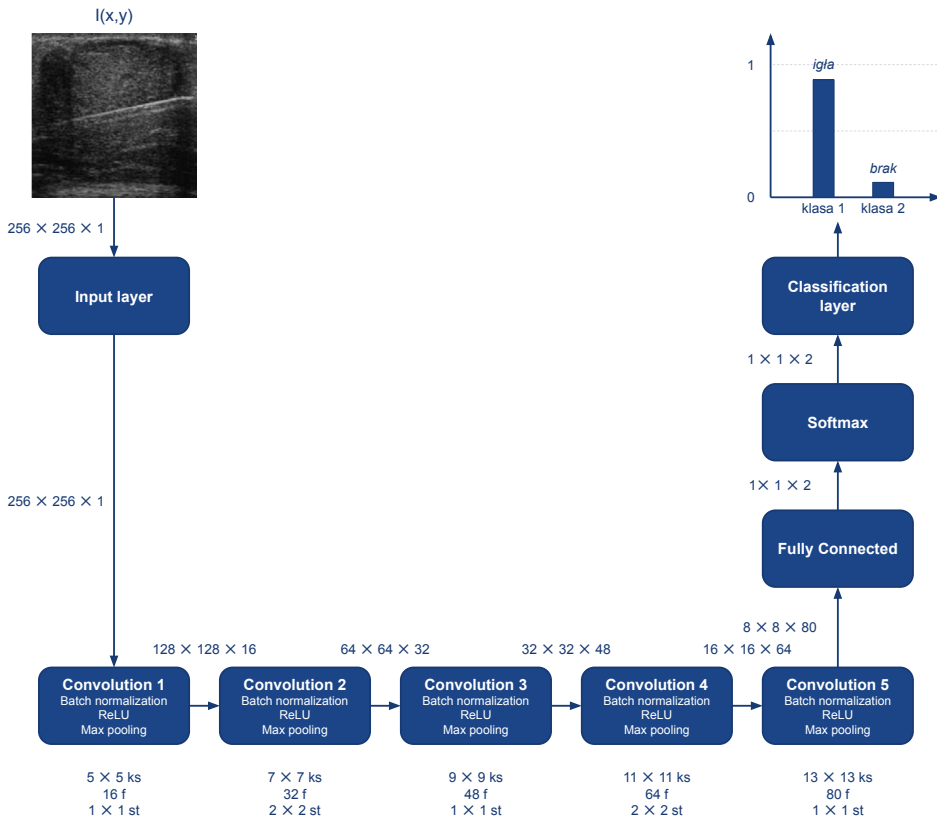
Przed określeniem lokalizacji igły biopsyjnej na obrazie USG konieczne jest określenie, czy igła w ogóle na tym obrazie jest widoczna. Do klasyfikacji obrazów USG pod kątem obecności igły zaimplementowano sieć CNN, której opis przedstawiono w dalszej części niniejszego podrozdziału.

4.4.1 Architektura sieci klasyfikującej obecność igły

Architektura sieci klasyfikującej obecność igły została oparta o architekturę jednogałęziową (Rys. 4.9). Sieć ta cechuje się efektywnością i na tyle szybkim czasem przetwarzania dwuwymiarowych obrazów w skali szarości, że możliwe jest zastosowanie klasyfikatora w czasie rzeczywistym na strumieniu obrazów z aparatu USG.

Architektura opisywanej sieci klasyfikującej składa się z warstwy wejściowej, warstw ukrytych oraz warstwy wyjściowej. Warstwa wejściowa reprezentuje jednokanałowy obraz wejściowy o wymiarach 256×256 . Za pomocą warstwy wyjściowej będącej równocześnie warstwą klasyfikującą [48], realizowana jest klasyfikacja binarna przynależności obrazu do jednej z dwóch klas: obecność igły albo brak igły.

Liczba warstw ukrytych opisywanej sieci wynosi pięć. Każda z warstw ukrytych składa się z warstwy konwolucyjnej, mechanizmu normalizacji, funkcji aktywacji typu ReLU oraz warstwy głosującej. Operacja splotu w pierwszej warstwie konwolucyjnej realizowana jest za pomocą jądra o rozmiarze 3×3 . W konwolucjach kolejnych warstwach ukrytych parametr rozmiaru jądra (ang. *kernel size*, ks) zwiększany jest każdorazowo o dwa. Wartość kroku (ang. *stride*, s)



Rys. 4.9: Architektura sieci CNN klasyfikującej obecność igły. Oznaczenia: ks – rozmiar jądra (ang. *kernel size*), f – liczba filtrów (ang. *filter*), st – krok (ang. *stride*). Obok strzałek podane są wymiary danych wchodzących do kolejnego bloku przetwarzania (szerokość \times wysokość \times liczba kanałów).

z jakim przesuwane jest jądro, wynosi 1×1 w pierwszej, trzeciej i piątej warstwie ukrytej, natomiast w drugiej i czwartej wynosi 2×2 . Liczba filtrów, na podstawie których generowane są mapy cech wyższego rzędu, stanowi w kolejnych warstwach następujące po sobie wielokrotności liczby 16. Inicjalizacja wartości trenowanych wag realizowana jest za pomocą metody zaproponowanej przez Glorota i wsp. [45]. Redukcja przestrzenna ekstrahowanych cech przeprowadzana jest z użyciem warstwy łączenia typującej wartości maksimum dla regionów o wymiarach 3×3 . Krok operacji łączenia w warstwie głośującej, w przeciwieństwie do warstw konwolucyjnych, w pierwszej, trzeciej i piątej warstwie ukrytej przyjmuje wartość 2×2 , natomiast w drugiej i czwartej 1×1 . Wartości te zostały dobrane celem lokalnej detekcji większych obiektów.

Klasyfikacja treści obrazu realizowana jest za pomocą warstw: w pełni połączonych (gęstej) z funkcją aktywacji typu *softmax* i klasyfikującej. Liczba jed-

nostek ukrytych warstwy w pełni połączonej wynosiła 5120. Funkcja straty została zdefiniowana za pomocą zmodyfikowanej ważonej entropii krzyżowej CE_w ze względu na nierównomierną liczebność klas.

W każdej z warstw ukrytych umieszczony został mechanizm normalizacji oraz funkcja aktywacji typu ReLU celem przyspieszenia treningu sieci. Mechanizm normalizacji BN zastosowany został również jako jedna z technik regularyzacji, która pozwala uniknąć nadmiernego dopasowania modelu do danych zbioru treningowego. Innymi regularyzatorami zapobiegającymi przeuczeniu modelu w przypadku autorskiej architektury sieci klasyfikującej są warstwy łączenia oraz trzy regularyzatory typu *dropout* z zadaniem prawdopodobieństwem p zerowania wag ukrytych jednostek oszacowanym na poziomie 10%. Regularyzatory typu *dropout* zostały umieszczone w pierwszej, trzeciej i piątej warstwie ukrytej.

4.4.2 Zbiory treningowe, walidacyjne i testowe

Przygotowanie danych do poprawnego wytrenowania modelu sieci klasyfikującej obejmuje ich podział na zbiór treningowy i testowy oraz, w technikach głębokiego uczenia, na zbiór walidacyjny. Zbiór walidacyjny wykorzystywany jest podczas procesu treningu do korygowania szacowanych wartości wag. Ocena działania modelu podczas treningu z danymi ze zbioru walidacyjnego wykorzystywana jest do określania optymalnego zestawu wartości wag, to jest takiego, dla którego wyniki uzyskiwane przez model są najwyższe.

Użycie zbioru walidacyjnego pozwala w pewnym stopniu wyeliminować ryzyko nadmiernego dopasowania modelu do danych ze zbioru treningowego, jednak, z drugiej strony, może wystąpić sytuacja nadmiernego dopasowania modelu do zbioru walidacyjnego. Z tego powodu konieczne jest określanie efektywności wytrenowanego modelu na takim zbiorze testowym, który nie był używany na etapie nauki tego modelu [12, 147].

Zbiór obrazów USG przygotowany do niniejszej pracy (opisany szczegółowo dalej w podrozdziale 5.2.1) został poddany losowemu podziałowi na zbiory treningowy, walidacyjny i testowy na podstawie unikatowych identyfikatorów pacjentów. Efektywność modelu oceniana była w oparciu o zbiór testowy. Podział został przeprowadzony zgodnie ze strategią k -krotniej walidacji krzyżowej (ang. *k-fold cross-validation*), która umożliwia losowe przydzielenie danych do poszczególnych zbiorów, przy założeniu, że każdy identyfikator pacjenta tylko raz będzie użyty w zbiorze testowym. Podejście to pozwala na wiarygodną weryfikację architektury pod względem dopasowywania do określonego typu danych [12].

Zgodnie z tą metodologią piętnastokrotnie powtórzono losowy podział obrazów źródłowych na trzy zbiory, przez co uzyskano piętnaście zestawów zbiorów treningowego, walidacyjnego i testowego. Pozwoliło to na trening i przete-

stowanie piętnastu modeli opartych o taką samą architekturę.

Te same zestawy danych wykorzystano zarówno w sieciach klasyfikujących obecność igły biopsyjnej, jak i w sieciach segmentujących tę igłę. W tym drugim przypadku zawężono zbiór danych tylko do tych obrazów, które zostały zakwalifikowane jako zawierające igłę.

4.4.3 Augmentacja danych

Poprawnie przeprowadzony proces treningu sieci CNN wymaga licznego zbioru danych źródłowych. Bazy danych zawierające od kilkuset do kilku tysięcy elementów poddawane są procesowi augmentacji w celu ich powiększenia [147]. Mały zbiór danych źródłowych zwykle wpływa negatywnie na proces treningu, ze względu na nadmierne dopasowanie (przeuczenie) modelu do danych treningowych. W celu eliminacji zjawiska przeuczenia sieci klasyfikującej stosuje się jedną z metod regularyzacji czyli technikę augmentacji.

Zbiory źródłowych danych obrazowych rozszerzane są między innymi poprzez zastosowanie złożenia elementarnych przekształceń geometrycznych, takich jak rotacji, odbić lustrzanych, translacji i skalowań [147]. Innym podejściem jest modyfikacja danych oryginalnych poprzez wprowadzenie do obrazów treningowych dodatkowych szumów albo zmianie kontrastu, jasności, balansu bieli, wyostrozaniu lub rozmywaniu krawędzi czy wyrównywaniu histogramu [101]. Oryginalne oraz nowo utworzone obrazy stanowią powiększony zbiór danych treningowych.

W niniejszej pracy przygotowana baza danych obrazowych (Tab. 5.6) została poddana augmentacji geometrycznej oraz augmentacji przez filtrację celem rozszerzenia liczebności zbioru treningowego.

Augmentacja geometryczna

Augmentacja geometryczna obejmowała translację obrazów o losową liczbę pikseli z zakresu ± 8 w każdym kierunku, skalowanie obrazów losowymi wartościami z zakresu 0,8–1,2, a także odbicie wzdłuż poziomej osi. Nie stosowano odbicia w pionie ani rotacji obrazu ze względu na charakter obrazów ultrasonograficznych przedstawiających biopsję, gdyż wprowadzenie rotacji skutkowałoby zaburzeniem relacji pomiędzy położeniem igły a czołem głowicy. Również charakterystyczny hyperechogeniczny obraz skóry przy samym czole głowicy nie powinien zmieniać swojego położenia.

Augmentacja przez filtrację

Augmentacja przez filtrację (Rys. 4.11) została użyta w celu zastosowania techniki *fine-tuning*, opisaney w podrozdziale 4.4.4. W augmentacji tej każdy z obrazów ze zbioru danych został przetworzony niezależnie za pomocą trzech

filtrów. Dla każdego z nich został zapisany osobny obraz wynikowy, czyli dla każdego obrazu ze zbioru wygenerowane zostały trzy nowe obrazy, tworząc tym samym drugą, trzykrotnie większą bazę obrazów. Zastosowanymi metodami filtrującymi była entropia, filtr *shock* oraz metoda średnich nielokalnych.

Dla obrazu w skali szarości, którego piksele mogą przyjmować wartości intensywności w zakresie od 0 do $n - 1$, entropia E stanowi miarę statystyczną, która jest zdefiniowana następująco:

$$E(i) = \sum_{i=0}^{n-1} P(i) \cdot \log_2 \frac{1}{P(i)} = - \sum_{i=0}^{n-1} P(i) \cdot \log_2 P(i) \quad (4.14)$$

gdzie $P(i)$ stanowi prawdopodobieństwo wystąpienia wartości intensywności piksela i w całym obrazie. Ponadto, gdy w obrazie są nieobecne piksele o pewnej intensywności i czyli $P(i) = 0$, wówczas w miejsce iloczynu $0 \cdot \log_2 0$ przyjmuje się wartość 0.

Entropia stosowana jest w ocenie charakterystyki tekstury obrazu i w pewnym stopniu określa stopień złożoności obrazu. Im więcej szczegółów widocznych jest w obrazie, tym wartość entropii jest wyższa. Z powyższego wzoru wynika, że największą wartość przyjmuje entropia obrazu, którego częstotliwości występowania wszystkich intensywności piksela są takie same, a najmniejszą, gdy obraz zawiera jeden poziom intensywności.

Filtracja z wykorzystaniem entropii umożliwia lokalną analizę stopnia skomplikowania obrazu (Rys. 4.10b), dzięki czemu znalazła ona zastosowanie w przetwarzaniu obrazów medycznych [25, 94, 153]. Obrazy pochodzące z bazy danych (Tab. 5.6) poddano filtracji z oknem o rozmiarze 11×11 pikseli.

Każdy z obrazów w zbiorze danych został przetworzony również za pomocą filtra *shock* (ang. *shock filter*), który często stosowany jest w celu redukcji szumów i wyostrenia krawędzi obiektów widocznych w obrazach ultrasonograficznych, RTG, tomografii komputerowej czy rezonansie magnetycznym [77, 171, 181].

Filtr ten wykorzystuje proces lokalnej operacji dylatacji i erozji analizowanego piksela obrazu w celu modyfikowania jego wartości zgodnie z zależnościami:

$$f^t(x_i, y_j) = \begin{cases} I^{t-1}(f(x_i, y_j)) \oplus D & \text{gdy } \Delta f^{t-1}(x_i, y_j) < 0 \\ I^{t-1}(f(x_i, y_j)) \ominus D & \text{gdy } \Delta f^{t-1}(x_i, y_j) > 0 \end{cases}, \quad (4.15)$$

gdzie dla każdego piksela $f^t(x_i, y_i)$ obrazu $f^t(x_i, y_j)$ w iteracji t , $\Delta f^t(x_i, y_i)$ jest operatorem Laplace'a na pikselu (x_i, y_i) . D jest elementem strukturalnym w kształcie dysku o promieniu 1. Symbole \oplus i \ominus oznaczają odpowiednio dylatację i erozję. Parametry filtracji stanowią rozmiar okna filtra oraz liczba iteracji.

W efekcie filtracji obraz igły zostaje wyostrzony, a część szumów wytłumiona (Rys. 4.10d). W przypadku przetwarzanych obrazów z dostępnej bazy danych (Tab. 5.6) zastosowano rozmiar okna 9×9 oraz 10 iteracji.

Obrazy w zbiorze danych poddano również przetwarzaniu metodą średnich nielokalnych (ang. *non-local means*, *NLM*) [16]. Filtr ten umożliwia usunięcie szumu z obrazu przy zachowaniu ostrości krawędzi obiektów widocznych w obrazie. Filtr średnich nielokalnych wyznacza wartość analizowanego piksela $NL(i)$ na podstawie wartości średniej ważonej wszystkich pikseli w zaszumionym obrazie $v = \{v(i) | i \in I\}$ zgodnie ze wzorem:

$$NL(i) = \sum_{j \in I} w(i, j)v(j) \quad (4.16)$$

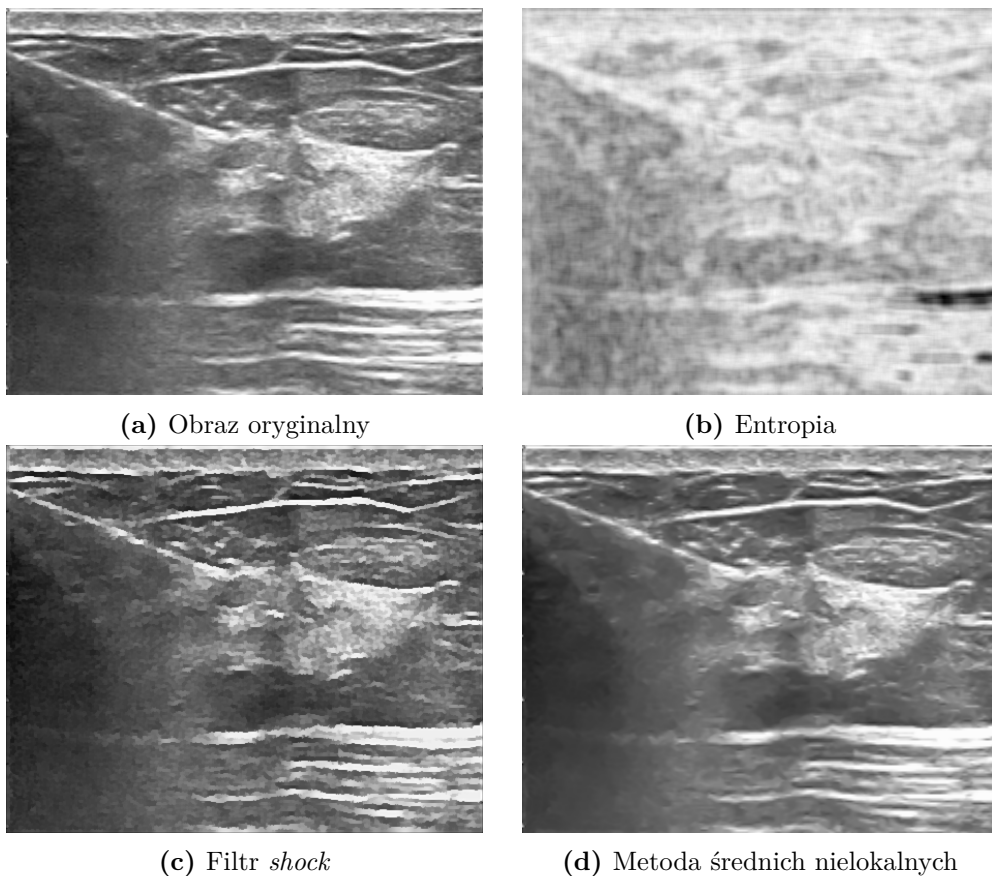
gdzie rodzina wag $\{w(i, j)\}_j$ jest miarą opartą o podobieństwo pomiędzy pikselami i i j , a także spełnia zależności $0 \leq w(i, j) \leq 1$ oraz $\sum_j w(i, j) = 1$. Parametrami filtracji są stopień wygładzenia, rozmiar okna wyszukiwania oraz rozmiar okna podobieństwa. Filtracja *NLM* pozwala na wytłumienie szumów bez rozmycia krawędzi igły (Rys. 4.10d). Obrazy ze źródłowej bazy danych (Tab. 5.6) przetwarzano ze stopniem wygładzenia równym 5 dla okna wyszukiwania o rozmiarze 21×21 , z oknem podobieństwa wynoszącym 5×5 .

Wyniki działania każdego z opisanych trzech filtrów na obrazie oryginalnym (Rys. 4.10a) przedstawiono na rysunkach 4.10b, 4.10c i 4.10d odpowiednio dla przetwarzania za pomocą entropii, filtrem *shock* oraz metodą średnich nielokalnych. Obrazy wynikowe przetwarzania opisanymi filtrami zostały połączone w jedną bazę danych, trzykrotnie większą w porównaniu do bazy oryginalnej. Zbiór przefiltrowanych obrazów poddano podziałowi z zachowaniem identyfikatorów pacjentów określonych metodą walidacji krzyżowej opisaną w podrozdziale 4.4.2 celem wykorzystania ich w pierwszym etapie treningu metodą *fine-tuning* proponowanego klasyfikatora oceniającego obecność igły biopsyjnej w dwuwymiarowym obrazie ultrasonograficznym. Proces treningu sieci klasyfikującej został opisany w podrozdziale 4.4.4.

4.4.4 Proces treningu sieci klasyfikującej

Architektura sieci klasyfikującej, zaprojektowana i opisana w niniejszej rozprawie, została poddana procesowi treningu, którego celem było dobranie optymalnych wartości wag minimalizujących funkcję straty (kosztu). Architektura projektowana z nastawieniem na skrócenie czasu uczenia względem istniejących już rozwiązań.

Powszechnie stosowanymi metodami optymalizacji sieci CNN są modyfikacje metody spadku wzdłuż gradientu (ang. *stochastic gradient descent*, *SGD*), wśród których najczęściej stosowane są modyfikacje *SGDM* i *ADAM*. Algorytm stochastycznego zejścia gradientu z pędem (ang. *stochastic gradient de-*



Rys. 4.10: Techniki augmentacji przez filtrację. Obraz oryginalny (4.10a), lokalna entropia obrazu w skali szarości (4.10b), filtr *shock* (4.10c) oraz filtracja metodą średnich nielokalnych (4.10d).

scent with momentum, SGDM) aktualizuje wartości wag na podstawie ostatniej dokonanej wartości wag [108]. Optymalizacja metodą szacowania momentu adaptacyjnego (ang. *adaptive moment estimation*, ADAM) zakłada aktualizację wag po każdej obserwacji wykorzystanej w treningu. W przypadku tego optymalizatora wagi aktualizowane są na podstawie wykładniczej średniej ruchomej gradientów. Metoda ta umożliwi efektywną optymalizację przy niewielkim zapotrzebowaniu na pamięć [81].

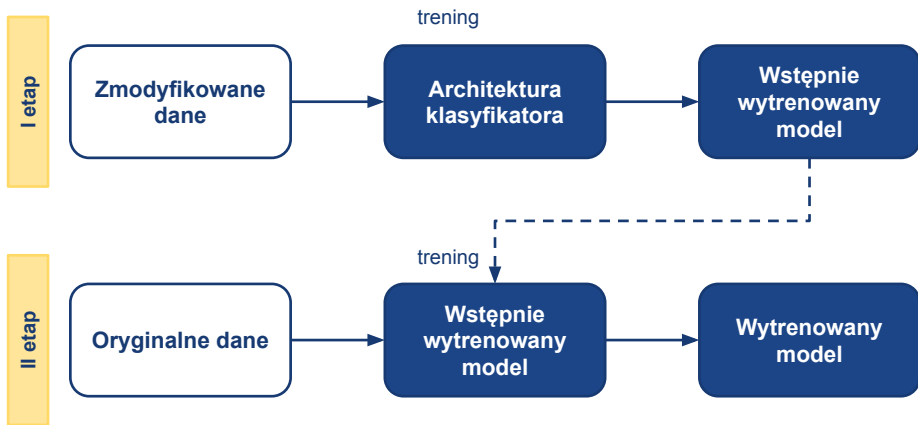
W opisywanym podejściu zastosowano optymalizator ADAM ze względu na małą liczbę warstw ukrytych zaprojektowanej sieci [112]. Hiperparametr optymalizatora ADAM czyli współczynnik zaniku gradientu B_1 został ustalony na poziomie 0,9. Zbiór treningowy podzielono na podzbiory (ang. *mini-batch*), co przelożyło się na zmniejszenie poziomu wariancji. Harmonogramowanie współczynnika uczenia celem zmniejszenia tempa w miarę postępu procesu treningu

uwzględniało początkowy współczynnik uczenia (ang. *initial learn rate*) równy 0,001 ze współczynnikiem spadku szybkości uczenia na poziomie 0,15. Tempo treningu aktualizowano co dziesięć iteracji.

Mechanizm inicjalizacji wag w poszczególnych warstwach konwolucyjnych i w warstwie w pełni połączonej stanowi istotny element w początkowym etapie procesu treningu. Niewłaściwie dobrana początkowa wartość wag może skutkować brakiem zbieżności procesu lub wydłużeniem czasu treningu. Efektem niepoprawnej inicjalizacji również może być inny zestaw końcowy wag, który cechuje się różną skutecznością dla modeli o tej samej architekturze i stałych hiperparametrach [57]. Eliminacja negatywnego wpływu doboru wag startowych osiągnięta jest za pomocą ich losowania z uwzględnieniem typu funkcji aktywacji używanych w inicjowanych węzłach (wartość wariancji funkcji aktywacji oraz wariancja obliczanych gradientów powinny być zbliżone). W proponowanej architekturze sieci klasyfikującej zarówno w warstwach konwolucyjnych jak i w pełni połączonej inicjalizacja wag realizowana była metodą opisaną przez Glorot i wsp. [45].

W celu eliminacji nadmiernego dopasowania trenowanego modelu do danych w zbiorze treningowym, poza mechanizmami typu *dropout* i normalizacji, zastosowana została również metoda regularyzacji L_2 dla warstw konwolucyjnych i w pełni połączonej [108]. Wykorzystanie tej techniki skutkuje dodaniem kary do funkcji straty [12]. Wartość parametru L_2 dla wszystkich warstw konwolucyjnych i w pełni połączonej była równa 0,01. Kolejny mechanizm mający zapobiec przeuczeniu modelu to metoda wcześniejszego zatrzymania (ang. *early stopping*) [200]. Metoda ta zakłada analizę zmian wartości funkcji straty wyznaczanej dla zbioru walidacyjnego. Jeżeli wartość funkcji straty wzrasta – sugerując przeuczenie modelu danymi pochodzącymi ze zbioru walidacyjnego – proces treningu zostaje przerwany.

Zaprojektowana architektura sieci klasyfikującej została również poddana treningowi techniką dostrajania (ang. *fine-tuning*) ze względu na małą bazę danych obrazowych [147]. W pierwszym etapie model uczony był danymi przekształcanymi geometrycznie oraz augmentowanymi przez filtrację zgodnie z opisem zawartym w podrozdziale 4.4.3. Następnie wstępnie wytrenowany model został poddany ponownemu procesowi treningu, tym razem z wykorzystaniem oryginalnego zbioru obrazów, które zostały poddane jedynie augmentacji geometrycznej. Dostrajanie modelu skutkuje jego lepszą efektywnością, o ile dane wykorzystane do przygotowania pretrenowanego modelu różnią się od danych treningowych wykorzystywanych w procesie treningu modelu finalnego [147]. Schemat treningu techniką dostrajania został przedstawiony na rysunku 4.11.



Rys. 4.11: Schemat treningu klasyfikatora obecności struktury w obrazach USG techniką dostrajania. W pierwszym etapie architektura jest trenowana i walidowana za pomocą zmodyfikowanego (z użyciem metod filtracji) zbioru obrazów USG. Następnie pretrenowany model poddawany jest nauce po raz drugi z oryginalnymi obrazami ultrasonograficznymi.



Rys. 4.12: Przykłady typów obrazów ultrasonograficznych gruczołu sutkowego.

4.5 Sieć CNN klasyfikująca zmianę ogniskową

Klasyfikacja obecności zmiany ogniskowej w obrazie ultrasonograficznym jest procedurą pozwalającą wyeliminować z dalszej analizy obrazy, na których nie jest widoczna zmiana ogniskowa. Przykładem są obrazy zawierające jedynie zdrową tkankę (Rys. 4.12a) lub obrazy zapisywane w chwili, gdy głowica nie przylega do powierzchni skóry. Ponadto dzięki klasyfikacji możliwe jest określenie typu zmiany ogniskowej: echodatniego (Rys. 4.12b) lub echoujemnego (Rys. 4.12c). Do każdego typu zmiany można dopasować odpowiednią metodę segmentacji.

4.5.1 Architektura sieci klasyfikującej obecność i typ zmiany ogniskowej

Architektura sieci klasyfikującej obecność i typ zmiany ogniskowej w obrazie USG jest zgodna z architekturą wykorzystaną do klasyfikacji obecności igły, przedstawionej wcześniej na schemacie 4.9. Jednogałęziowa struktura została w tym przypadku zmodyfikowana w obrębie warstwy klasyfikującej. O ile w przypadku igły biopsyjnej występowała klasyfikacja binarna samej obecności igły w obrazie, tak w przypadku klasyfikacji zmiany ogniskowej występuje klasyfikacja wieloklasowa do jednej z trzech klas: zmiana echododatnia, zmiana echoujemna albo brak zmiany.

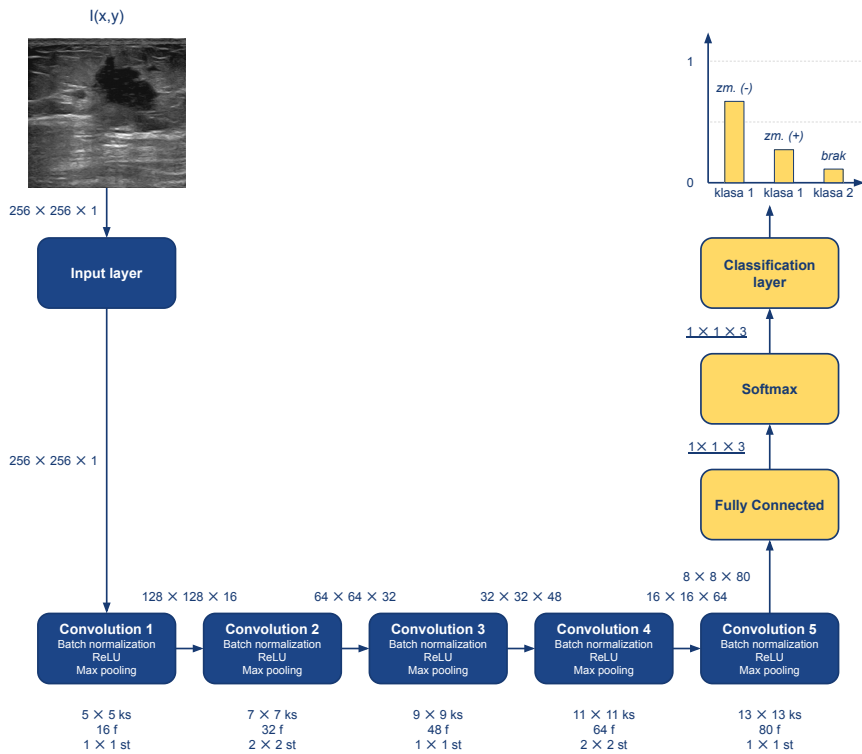
Obraz wejściowy o wymiarach 256×256 przetwarzany był za pomocą pięciu warstw ukrytych, gdzie każda z warstw składała się z warstwy konwolucyjnej, mechanizmu normalizacji, funkcji aktywacji typu ReLU oraz warstwy głosującej. Szczegóły dotyczące rozmiaru jądra, liczby map cech oraz kroku przedstawiono na schemacie 4.13.

Klasyfikacja treści obrazu realizowana była za pomocą warstwy w pełni połączonej (gęstej) z funkcją aktywacji typu *softmax* oraz warstwy klasyfikującej. Funkcję straty szacowano za pomocą ważonej entropii krzyżowej CE_w .

Tak jak w przypadku architektury klasyfikującej obecność igły, zastosowano mechanizmy przyspieszające trening takie jak mechanizm normalizacji oraz funkcję aktywacji typu ReLU. Ryzyko wystąpienia nadmiernego dopasowania sieci do zbioru danych regulowane było za pomocą mechanizmów normalizacji, warstw łączenia oraz regularyzatorów typu *dropout* z zadaniem prawdopodobieństwem p zerowania wag jednostek ukrytych ustalonym na poziomie 10%.

4.5.2 Zbiory treningowe, walidacyjne i testowe

Zbiór dwuwymiarowych obrazów ultrasonograficznych przygotowany celem przeprowadzenia procesu uczenia i testowania architektury klasyfikującej został szczegółowo opisany w podrozdziale 5.8. Dane zostały podzielone na trzy klasy: zmiana echoujemna, zmiana echododatnia oraz brak zmiany. Połączony zbiór danych został poddany podziałowi na trzy podzbiory: treningowy, walidacyjny oraz testowy na poziomie unikalnych identyfikatorów pacjentów, zgodnie ze strategią k -krotnej walidacji krzyżowej (ang. *k-fold cross-validation*, *CV*). Strategia ta pozwala na odseparowanie danych pomiędzy etapami treningu i testowania oraz pozwala na obiektywną i wiarygodną ocenę skuteczności sieci ze względu na wykorzystanie wszystkich danych (obarczonych różnymi artefaktami) w kolejnych iteracjach testowania [12, 147]. Zgodnie z zasadami walidacji krzyżowej dane podzielono losowo piętnastokrotnie, co pozwoliło na przeprowadzenie 15 eksperymentów uwzględniających proces uczenia z wykorzystaniem zbioru treningowego i walidacyjnego oraz niezależny etap ewaluacji modelu za pomocą zbioru testowego.



Rys. 4.13: Architektura sieci CNN wieloklasowej klasyfikującej typ zmiany ogniskowej. Oznaczenia: ks – rozmiar jądra (ang. *kernel size*), f – liczba map cech (ang. *feature map*), st – krok (ang. *stride*). Obok strzałek podane są wymiary danych wchodzących do kolejnego bloku przetwarzania (szerokość \times wysokość \times liczba kanałów). Istotne zmiany względem klasyfikatora binarnego zaznaczono kolorem żółtym.

4.5.3 Augmentacja danych

Zbiór danych wykorzystywany w przypadku zmiany ogniskowej stanowi przykład małego zbioru danych [147]. Dane, podobnie jak w przypadku zbioru przygotowywanego do klasyfikacji obecności igły biopsyjnej, zostały poddane rozszerzeniu z wykorzystaniem augmentacji geometrycznej (Podrozdz. 4.4.3) oraz przez filtrację (Podrozdz. 4.4.3).

W przypadku augmentacji geometrycznej obrazu poddano przetwarzaniu za pomocą translacji o losową liczbę pikseli z zakresu ± 8 w każdym kierunku, skalowaniu wartościami losowanymi z przedziału 0,8–1,2 i losowemu odbiciu względem osi pionowej obrazu.

Przygotowany został również zbiór danych, które zostały poddane filtracji metodami szczegółowo opisanymi w podrozdziale 4.4.3: filtracją za pomocą entropii, filtrem *shock* oraz metodą średnich nielokalnych. Każdy z oryginalnych

obrazów został niezależnie poddany filtracji każdą z wymienionych metod. Podejście to pozwoliło uzyskać zbiór trzykrotnie większy w odniesieniu do oryginalnego, który został wykorzystany w pierwszym etapie treningu architektury metodą dostrajania, w którym to etapie dane były również augmentowane geometrycznie zgodnie z powyższymi parametrami. Drugi etap treningu metodą *fine-tuning* zakładał wykorzystanie zbioru obrazów oryginalnych rozszerzanych jedynie geometrycznie.

4.5.4 Proces treningu sieci klasyfikującej

Architektura sieci klasyfikującej obecność i typ zmiany ogniskowej została poddana procesowi uczenia analogicznie do klasyfikacji igły biopsyjnej (Podrozdz. 4.4.4).

Proponowana architektura sieci poddana została treningowi metodą dostrajania ze względu na klasyfikację wieloklasową ograniczonym zbiorem danych [147], zgodnie z ogólnym schematem przedstawionym na rysunku 4.11. Autorska architektura klasyfikująca obrazy ultrasonograficzne wykorzystuje optymalizację metodą szacowania momentu adaptacyjnego – ADAM [81] ze względu na małą pojemność proponowanego rozwiązania. Dobór wartości hiperparametrów, takich jak współczynnik zaniku gradientu, początkowy współczynnik uczenia oraz współczynnik spadku uczenia, zostały dobrane na podstawie eksperymentów (opisanych w podrozdziale 5.3.1) i ustalone odpowiednio na poziomie 0,9, 0,001 oraz 0,15 dla obrazu o wymiarach początkowych 256 ± 256 .

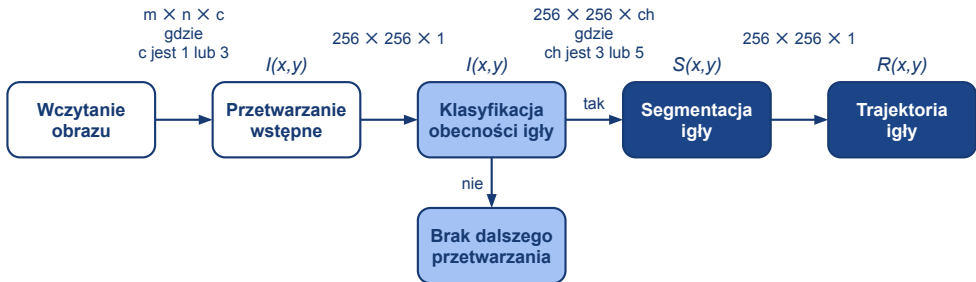
Inicjalizacja wag w poszczególnych warstwach ukrytych w pierwszym etapie treningu realizowana była za pomocą metody zaproponowanej przez Glorota i wsp. [45] celem zapewnienia poprawności inicjalizacji i zbieżności procesu. W drugim etapie treningu za wartości początkowe przyjęto wartości końcowe wag etapu pierwszego.

Zarówno w pierwszym jak i drugim etapie treningu wdrożono mechanizmy minimalizujące wpływ nadmiernego dopasowania trenowanego modelu do danych w zbiorze treningowym w postaci mechanizmu normalizacji oraz regularyzatorów typu *dropout* i L_2 [12, 108]. Każdy z tych mechanizmów oddziałuje na wartości wag w trakcie całego procesu uczenia. Wdrożono również mechanizm wcześniejszego zatrzymania (ang. *early stopping*) na każdym z etapów treningu [200], który przerywa proces treningu w przypadku wzrostu wartości funkcji straty.

4.6 Sieć CNN do segmentacji semantycznej regionu igły biopsyjnej

Automatyzacja algorytmów lokalizacji i śledzenia igieł jest zabiegiem ceLOWANYM ze względu na potencjalne zmniejszenie liczby wykonywanych na-

kluc, skrócenie czasu zabiegu, ograniczenie ryzyka uszkodzenia tkanek i nerwów [129], jak również pomaga śledzić narząd, który został obciążony ryzykiem rozprzestrzenienia się komórek nowotworowych podczas usuwania igły biopsyjnej [1]. W metodach programowych śledzenia igły biopsyjnej brakuje prac opartych na danych klinicznych, które pozwoliłyby na wiarygodne i powtarzalne lokalizowanie jej położenia. W poniższym podrozdziale przedstawiona jest autorska metoda szacowania trajektorii igły biopsyjnej oparta o proces uczenia i testowanie unikalnym, heterogenicznym zbiorem danych klinicznych.



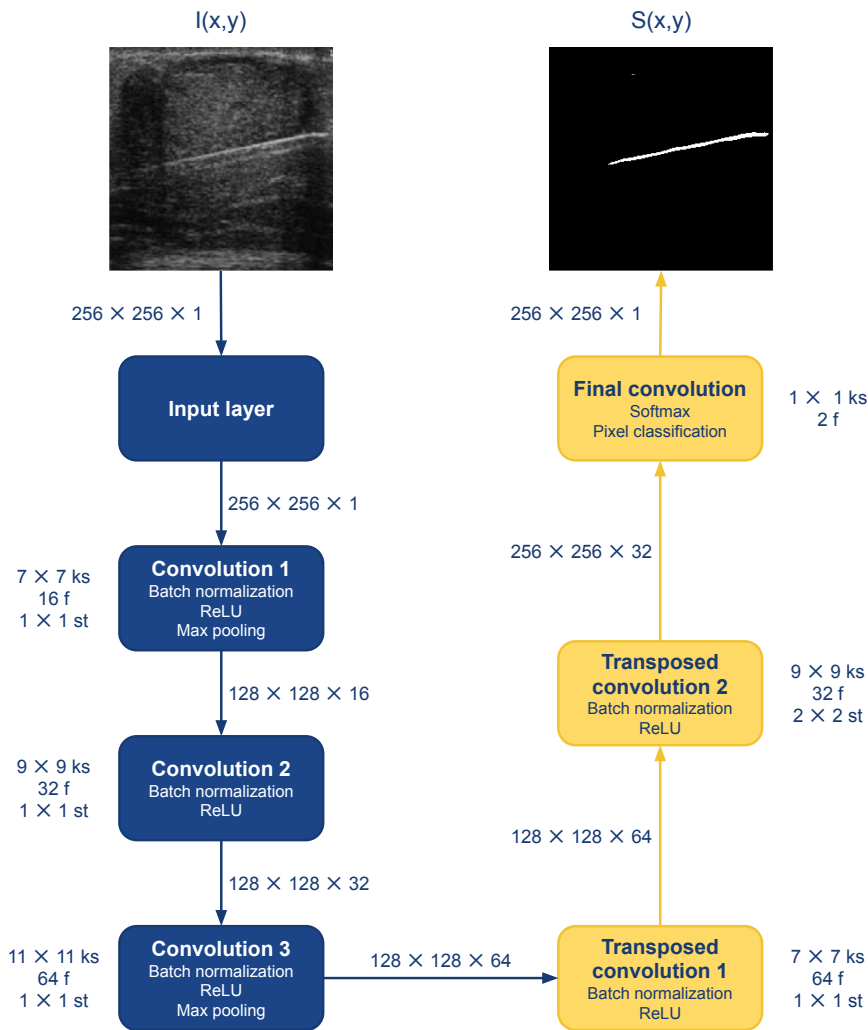
Rys. 4.14: Schemat proponowanej metody detekcji i lokalizacji igły biopsyjnej w dwuwymiarowych obrazach ultrasonograficznych.

4.6.1 Architektura sieci do segmentacji semantycznej

W trakcie przeglądu istniejących sieci konwolucyjnych, poddano testom różne typy architektury segmentacji semantycznej (U-Net [57, 138], SegNet [7, 57]). Na ich podstawie została przygotowana autorska metoda segmentacji zarówno igły biopsyjnej, jak i zmian ogniskowych w obrazach ultrasonograficznych. Autorskie rozwiązanie przedstawione na rysunku 4.15 cechujące się efektywnością i niewielkim czasem przetwarzania dwuwymiarowego obrazu w skali szarości oparte zostało o architekturę jednogłęziową.

Autorska architektura sieci segmentującej w dwuwymiarowym obrazie ultrasonograficznym zbudowana została z warstwy wejściowej, warstw ukrytych oraz warstwy wyjściowej. Warstwa wejściowa stanowi reprezentację obrazu w skali szarości $I(x, y)$ o wymiarach $256 \times 256 \times 1$ pikseli. Wynik działania warstwy wyjściowej stanowi binarna maska $S(x, y)$ o tych samych wymiarach.

Segmentacja semantyczna za pomocą konwolucyjnej sieci neuronowej realizowana jest w dwóch kolejnych etapach: konwolucji i transponowanej konwolucji. Etap konwolucji realizowany jest przez trzy warstwy ukryte. Każda z warstw ukrytych składa się z warstwy konwolucyjnej, mechanizmu normalizacji i funkcji aktywacji wykorzystującej rektyfikowaną jednostkę liniową. Operacja splotu realizowana jest kolejno za pomocą jąder o rozmiarach 7×7 , 9×9 i 11×11 w kolejnych warstwach ukrytych. Rozmiar jądra został dobrany celem



Rys. 4.15: Architektura konwolucyjnej sieci neuronowej segmentującej struktury na obrazie USG. Oznaczenia: ks – rozmiar jądra (ang. *kernel size*), f – liczba filtrów (ang. *filter*), st – krok (ang. *stride*). Obok strzałek podane są wymiary danych wchodzących do kolejnego bloku przetwarzania (szerokość \times wysokość \times liczba kanałów).

lokalnego wykrycia większych obiektów. Liczba filtrów dla kolejnych warstw konwolucyjnych (ang. *filter*, f), na podstawie których generowane są mapy cech wyższego rzędu, jest równa kolejnym potęgom liczby dwa począwszy od 16. Wartości trenowanych wag inicjalizowane są za pomocą metody Glorota i wsp. [45]. Wartość parametru kroku (ang. *stride*, s) z jakim przesuwane jest jądro wynosi 1×1 w każdej z warstw ukrytych. Po pierwszej i trzeciej warstwie ukrytej następuje warstwa łączenia typująca wartości maksimum filtrem

3×3 z krokiem 2×2 . Wypełnienie (ang. *padding*) przyjmuje te same wartości co piksele brzegowe obrazu.

Etap transponowanej konwolucji oparty jest o dwie warstwy ukryte. Po każdej z nich następuje mechanizm normalizacji oraz funkcja aktywacji ReLU. Operacja transponowanego splotu umożliwia generowanie wyjściowej mapy cech o wyższej rozdzielczości, co jest niezbędnym krokiem w przygotowaniu binarnej maski wynikowej będącej efektem segmentacji semantycznej. Splot transponowany realizowany jest za pomocą jądra o rozmiarze 7×7 i 9×9 oraz liczbą filtrów 64 i 32 w kolejnych warstwach transponowanej konwolucji. Operacja próbkowania w górę (ang. *up-sampling*) mapy cech celem zwiększenia jej rozdzielczości w drugiej warstwie transponowanej konwolucji realizowana jest z krokiem o rozmiarze 2×2 . Inicjalizacja trenowanych wag przeprowadzana jest za pomocą metody Glorota i wsp. [45].

Segmentacja pikseli realizowana jest za pomocą warstwy konwolucyjnej z funkcją aktywacji typu *softmax* (tzw. warstwy finalnej) oraz warstwy klasyfikacji pikseli. Operacja splotu w ostatniej warstwie konwolucyjnej przeprowadzana jest rozmiarem jądra 1×1 oraz liczbą filtrów odpowiadającą liczbie klas. W opisywanej tu segmentacji semantycznej klasy są dwie: obiekt i tło. Wartości wag są wyznaczone losowo za pomocą metody Glorota i wsp. [45]. Na podstawie map cech realizowana jest klasyfikacja pikseli z uwzględnieniem funkcji straty zdefiniowanej za pomocą uogólnionego indeksu Dice'a *DL* [159]. Ten typ funkcji straty uwzględnia ważenie klas na podstawie liczebności pikseli do nich przypisanych [73]. Wynikiem segmentacji autorską metodą jest obraz binarny $S(x, y)$.

Autorska architektura sieci segmentującej igłę biopsyjną uwzględnia w każdej z warstw konwolucyjnych i transponowanej konwolucji mechanizm normalizacji i funkcję aktywacji ReLU. Zastosowanie tych narzędzi umożliwia skrócenie czasu treningu sieci. Mechanizm normalizacji *BN* będący jedną z technik regularyzacji zastosowany został w celu zmniejszenia wrażliwości treningu na losowe wartości inicjalizacji wag. Mechanizm też również prowadzi do obniżenia ryzyka nadmiernego dopasowania modelu do zbioru treningowego [123, 147]. W proponowanej architekturze zastosowano również warstwy łączenia celem uniknięcia przeuczenia poprzez redukcję nadmiarowych cech.

4.6.2 Zbiory treningowe, walidacyjne i testowe

Do segmentacji igły biopsyjnej wykorzystano tylko te obrazy, które wcześniej zostały wskazane przez sieć klasyfikującą. Na obrazach tych igła została ręcznie obrysowana przy użyciu oprogramowania graficznego [182]. Obrazy wraz z obrysami następnie piętnaście razy podzielono na zbiory treningowy, walidacyjny i testowy z wykorzystaniem *k*-krotnej walidacji krzyżowej na podstawie unikatowych identyfikatorów pacjentów zgodnie z opisem zawartym

w podrozdziale 4.4.2. Podział danych pozwolił na wyznaczenie 15 zestawów składających się z trzech zbiorów każdy, które zostały wykorzystane w procesie treningu i walidacji.

4.6.3 Augmentacja danych

Augmentacja geometryczna danych źródłowych została przeprowadzona w ten sam sposób, jak w przypadku danych obrazowych użytych w sieci klasyfikującej obecność igły biopsyjnej. Opis tych metod przedstawiono w podrozdziale 4.4.3.

4.6.4 Proces treningu sieci segmentującej

Głównym elementem treningu sieci segmentującej jest algorytm minimalizujący funkcję straty poprzez dobór optymalnych wartości wag.

W zaproponowanej metodzie segmentacji zastosowano optymalizator szacowania momentu adaptacyjnego (ang. *adaptive moment estimation*, ADAM). Wybór optymalizatora powodowany był niewielką liczbą warstw ukrytych architektury oraz jego efektywną i skuteczną optymalizacją przy zachowaniu niewielkiego zapotrzebowania na pamięć. Metoda optymalizacji ADAM uwzględnia aktualizację wartości wag w oparciu o każdą próbkę zbioru treningowego wykorzystywaną w procesie uczenia. Optymalizator ten szacuje nowe wartości wag na podstawie wykładniczej średniej ruchomej gradientów [81]. Współczynnik zaniku gradientu B_1 będący jednym z hiperparametrów optymalizatora został ustalony na poziomie 0,9.

Zbiór treningowy podzielono na mniejsze podzbiory (ang. *mini-batch*), co przełożyło się na zmniejszenie poziomu wariancji. Harmonogramowanie współczynnika uczenia celem zmniejszenia tempa w miarę postępu procesu treningu uwzględniało początkową wartość współczynnika uczenia (ang. *initial learn rate*) równą 0,005 ze współczynnikiem spadku szybkości uczenia na poziomie 0,5. Tempo treningu aktualizowano co 3 iteracje.

Mechanizm inicjalizacji wag w poszczególnych warstwach konwolucyjnych autorskiej metody segmentacji igły biopsyjnej realizowany był z użyciem metody Głorota i wsp. [45] w celu eliminacji ryzyka braku zbieżności procesu, wydłużenia czasu treningu lub braku powtarzalności dla modeli o tej samej architekturze i stałych hiperparametrach [57], ale trenowanych innymi zbiorami danych.

Celem eliminacji nadmiernego dopasowania trenowanego modelu do danych w zbiorze treningowym, zastosowano podobne elementy co w sieci klasyfikującej: mechanizmy normalizacji warstwy łączenia, technikę regularyzacji L_2 [12, 108] oraz mechanizm wcześniejszego zatrzymania (ang. *early stopping*) [200]. Wartość parametru L_2 dla wszystkich warstw konwolucyjnych w zaproponowanym rozwiązaniu wyniósł 0,01.



Rys. 4.16: Schemat treningu metody segmentacji z wykorzystaniem sieci CNN. Trening i walidacja modelu realizowana za pomocą obrazów klinicznych obrysowanych przez eksperta.

Schemat procesu treningu autorskiej architektury segmentacji igły biopsyjnej został przedstawiony na rysunku 4.16.

4.6.5 Wyznaczanie trajektorii igły za pomocą transformaty Radona

Wyznaczenie trajektorii igły na obrazie realizowane jest poprzez zastosowanie transformaty Radona [166] na binarnym obrazie pochodzącym z segmentacji $S(x, y)$. Obraz zostaje przekształcony do przestrzeni Radona $\mathcal{R}\{S\}$ zgodnie z definicją:

$$\mathcal{R}\{S\} = \int S(d \cos(\phi) - s \sin(\phi), d \sin(\phi) + s \cos(\phi)) ds \quad (4.17)$$

gdzie d odpowiada najkrótszemu odcinkowi łączącemu początek układu współrzędnych z analizowanym punktem, natomiast ϕ jest kątem pomiędzy tym odcinkiem a osią OX ; s stanowi wyraz wolny korespondujący ze specyficznym punktem kształtu [119].

Powstała w ten sposób przestrzeń Radona $\mathcal{R}\{S\}$ jest sinogramem, w którym widoczne jasne obszary odpowiadają obiektom w segmentowanym obrazie. Im jaśniejszy obszar, tym większa kolinearność obiektów widocznych na analizowanym obrazie. Zakłada się, że położenie i orientacja odpowiadające maksymalnej wartości sinogramu reprezentują prostą przechodzącą przez igłę, dlatego z obrazu usuwane są inne segmentowane obszary [29, 55]. Ostateczny obraz zawierający informacje o lokalizacji igły jest oznaczony jako $R(x, y)$.

4.7 Komitety klasyfikujące do segmentacji zmiany ogniskowej

Obrazowanie ultrasonograficzne pozwala w nieinwazyjny sposób zwizualizować zmianę ogniskową w badanym narządzie i oszacować jej parametry. Obrazowanie zmienionej chorobowo tkanki pozwala na przeprowadzenie zabiegu biopsji gruczołowej. Ograniczeniem biopsji przeprowadzanej pod kontrolą obrazowania USG jest fakt, że jest ono tylko prezentacją dwuwymiarową, a ope-

rator musi w trakcie zabiegu równocześnie manipulować zarówno głowicą ultrasonograficzną, jak i pistoletem biopsyjnym. Dlatego też istotnym wydaje się zaproponowanie systemu wspomagania biopsji gruboigłowej wskazującego region zainteresowania odpowiadający zmianie ogniskowej w celu uniknięcia pobrania tkanek niezmiennych chorobowo.

W kolejnych podrozdziałach przedstawiono propozycję metody segmentacji zmiany ogniskowej opartej o metody klasyczne, a następnie agregowanej informacji za pomocą komitetu klasyfikującego.

4.7.1 Filtracja dwuwymiarowych obrazów ultrasonograficznych

Segmentacja zmian ogniskowych w obrazach ultrasonograficznych poprzedzona jest filtracją w celu usunięcia charakterystycznego dla tej modalności multiplikatywnego szumu plamkowego [36, 99], który prowadzi do rozmycia krawędzi [59] utrudniając wskazanie granicy patologicznej tkanki. Filtracja realizowana jest za pomocą wygładzenia gaussowskiego [102, 190], dyfuzji anizotropowej [43, 102], filtracji medianowej [90, 180] czy metody średnich nielokalnych (ang. *non-local means*, NLM) [183]. W niniejszej pracy zweryfikowano również filtrację uśredniającą i bilateralną [120].

Filtracja uśredniająca

Filtracja uśredniająca (ang. *mean filter*) jest przykładem filtracji dolno-przepustowej, stosowanej w przetwarzaniu obrazów. Zastosowanie filtra uśredniającego umożliwiłoby łagodne wygładzenie obrazu. Celem filtracji uśredniającej jest, aby intensywność analizowanego piksela przybrała wartość średnią intensywności sąsiadujących pikseli w odległości wskazanej oknem filtra.

$$J(i, j) = \frac{1}{W} \sum_{i, j \in K} I(x + i, y + j) \cdot w(i, j) \quad (4.18)$$

gdzie: K stanowi okno (maskę) filtra, $w(i, j)$ jest wagą przypisaną elementowi okna, natomiast W jest współczynnikiem normalizacji sumy wag okna filtra. W przypadku filtra uśredniającego wszystkie elementy okna mają przypisane wagi w , natomiast współczynnik normalizacji przyjmuje wartość odpowiadającą liczbie elementów okna [97]. Przykładowy rezultat filtracji przedstawiono na rysunku 4.17a.

Filtracja medianowa

Nieliniowy filtr medianowy (ang. *median filter*) stosowany jest w celu usunięcia szumów z obrazu przy jednoczesnym zachowaniu kluczowych informacji. Filtracja medianowa pozwala na usunięcie większości lokalnych zakłóceń typu

„sól i pieprz” bez znaczącego rozmycia obrazu, co może powodować utratę drobnych detali. Metoda ta nie wpływa na rozmycie krawędzi. Zasada działania filtracji medianowej oparta jest o wyznaczenie dla współrzędnych analizowanego piksela wyjściowej wartości intensywności, która odpowiada medianie wartości intensywności w sąsiedztwie wskazywanych oknem filtra [97]. Przykładowy efekt filtracji przedstawiono na rysunku 4.17b

Filtracja Gaussa

Filtr Gaussa (ang. *Gauss filter*) również jest przykładem liniowej filtracji dolnoprzepustowej z wykorzystaniem okna operującego średnią ważoną [97]. Wagi w filtrze określone są zgodnie z rozkładem Gaussa, który definiowany jest następująco:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (4.19)$$

gdzie σ^2 stanowi wariancję. Filtr ten nakładany jest zgodnie ze wzorem 4.18, różniąc się co do wartości wag w . Efektem zastosowania filtracji gaussowskiej jest mocniejsze niż w przypadku filtracji uśredniającej wygładzenie obrazu poprzez jego rozmycie ale z mniejszym oddziaływaniem na krawędzie obiektu. Przykład filtracji gaussowskiej przedstawiono na rysunku 4.17c

Dyfuzja anizotropowa

Dyfuzja anizotropowa (ang. *anisotropic diffusion*, AF) określana również mianem dyfuzji Perony-Malika jest metodą iteracyjną stosowaną w celu minimalizacji zaszumienia obrazu bez usuwania istotnych jego treści, takich jak krawędzie, linie i inne kluczowe detale [126]. Przykładowy rezultat filtracji dyfuzją anizotropową pokazano na rysunku 4.17d. Metoda ta definiowana jest następująco:

$$I^{t+1}(x) = I^t(x) + \lambda \sum_{y \in N(x)} c^t(x, y) \cdot (I^t - I^t(y)) \quad (4.20)$$

gdzie I stanowi intensywność obrazu, t odpowiada liczbie iteracji, x jest współrzędną analizowanego piksela, λ określa szybkość dyfuzji, natomiast $N(x)$ sąsiedztwo piksela x . Współczynnik dyfuzji $c^t(x, y) = g(|\nabla I|)$, który składa się z dwóch składowych: c_1 i c_2 , które odpowiednio nadają priorytet krawędom i intensywnościom, zgodnie z definicjami:

$$c_1(x, y) = \exp\left(-\frac{|\nabla I|^2}{2k^2}\right) \quad (4.21)$$

$$c_2(x, y) = \frac{1}{1 + \frac{|\nabla I|^2}{2k^2}} \quad (4.22)$$

gdzie k jest parametrem wielkości krawędzi i ∇ stanowi lokalny gradient.

Filtracja metodą średnich nielokalnych

Filtracja metodą średnich nielokalnych (ang. *non-local means*, *NLM*) [16] pozwala na eliminację szumu z obrazu przy zachowaniu ostrości krawędzi obiektów obecnych w obrazie. Przykładowy efekt filtracji przedstawiono na rysunku 4.17e. Filtr średnich nielokalnych wyznacza wartość analizowanego piksela $NL(i)$ na podstawie wartości średniej ważonej wszystkich pikseli w zaszumionym obrazie $v = \{v(i)|i \in I\}$ zgodnie ze wzorem:

$$NL(i) = \sum_{j \in I} w(i, j)v(j), \quad (4.23)$$

gdzie rodzina wag $\{w(i, j)\}_j$ jest miarą opartą o podobieństwo pomiędzy pikselami i i j , a także spełnia zależności $0 \leq w(i, j) \leq 1$ oraz $\sum_j w(i, j) = 1$. Parametrami filtracji są stopień wygładzenia, rozmiar okna wyszukiwania oraz rozmiar okna podobieństwa. Filtracja *NLM* pozwala na wytłumienie szumów bez rozmycia krawędzi igły.

Filtracja bilateralna

Filtracja bilateralna (ang. *bilateral filter*, *BF*) stanowi przykład nieliniowej metody wygładzania obrazu przy zachowaniu krawędzi struktur. Intensywność piksela $BF(x)$ definiowana jest na podstawie sumy ważonej intensywności pikseli sąsiadujących z uwzględnieniem w wartości wag kryterium odległości pomiędzy pikselami [167]. Filtr bilateralny przedstawiony jest wzorem:

$$BF(x) = \frac{1}{W_p} \sum_{x_i \in \Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|) \quad (4.24)$$

gdzie: Ω jest oknem o środku we współrzędnej x , f_r stanowi jądro zakresu definiowanego celem wygładzenia różnic w intensywnościach, natomiast g_s jest jądrem przestrzennym definiowanym do wygładzania różnic we współrzędnych [9]. W_p jest współczynnikiem normalizacji sumy wag okna filtra definiowanym następująco:

$$W_p = \sum_{x_i \in \Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|) \quad (4.25)$$

Metoda ta jest metodą lokalną i nieiteracyjną, pozwalającą na łączenie

poziomów szarości w oparciu o kryterium odległości geometrycznej oraz podobieństwo fotometryczne [9, 167]. Przykładowy rezultat filtracji przedstawiono na rysunku 4.17f

4.7.2 Metody oceny jakości filtracji obrazów

Miary podobieństwa stanowią istotny element oceny filtracji obrazów, ale również w ich klasyfikacji, grupowaniu, detekcji (segmentacji) oraz rejestracji [124]. Szereg narzędzi stosowanych w ewaluacji podobieństwa obrazów można podzielić na następujące kategorie: pomiary zmiany (błędu) pikseli, ocena poziomu szumów, wskaźniki zachowania krawędzi struktur widocznych w obrazach czy parametry jednorodności strukturalnej obrazu [100, 145]. Szczególnie w przypadku oceny filtracji, zastosowanie metryk wydaje się być kluczowe do wytypowania rozwiązania pozwalającego na zachowanie elementów istotnych dla potoku proponowanego rozwiązania.

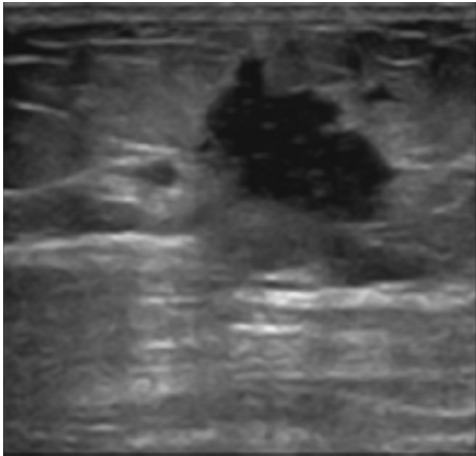
Ocena porównawcza pikseli obrazu oryginalnego z obrazem poddanym filtracji pozwalająca na oszacowanie redukcji szumu, realizowana jest z zastosowaniem podstawowych miar oceny jakości filtracji, których definicje przytoczono poniżej. Pierwszą ze stosowanych metryk jest błąd średniokwadratowy (ang. *mean squared error*, *MSE*) [124, 144], który jest opisany następującym wzorem:

$$MSE(I, J) = \frac{1}{|D|} \sum_{x_i \in D} (I(x_i) - J(x_i))^2 \quad (4.26)$$

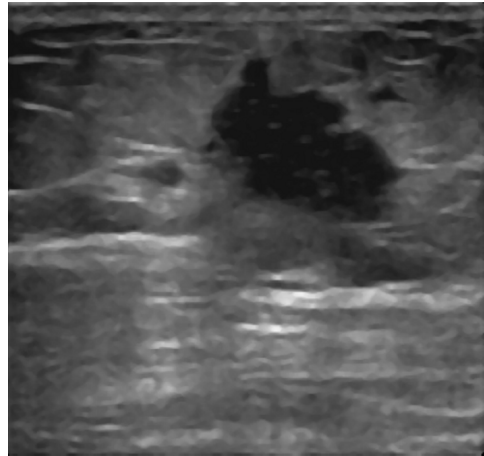
gdzie I i J stanowią odpowiednio oznaczenia obrazu odniesienia (oryginalnego) oraz obrazu wynikowego (poddanego filtracji); D jest domeną obrazu (zbiorem wszystkich pikseli x_i), $|D|$ jest mocą domeny obrazu (liczbą pikseli w obrazie). Stosowaną modyfikacją miary MSE jest znormalizowany średni błąd kwadratowy (ang. *normalized mean square error*, *NMSE*), którego definicja jest następująca:

$$NMSE(I, J) = \frac{\sum_{x_i \in D} (I(x_i) - J(x_i))^2}{\sum_{x_i \in D} I(x_i) \cdot \sum_{x_i \in D} J(x_i)} \quad (4.27)$$

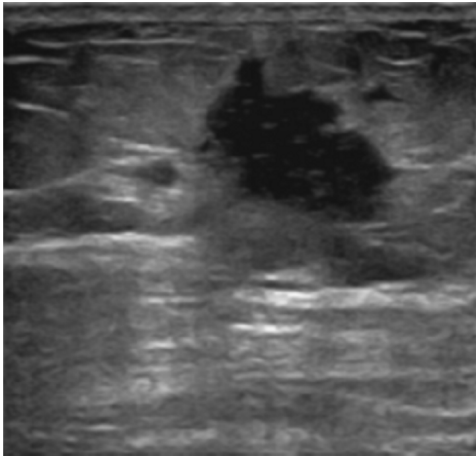
Kolejną stosowaną metryką oceny jakości filtracji jest szczytowy stosunek sygnału do szumu (ang. *peak signal-to-noise ratio*, *PSNR*) [144], wyrażany na-



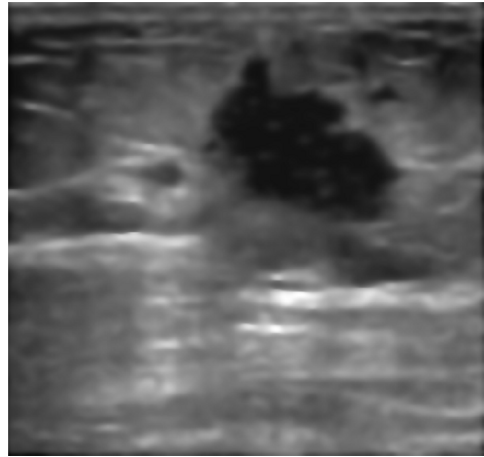
(a) Filtracja uśredniająca



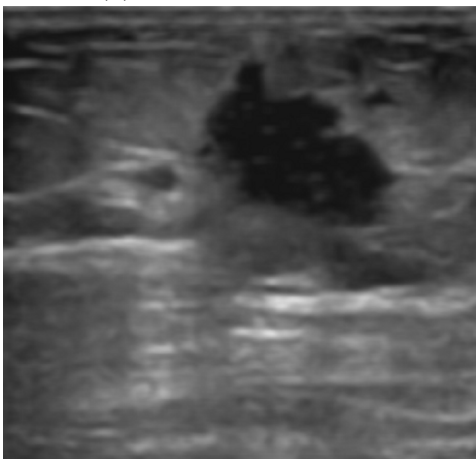
(b) Filtracja medianowa



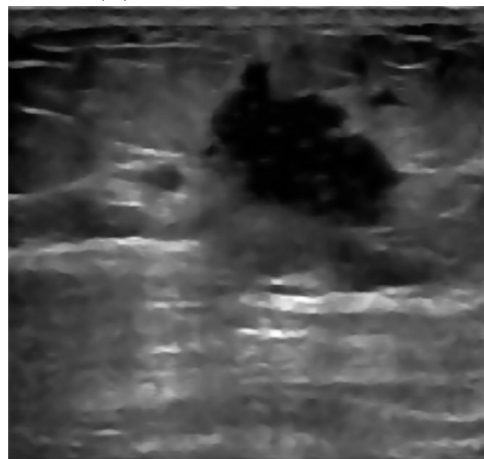
(c) Filtracja gaussowska



(d) Dyfuzja anizotropowa



(e) Metoda średnich nielokalnych



(f) Filtr bilateralny

Rys. 4.17: Przykłady filtracji obrazów ultrasonograficznych zmiany ogniskowej gruczołu sutkowego.

stępująco:

$$PSNR(I, J) = 10 \cdot \log \left(\frac{(L-1)^2}{MSE(I, J)} \right) \text{ dB} \quad (4.28)$$

gdzie L jest dynamicznym zakresem wartości pikseli (np. dla obrazów monochromatycznych o reprezentacji 8-bitowej wynosi 256) [75]. Prostota i niska złożoność matematyczna umożliwia szybką i wygodną ocenę straty jakości obrazu, niemniej jednak miary te nie są wystarczające do oceny jakości wizualnej obrazu.

Wśród często stosowanych miar można wskazać wraz z definicjami również średni błąd bezwzględny (ang. *mean absolute error*, *MAE*):

$$MAE(I, J) = \frac{1}{|D|} \sum_{x_i \in D} |I(x_i) - J(x_i)|, \quad (4.29)$$

jak również stosunek sygnału do szumu (ang. *signal to noise ratio*, *SNR*):

$$SNR(I, J) = \frac{\sum_{x_i \in D} I(x_i)^2}{\sum_{x_i \in D} (I(x_i) - J(x_i))^2}, \quad (4.30)$$

oraz pierwiastek kwadratowy z *MSE* (ang. *root mean square error*, *RMSE*) definiowany jako standardowe odchylenie wartości intensywności pikseli [125, 144]:

$$RMSE = \sqrt{\frac{1}{|D|} \sum_{x_i \in D} (I(x_i) - J(x_i))^2}. \quad (4.31)$$

Ze względu na silne przestrzenne zależności pikseli obrazu jako alternatywną metodę oceny jakości struktury stosowany jest indeks jakości podobieństwa strukturalnego (ang. *structural similarity quality index*, *SSIM*) [144, 175], którego definicja jest następująca:

$$SSIM(I, J) = \frac{(2\mu_I\mu_J + C_1)(2\sigma_{IJ} + C_2)}{(\mu_I^2 + \mu_J^2 + C_1)(\sigma_I^2 + \sigma_J^2 + C_2)} \quad (4.32)$$

gdzie μ_I i μ_J są odpowiednio średnimi porównywanych obrazów I i J :

$$\mu_I = \frac{1}{|D|} \sum_{x_i \in D} I(x_i), \quad (4.33)$$

$$\mu_J = \frac{1}{|D|} \sum_{x_i \in D} J(x_i) \quad (4.34)$$

natomiast σ_I i σ_J stanowią wartości wariancji, definiowane następująco:

$$\sigma_I^2 = \frac{1}{|D|-1} \sum_{x_i \in D} (I(x_i) - \mu_I)^2, \quad (4.35)$$

$$\sigma_J^2 = \frac{1}{|D|-1} \sum_{x_i \in D} (J(x_i) - \mu_J)^2 \quad (4.36)$$

Kowariancja σ_{IJ} wyznaczana jest następująco:

$$\sigma_{IJ} = \sum_{x_i \in D} (I(x_i) - \mu_I)(J(x_i) - \mu_J) \quad (4.37)$$

gdzie D przyjmuje wartość równą mocy zbioru obrazu. C_1 i C_2 są współczynnikami stabilizującymi równanie, wyznaczanymi za pomocą wzorów:

$$C_1 = ((L - 1) \cdot k_1)^2 \quad (4.38)$$

$$C_2 = ((L - 1) \cdot k_2)^2 \quad (4.39)$$

gdzie L jest dynamicznym zakresem wartości pikseli, który zazwyczaj przyjmuje wartości $2^{\#bits \text{ per pixel}} - 1$, natomiast współczynniki $k_1, k_2 \ll 1$.

Kolejną miarą oceny jakości struktury jest wskaźnik jakości Q (ang. *quality index* Q , Q) [174] wyrażony wzorem:

$$Q = \frac{4\sigma_{IJ}\mu_I\mu_J}{(\sigma_I^2 + \sigma_J^2)(\mu_I^2 + \mu_J^2)} \quad (4.40)$$

Zakres dynamiczny wskaźnika Q wynosi $[-1; 1]$, a najlepsza wartość to 1.

Metryki wymienione powyżej nie umożliwiają oceny jakości zachowania krawędzi, co można przeprowadzić poprzez pomiar tłumienia szumu i wyrazić za pomocą wskaźnika ρ (ang. *noise suppression* ρ , ρ) [145]:

$$\rho = \frac{\Gamma(I - \bar{I}, J - \bar{J})}{\sqrt{\Gamma(I - \bar{I}, I - \bar{I}) \cdot \Gamma(J - \bar{J}, J - \bar{J})}} \quad (4.41)$$

gdzie \bar{I} i \bar{J} są wartościami średnimi odpowiednio w I i J oraz:

$$\Gamma(s_1, s_2) = \sum_{x_i \in D} s_1(x_i) \cdot s_2(x_i) \quad (4.42)$$

Druga z metryk, to miara zachowania krawędzi, która jest definiowana następującym wzorem:

$$\beta = \frac{\Gamma(\Delta I - \bar{\Delta I}, \Delta J - \bar{\Delta J})}{\sqrt{\Gamma(\Delta I - \bar{\Delta I}, \Delta I - \bar{\Delta I}) \cdot \Gamma(\Delta J - \bar{\Delta J}, \Delta J - \bar{\Delta J})}} \quad (4.43)$$

gdzie ΔI i ΔJ stanowią wynik filtracji górnoprzepustowej odpowiednio I i J za pomocą standardowego przybliżenia operatora Laplace'a o wymiarach 3×3 ; $\bar{\Delta I}$ i $\bar{\Delta J}$ są średnimi wartościami I i J [100, 145]

4.7.3 Dobór metody filtracji i jej parametrów

Każda z metod filtracji cechuje się parametrami, które wpływają na jakość przetwarzania obrazu. Celem doboru optymalnego narzędzia oraz zestawu parametrów poddano je testowaniu w zakresach przedstawionych w tabeli 4.2. Ewaluacja metod filtracji zrealizowana została z wykorzystaniem zbioru danych składającego się z 350 dwuwymiarowych obrazów ultrasonograficznych zebranych podczas 20 badań klinicznych zmian ogniskowych obecnych w gruczołach sutkowych. Obrazy te zapisano za pomocą systemu rejestracji danych, którego opis jest zawarty w rozdziale 3. Obrazy kliniczne pochodziły z ultrasonografu Hitachi EZU-MT25-s1 wyposażonego w głowicę liniową Hitachi EUP-L65. Głębokość wiązki ultradźwięków była stała dla każdego pacjenta, a zakres wynosił od 25 do 50 mm w zależności od lokalizacji zmiany. Wartość częstotliwości była związana z głębokością wiązki ultradźwiękowej i wynosiła od 6 do 14 MHz.

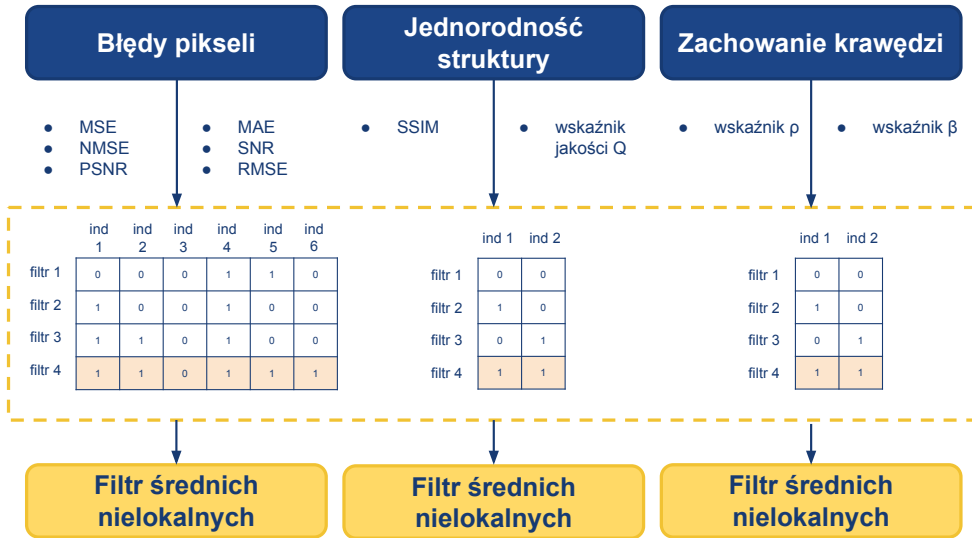
W oparciu o metody oceny jakości filtracji (Podrozdz. 4.7.2), został przeprowadzony eksperyment, którego celem było wskazanie metody filtracji oraz zestawu jej parametrów, które będą najkorzystniejsze w kontekście dostępnych danych obrazowych. Wartości otrzymanych parametrów zostały poddane ocenie binarnej w następujący sposób:

$$f(x) = \begin{cases} 0 & \text{gdy } g(x) \neq \max(\text{metryka}) \\ 1 & \text{gdy } g(x) = \max(\text{metryka}) \end{cases}$$

Następnie każdy z testowanych filtrów został poddany ewaluacji i zestawiony w rankingu (Rys. 4.18). Finalnie wybrany został filtr, który najczęściej otrzymywał wartość 1 dla poszczególnych metryk. W każdej z analizowanych kategorii był to filtr średnich nielokalnych. Różnice występowały jedynie w parametrach wejściowych i zostały zweryfikowane na etapie segmentacji.

Tab. 4.2: Testowane parametry różnych metod filtracji (pierwsza wartość to początek testowanego zakresu dla danego parametru, ostatnia koniec, natomiast środkowa to krok).

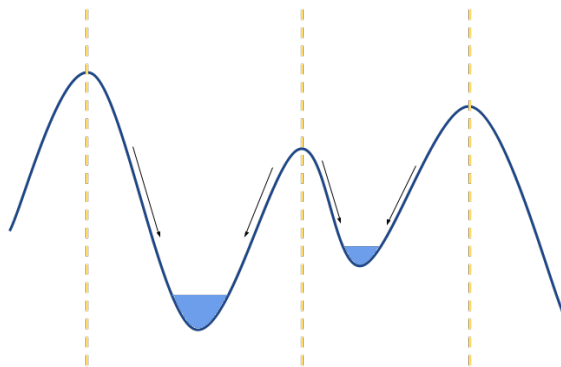
Typ filtracji	Parametry
filtr uśredniający	$n = 3 : 2 : 9$
filtr medianowy	$n = 3 : 2 : 9$
filtr gaussowski	$\sigma = 0.25 : 0.25 : 3$
dyfuzja anizotropowa	$iter = 5 : 5 : 25$
	$\kappa = 5 : 5 : 60$
filtr średnich nielokalnych	$t = 1 : 1 : 6$
	$f = 1 : 1 : 10$
	$h = 0.25 : 0.25 : 3$
filtr bilateralny	$w = 0.25 : 0.25 : 3$
	$\sigma = 0.25 : 0.25 : 3$



Rys. 4.18: Ocena metod filtracji obrazów ultrasonograficznych zmiany ogniskowej gruczołu sutkowego.

4.7.4 Segmentacja zmiany ogniskowej metodami klasycznymi

Przefiltrowany dwuwymiarowy obraz ultrasonograficzny stanowi obraz wejściowy w przetwarzaniu klasycznymi metodami segmentacji. W ramach niniej-



Rys. 4.19: Algorytm wododziałowy bazuje na detekcji wododziałów, które są zaznaczone kolorem żółtym i odpowiadają grzbietom. Wododział rozgranicza rozlewiska zaznaczone kolorem niebieskim.

szej rozprawy przetestowano cztery metody segmentacji, które można podzielić na dwie kategorie: (i) metoda automatyczna będąca algorytmem wododziałowym (ang. *watershed segmentation*, WS), oraz (ii) metody półautomatyczne, do których należą rozrost obszaru (ang. *region growing*, RG), aktywny kontur (ang. *active contour*, AC) oraz aktywny kontur wykorzystujący obraz gradientowy (ang. *gradient vector flow*, GVF).

Metoda automatyczna pozwala na segmentację maski zmiany ogniskowej bazując jedynie na wczytanym obrazie. Natomiast w przypadku metod półautomatycznych konieczne jest podanie przez użytkownika przynajmniej jednego punktu startowego, natomiast w przypadku aktywnego konturu i jego odmian wymagane są przynajmniej cztery punkty, ponieważ jest to minimalna liczba punktów, która pozwala na wyznaczenie interpolacją zamkniętej krzywej (w przypadku trzech punktów obiekt wyznaczony za pomocą krzywej pozostaje otwarty).

Poniżej opisano kolejne przetestowane metody segmentacji.

Segmentacja wododziałowa

Segmentacja wododziałowa (ang. *watershed segmentation*) jest rodzajem segmentacji krawędziowej bazującej na gradiencie obrazu. Inspiracją algorytmu był proces zalewania obszarów o różnej wysokości. Najwyższe obiekty będące maksimami lokalnymi wyznaczają grzbiety, które rozgraniczają obiekty. W obrazowaniu odpowiednik grzbietu stanowi amplituda gradientu [67, 180].

Zaletą tej metody jest automatyzacja oraz krótki czas działania, co umożliwia jej aplikację w systemach wymagających niemal natychmiastowego wyniku.

Rozrost obszaru

Metoda rozrostu obszaru (ang. *region growing*) stanowi metodę segmentacji obrazów bazującą na analizie pikseli sąsiadujących z punktami początkowymi wskazanymi przez użytkownika i decyzji, czy dany piksel powinien zostać dodany do tworzonej maski binarnej [90]. Metoda ta bazuje na technice progowania rozszerzonej o analizę położenia piksela (uwzględnia sąsiedztwo). Zaletą tej metody jest możliwość poprawnego działania już przy jednym punkcie startowym.

Aktywny kontur

Parametryczny aktywny kontur (ang. *active contour*) jest metodą lokalnej segmentacji krawędziowej. Kontur maski binarnej jest wyznaczany iteracyjnie na podstawie przyłożonych do aktualnego stanu sił wewnętrznych, których celem jest wygładzenie kształtu konturu, i zewnętrznych służących dostosowaniu konturu do środowiska. Kształt konturu ulega zmianie poprzez dążenie do minimalizacji energii, której źródłem są siły [78]. Proces modyfikacji węzłów jest powtarzany iteracyjnie aż do spełnienia warunku zakończenia. Dla każdego punktu $v(s) = (x(s), y(s))$ energia E_{snake} definiowana jest następująco:

$$E_{snake} = \int_0^1 E_{int}(v(s))ds + \int_0^1 E_{image}(v(s))ds + \int_0^1 E_{con}(v(s))ds, \quad (4.44)$$

gdzie: E_{int} jest energią wewnętrzną kontrolującą gładkość i sztywność konturu (węża), E_{image} stanowi energię obrazu odpowiadającą za położenie konturu w obrazie, natomiast E_{con} jest energią zewnętrznych sił ograniczających (definiowanych przez użytkownika).

Wartość energii konturu wewnętrznego obliczana jest zgodnie z zależnością:

$$E_{int} = \frac{1}{2} \left(\alpha(s) |v_s(s)|^2 \right) + \frac{1}{2} \left(\beta(s) |v_{ss}(s)|^2 \right), \quad (4.45)$$

gdzie α jest parametrem wrażliwości na rozciąganie, a β odpowiada wrażliwości na krzywiznę konturu. Energia obrazu E_{image} jest wyznaczana za pomocą wzoru:

$$E_{image} = w_{line} E_{line} + w_{edge} E_{edge} + w_{term} E_{term}, \quad (4.46)$$

gdzie w_{line} , w_{edge} i w_{term} są wagami, $E_{line} = I(x, y)$, $E_{edge} = -|\nabla I(x, y)|^2$, natomiast E_{term} zależy od kierunku gradientu.

Efektem przetwarzania jest maska binarna odwzorowująca obiekt o dowolnym wypukłym kształcie, co stanowi zaletę w przypadku segmentacji zmian

ogniskowych w obrazach ultrasonograficznych, które cechują się różnymi kształtami [36, 99, 120].

Aktywny kontur uwzględniający rozszerzenie zasięgu gradientu

Modyfikacja metody aktywnego konturu poprzez wykorzystanie gradientu obrazu (ang. *gradient vector flow*, GVF) zamiast jego energii rozszerza możliwość metody o segmentację struktur cechujących się wklęsłym kształtem [191]. GVF stanowi definiowane na podstawie gradientu obrazu $|\nabla I|$ pole wektorowe $V(x, y) = (u(x, y), v(x, y))$, które oddziałuje na oryginalne pole wektorowe $|\nabla f|$ stosowane w klasycznym aktywnym konturze i minimalizuje następująco energię:

$$E_{\text{GVF}} = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V - \nabla f|^2 dx dy \quad (4.47)$$

gdzie współczynnik μ kontroluje zgodność pól wektorowych [192]. Podejście to pozwala na segmentację zmian ogniskowych cechujących się obecnością wklęsłości [137, 183].

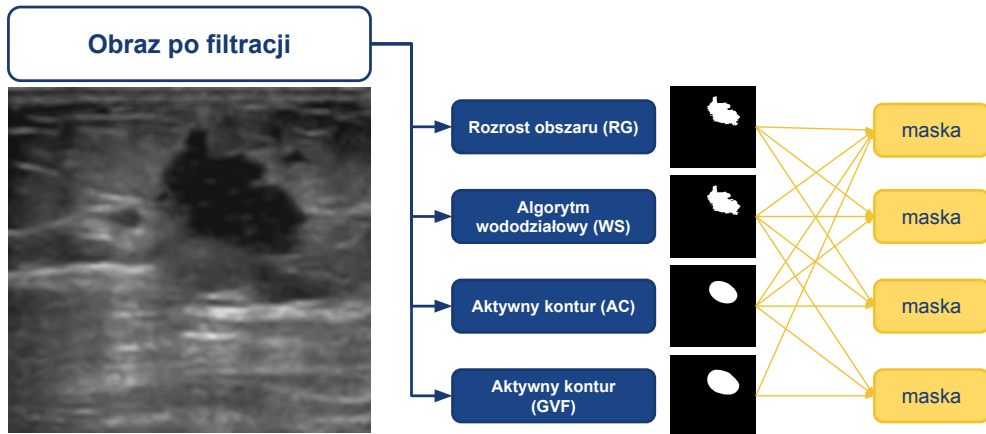
4.7.5 Komitety klasyfikujące

Segmentacja zmiany ogniskowej gruczołu sutkowego realizowana metodami klasycznymi jest obciążona błędem związanym z doбором parametrów, które nie mogą być wskazane uniwersalnie ze względu na heterogeniczny charakter obrazów ultrasonograficznych. W niniejszej rozprawie proponowany jest komitet klasyfikujący (ang. *ensemble*), który pozwala na agregację informacji pochodzącej z kilku źródeł. Maskę wynikową jest wyznaczona w oparciu o połączenie masek otrzymanych za pomocą poszczególnych metod segmentacji: metody rozrostu, aktywnego konturu, aktywnego konturu uwzględniającego rozszerzenie gradientu oraz segmentacji wododziałowej. Kryterium przyłączenia piksela do maski wynikowej jest następujące:

$$J(x, y) = \begin{cases} 0 & \text{gdy } \sum_{m=0}^n I_m(x_i) < \lceil \frac{n}{2} \rceil \\ 1 & \text{gdy } \sum_{m=0}^n I_m(x_i) \geq \lceil \frac{n}{2} \rceil \end{cases} \quad (4.48)$$

gdzie n stanowi liczbę wykorzystanych metod segmentacji, I_m jest maską otrzymaną z każdej z poszczególnych metod segmentacji, a J stanowi hybrydową maskę wynikową.

Przetestowano wszystkie kombinacje na danych obrazowych i najkorzystniejsze wyniki odnotowano w przypadku następujących, testowanych połączeń:

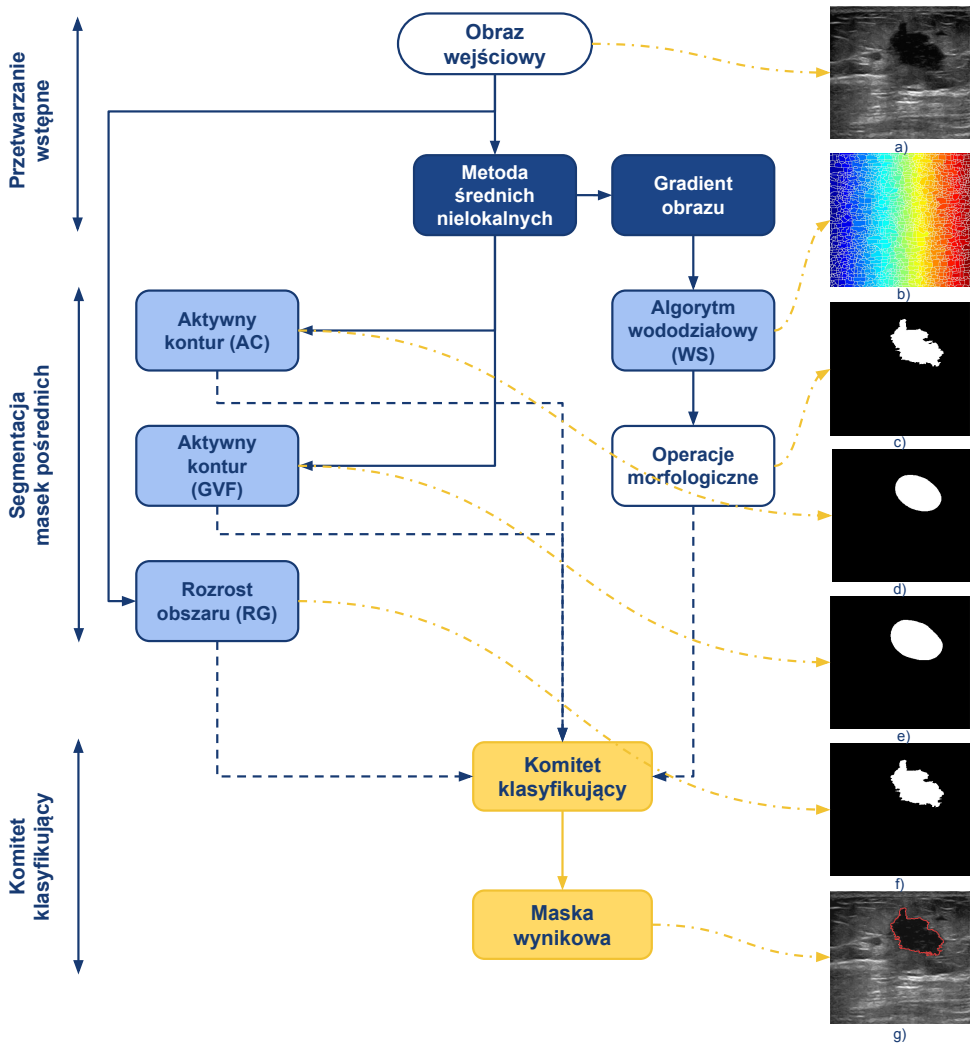


Rys. 4.20: Komitet klasyfikatorów klasycznych metod segmentacji zmiany ogniskowej w obrazach ultrasonograficznych.

- aktywny kontur bazujący na energii obrazu z aktywnym konturem wykorzystującym informację gradientową,
- aktywny kontur bazujący na energii obrazu z aktywnym konturem wykorzystującym informację gradientową oraz rozrostem obszaru,
- aktywny kontur bazujący na energii obrazu z rozrostem obszaru,
- aktywny kontur wykorzystujący informację gradientową oraz rozrost obszaru,
- połączenie wszystkich metod.

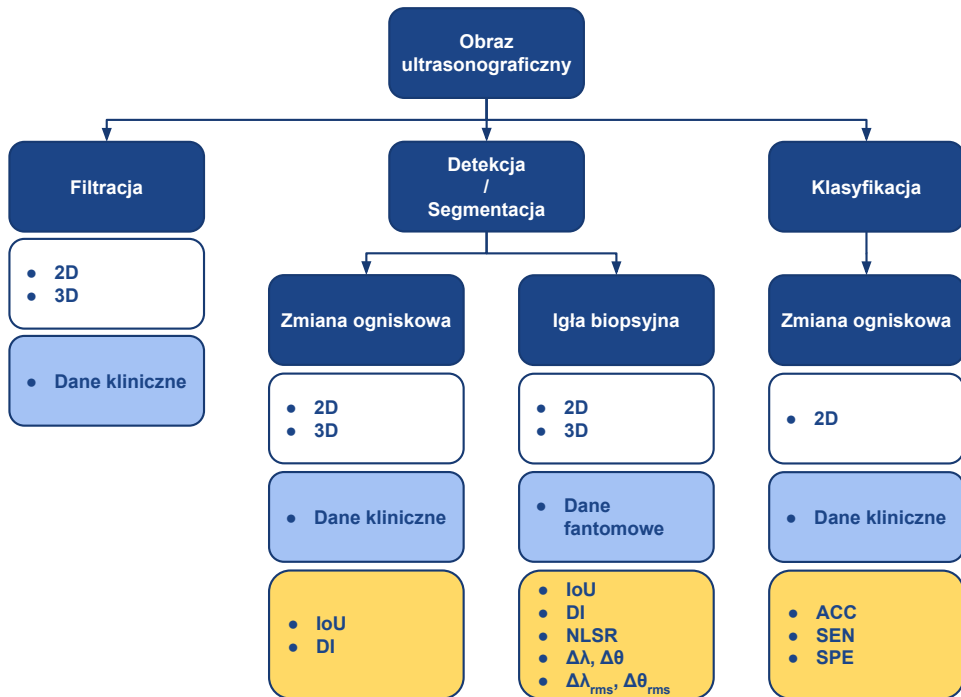
4.8 Ewaluacja metod klasyfikacji i segmentacji obrazów USG

Każda z metod zarówno klasyfikacji jak i segmentacji, niezależnie od sposobu przygotowania danych czy złożoności proponowanych rozwiązań musi zostać poddana ewaluacji za pomocą narzędzi zapewniających niezależną i obiektywną ocenę pozwalającą na jednoczesne porównanie różnych metod dedykowanych temu samemu zagadnieniu. Wyróżnia się ocenę jakościową oraz ilościową. Pierwszy z typów uzależniony jest od doświadczenia użytkownika, drugi wymaga definicji złotego standardu [156]. W poniższym podrozdziale dokonano przeglądu i analizy wskaźników liczbowych pozwalających dokonać oceny ilościowej autorskiej metody klasyfikującej wykorzystanej w dwóch problemach: (i) klasyfikacji obecności igły w obrazie USG oraz (ii) klasyfikacji typu zmiany



Rys. 4.21: Komitet klasyfikatorów klasycznych metod segmentacji zmiany ogniskowej w obrazach ultrasonograficznych.

ogniskowej lub jej braku w oparciu o przegląd literatury (Podrozdz. 2.2 i 2.3). Przedstawione są również narzędzia dedykowane ocenie algorytmów segmentacji w kontekście oceny efektywności segmentacji struktur takich jak zmiana ogniskowa lub igła biopsyjna w obrazach ultrasonograficznych, jak również wskazano stosowane metryki ewaluacji szacowania dokładności trajektorii igły biopsyjnej (Rys. 4.22).



Rys. 4.22: Podsumowanie metod ewaluacji z zakresu przetwarzania obrazów ultrasonograficznych.

4.8.1 Miary oceny klasyfikacji

Efektywność klasyfikatora mierzona jest różnymi wskaźnikami w odniesieniu do etykiet, które zostały wskazane dla poszczególnych próbek w zbiorze. Do najpopularniejszych należą dokładność, czułość oraz swoistość. Miary te wykorzystują pojęcia wykryć próbek prawdziwie dodatnich (ang. *true positive*, *TP*), prawdziwie ujemnych (ang. *true negative*, *TN*), fałszywie dodatnich (ang. *false positive*, *FP*) oraz fałszywie ujemnych (ang. *false negative*, *FN*) [121, 130]. Wykrycie próbki jest zaliczane do jednej z czterech powyższych grup na podstawie rzeczywistej przynależności (bądź jej braku) próbki do klasy (Rys. 4.23, kolumny) oraz wyniku działania czyli predykcji klasyfikatora (ibidem, wiersze) [108].

Poza wynikami pozytywnymi i negatywnymi w literaturze wyróżniane są wyniki niejednoznaczne, które dzielone są na nieważne i ważne. Pierwsze z nich to takie, gdzie ze względu na brak spełnienia minimalnych kryteriów próbki nie ma możliwości ich interpretacji. Drugi przypadek dotyczy próbek dla których otrzymany wynik nie jest wyraźnie dodatni lub ujemny (prawdopodobieństwo przynależności do analizowanych klas jest porównywalne). Wyniki niejedno-

Klasa		Rzeczywista	
		dodatnia	ujemna
Predykowana	dodatnia	<p>prawdziwie dodatnia TP</p>	<p>fałszywie dodatnia FP</p>
	ujemna	<p>fałszywie ujemna FN</p>	<p>prawdziwie ujemna TN</p>

Rys. 4.23: Tablica pomyłek.

znaczące są często całkowicie pomijane przy ocenie narzędzia, pomijanie przy ocenie ale raportowane niezależnymi miarami lub traktowane jako wyniki negatywne (przeciwnie do etykiety klasy do której należy próbka) [149]. Problem niejednoznacznych wyników dotyczy głównie testów laboratoryjnych (analiza próbek krwi, tkanek itp.), w przypadku prac z zakresu obrazowania nie jest uwzględniany.

Poniżej przedstawiono definicję wszystkich miar klasyfikacji wykorzystanych w niniejszej rozprawie.

Dokładność klasyfikatora (ang. *accuracy*, *ACC* lub ang. *correct rate*) definiowana jest jako iloraz liczby poprawnie sklasyfikowanych próbek i liczby wszystkich sklasyfikowanych próbek [121, 130]. Dokładność wyrażana jest następującym wzorem:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.49)$$

Dopełnieniem wskaźnika dokładności klasyfikatora jest poziom błędów klasyfikatora (ang. *error rate*, *ER*) [108]. Wskaźnik ten pozwala oszacować liczbę nieprawidłowo sklasyfikowanych próbek w odniesieniu do wszystkich sklasyfikowanych próbek zgodnie ze wzorem:

$$ER = \frac{FP + FN}{TP + TN + FP + FN} \quad (4.50)$$

Czułość klasyfikatora (ang. *sensitivity*, *SEN*) określana również mianem współczynnika wyników prawdziwie pozytywnych (ang. *true positive rate*, *TPR*) stanowi iloraz liczby sklasyfikowanych prawdziwie dodatnie próbek oraz sumy prawdziwie dodatnich i fałszywie ujemnych próbek [121, 130]. Wzór definiujący

miarę TPR jest następujący:

$$SEN = TPR = \frac{TP}{TP + FN} \quad (4.51)$$

Swoistość klasyfikatora (również zwana specyficznością, z ang. *specificity*, SPE), także określana mianem współczynnika wyników prawdziwie negatywnych (ang. *true-negative rate*, TNR), szacowana jest na podstawie stosunku liczby prawidłowo sklasyfikowanych próbek ujemnych do sumy liczby próbek prawdziwie ujemnych i fałszywie dodatnich, zgodnie z zapisem [121, 130]:

$$SPE = TNR = \frac{TN}{TN + FP} \quad (4.52)$$

Współczynnik wyników fałszywie pozytywnych (ang. *false positive rate*, FPR) jest obliczany jako iloraz liczby próbek fałszywie dodatnich oraz sumy próbek fałszywie dodatnich i prawdziwie ujemnych [130]. Definicja współczynnika FPR jest następująca:

$$FPR = \frac{FP}{FP + TN} \quad (4.53)$$

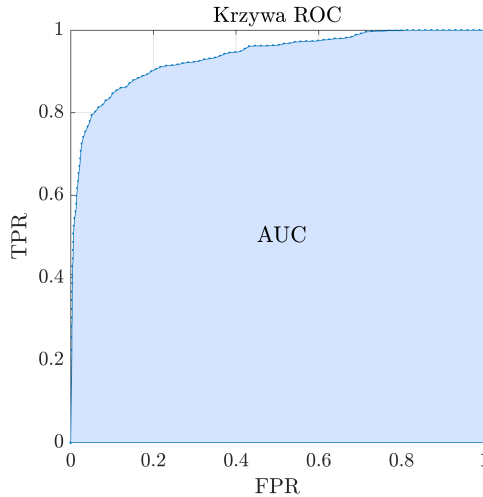
Z powyższych wzorów wynikają wprost następujące zależności:

$$ER = 1 - ACC \quad (4.54)$$

$$FPR = 1 - SPE \quad (4.55)$$

W poprawnie działającym klasyfikatorze wartości zarówno czułości, jak i swoistości powinny osiągać jak największe wartości dla wszystkich przypadków ze zbioru testowego. Krzywa ROC (ang. *receiver operating characteristic*) odzwierciedla odsetek wyników prawdziwie dodatnich w stosunku do odsetka wyników fałszywie dodatnich poprzez prezentację zależności pomiędzy miarami czułości i swoistości [40, 54]. Krzywa ROC stosowana jest celem wytypowania optymalnego punktu odcięcia klasyfikatora, dla którego dokładność klasyfikatora jest maksymalna, jak również do oceny metody klasyfikacji w obszarach cechujących się wysoką czułością i specyficznością (Rys. 4.24) oraz porównania z innymi modelami. Dla poprawy czytelności swoistość przedstawiona jest na takim wykresie jako FPR (Wzór 4.55). Pole pod wykresem krzywej ROC , przyjmujące wartości z zakresu $[0, 1]$, również stosowane jest do oceny klasyfikatora jako miara AUC (ang. *area under curve*) [40]. Im większa wartość pola AUC z uwzględnieniem kształtu krzywej ROC , tym model jest lepszy [108, 130].

Innymi miarami jakości klasyfikacji są wartość predykcyjna dodatniego wy-



Rys. 4.24: Przykład konstrukcji krzywej ROC.

niku oraz wartość predykcyjną ujemnego wyniku. Wartość predykcyjną dodatniego wyniku klasyfikatora (ang. *positive predictive value*, *PPV*) określana również mianem precyzji (ang. *precision*, *PRE*), definiowana jest jako liczba poprawnie sklasyfikowanych próbek prawdziwie dodatnich podzielona przez liczbę próbek sklasyfikowanych jako pozytywne (dodatnie) [108, 121, 130]. Miara *PPV* pozwala oszacować dokładność klasyfikacji próbki w obrębie klasy. Wzór definiujący *PPV* jest następujący:

$$PPV = \frac{TP}{TP + FP} \quad (4.56)$$

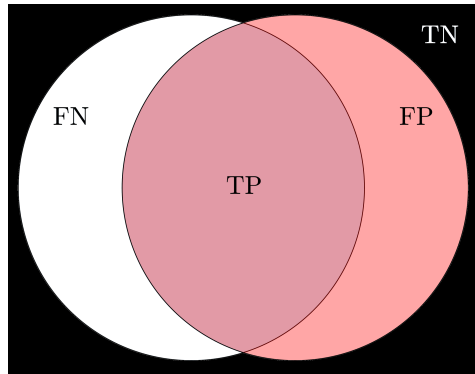
Wartość predykcyjną ujemnego wyniku klasyfikatora (ang. *negative predictive value*, *NPV*) jest stosunkiem liczby prawidłowo sklasyfikowanych próbek prawdziwie ujemnych do liczby próbek sklasyfikowanych negatywnie (ujemnie) [108, 121], zgodnie z zależnością:

$$NPV = \frac{TN}{TN + FN} \quad (4.57)$$

Miara F_1 (ang. *F₁ score*), określana również mianem indeksu (współczynnika) Dice'a (ang. *Dice Index*, *DI*), to średnia harmoniczna wartości *PPV* i *TPR*. Miara F_1 definiowana jest wzorem:

$$F_1 = DI = \frac{2}{PPV^{-1} + TPR^{-1}} = 2 \cdot \frac{PPV \cdot TPR}{PPV + TPR} = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \quad (4.58)$$

Na jej podstawie możliwe jest szacowanie dokładności wykonanego ekspery-



Rys. 4.25: Definicja klasyfikacji pikseli. Biały obiekt odpowiada złotemu standardowi GT , czerwony wynikowi segmentacji S .

mentu, szczególnie w przypadku wykorzystywania danych opisywanych przez człowieka (etykietowanie, obrysy eksperckie) [108, 130].

4.8.2 Miary oceny segmentacji

Ocena efektywności segmentacji widocznego w obrazie obiektu realizowana jest w odniesieniu do tego obiektu obrysowanego przez eksperta, który to obrys stanowi złoty standard (ang. *gold standard* lub ang. *ground truth*, GT). Podobieństwo i różnicowanie pomiędzy maską segmentacji a złotym standardem (obrysem eksperckim) można określić za pomocą indeksu Dice'a (ang. *Dice Index*, DI) [89, 121, 194] opisanego wzorem 4.58 oraz indeksu (współczynnika) Jaccarda (ang. *Jaccard Indeks*, JI) określanego mianem przecięcia sumy (ang. *intersection over union*, IoU), definiowanego następująco [89, 148]:

$$JI = IoU = \frac{|S \cap GT|}{|S \cup GT|} = \frac{DI}{2 - DI} = \frac{TP}{(TP + FP + FN)} \quad (4.59)$$

gdzie: S odpowiada wysegmentowanemu obiektowi, GT obrysowi eksperckiemu, natomiast TP , TN , FP , FN oznaczają odpowiednio liczbę prawdziwie dodatnich (ang. *true positive*), prawdziwie ujemnych (ang. *true negative*), fałszywie dodatnich (ang. *false positive*) i fałszywie ujemnych (ang. *false negative*) wykryć pikseli (Rys. 4.25). Zarówno indeks Dice'a określany również mianem miary F_1 , jak i indeks Jaccarda są wyrażane w postaci ułamkowej lub skali procentowej.

Stosowaną modyfikacją miary JI jest ważony indeks Jaccarda (ang. *weighted IoU*, IoU_w) ważony jest poprzez liczbę pikseli w analizowanej klasie [148, 198], zgodnie z zależnością:

$$IoU_w = \frac{|S \cap (GT \circ W)|}{|S \cup (GT \circ W)|} \quad (4.60)$$

gdzie W jest wartością wagi. Miara ta jest stosowana w przypadku klas o nierównomiernych rozmiarach i pozwala na zmniejszenie błędów w małych klasach. Taka sytuacja występuje w przypadku igły lub małych zmian ogniskowych, gdzie pole powierzchni segmentowanego obiektu jest znacząco mniejsze w porównaniu z tłem.

Podobnie jak w przypadku oceny klasyfikatora, również jakość segmentacji ocenia się za pomocą miar dokładności (wzór 4.49), czułości (wzór 4.51) i swoistości (wzór 4.52), gdzie ocenie podawana jest poprawność identyfikacji pikseli jako segmentowany obiekt lub tło.

4.8.3 Miary dokładności detekcji igły

W przypadku lokalizacji igły metoda oceniana jest w dwojaki sposób: poprzez określenie jakości przeprowadzonej segmentacji gdzie jej wynik ($S(x, y)$) porównywany jest z obrysem eksperckim ($GT(x, y)$) za pomocą miar opisanych w podrozdziale 4.8.2 oraz poprzez zbadanie poprawności wyznaczonej trajektorii igły za pomocą porównania linii prostej obliczonej na podstawie wyniku segmentacji ($R(x, y)$) z linią prostą wytyczoną na podstawie obrysu eksperta ($R_{GT}(x, y)$), których przykład przedstawiono na rysunku 4.26. Wśród miar pozwalających ocenić dokładność detekcji igły w obrazie wymienić można wskaźnik sukcesu lokalizacji igły, wartości błędów trajektorii, średniokwadratowego błędu odległości, średniokwadratowego błędu kąta oraz średniego błędu bezwzględnego.

Wskaźnik sukcesu lokalizacji igły (ang. *needle localization success rate*, $NLSR_i$) informuje o stosunku N obrazów, dla których metoda poprawnie wskazuje linię przechodzącą przez igłę [55].

$$NLSR = \frac{1}{N} \sum_{i=1}^N NLSR_i \quad (4.61)$$

Przypadek $NLSR_i$ jest traktowany jako poprawny, jeśli przecięcie wyniku segmentacji igły ($S_i(x, y)$), obrysu eksperckiego ($GT_i(x, y)$) i obrazu trajektorii ($R_i(x, y)$) jest zbiorem niepustym.

$$NLSR(i) = \begin{cases} 0 & \text{gdzie } (GT_i(x, y) \cap S_i(x, y) \cap R_i(x, y)) = \emptyset \\ 1 & \text{gdzie } (GT_i(x, y) \cap S_i(x, y) \cap R_i(x, y)) \neq \emptyset \end{cases} \quad (4.62)$$

Ocena poprawności lokalizacji igły realizowana jest poprzez porównanie wartości odległości i kątów pomiędzy prostą odzwierciedlającą trajektorię wy-

segmentowanej igły a trajektorią wyznaczoną dla obrysu eksperckiego. Błędy trajektorii odpowiadające odległości $\Delta\lambda$ [55, 111] oraz kątom $\Delta\theta$ [111] są definiowane następująco:

$$\Delta\lambda = |\lambda_{R_{GT}} - \lambda_R| \quad (4.63)$$

$$\Delta\theta = |\theta_{R_{GT}} - \theta_R| \quad (4.64)$$

gdzie $\lambda_{R_{GT}}$ jest najmniejszą odległością pomiędzy linią odpowiadającą igle (obrysowi eksperckiemu) a środkowym pikselem obrazu; λ_R jest najmniejszą odległością pomiędzy linią odpowiadającą wynikowi segmentacji a środkowym pikselem obrazu; $\theta_{R_{GT}}$ jest kątem pomiędzy linią odpowiadającą igle zaznaczonej przez eksperta a osią horyzontalną obrazu, natomiast θ_R jest kątem pomiędzy linią odpowiadającą wysegmentowanemu obiektowi a osią horyzontalną obrazu. Parametry te zostały pokazane na rysunku 4.26.

Błędy lokalizacji trajektorii są klinicznie istotnymi miarami związanymi ze wskazaniem segmentowanej struktury, niemniej jednak silnie zależą od błędu orientacji, dlatego należy je stosować w przypadku odpowiedniego wskazania struktury (w kontekście orientacji) w celu wyeliminowania błędnej interpretacji wyników na podstawie statystyk wskazanych dla niezgodnych lokalizacji [55, 111].

Ocena metody lokalizacji igły dokonywana jest również za pomocą średniokwadratowego błędu odległości ($\Delta\theta_{rms}$) [50, 89] i średniokwadratowego błędu kąta ($\Delta\lambda_{rms}$) [89], które są definiowane następująco:

$$\Delta\lambda_{rms} = \sqrt{\frac{1}{N} \sum_{k=1}^N (\lambda_{R_{GT}} - \lambda_R)^2} \quad (4.65)$$

$$\Delta\theta_{rms} = \sqrt{\frac{1}{N} \sum_{k=1}^N (\theta_{R_{GT}} - \theta_R)^2}. \quad (4.66)$$

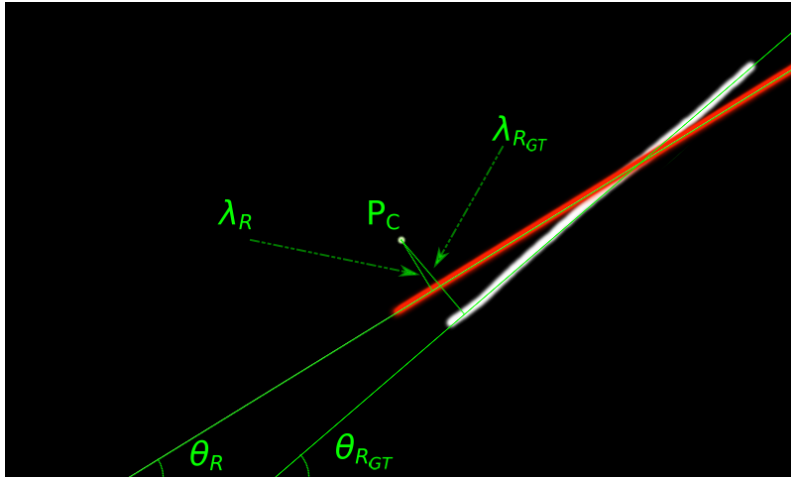
Ocena efektywności detekcji trajektorii igły może zostać zrealizowana również poprzez bezwzględne i względne miary błędów (ang. *absolute error*, Δ i ang. *approximation error*, δ) wartości wyznaczonych za pomocą λ_R :

$$\Delta\lambda = \lambda_R - \lambda_{R_{GT}}, \quad (4.67)$$

$$\delta\lambda = \frac{\Delta\lambda}{\lambda_{R_{GT}}} \cdot 100\%, \quad (4.68)$$

i θ_R :

$$\Delta\theta = \theta_R - \theta_{R_{GT}}, \quad (4.69)$$



Rys. 4.26: Wyjaśnienie parametrów oceny wyznaczonej trajektorii igły. Linia czerwona odpowiada trajektorii igły wyznaczonej za pomocą proponowanej metody. Biała linia odpowiada obrysowi eksperckiemu (złoty standard). P_C jest centralnym punktem obrazu. λ_R i λ_{RGT} stanowią najmniejsze odległości pomiędzy P_C i porównywanymi trajektoriami. θ_R i θ_{RGT} są kątami pomiędzy osią horyzontalną obrazu a odpowiednio wyznaczoną trajektorią na podstawie wyniku segmentacji oraz złotym standardem.

$$\delta_\theta = \frac{\Delta_\theta}{\theta_{RGT}} \cdot 100\%, \quad (4.70)$$

gdzie wyniki wyrażone są odpowiednio w pikselach w przypadku Δ_λ i stopniach dla Δ_θ oraz procentowo dla δ_λ i δ_θ . Wartości bezwzględne i względne miar błędów stanowią podstawę do oszacowania średniego błędu bezwzględnego (ang. *mean absolute error*, *MAE*), który definiowany jest wzorem

$$MAE = \frac{1}{N} \sum_{i=1}^N |\Delta_i|, \quad (4.71)$$

i średniego bezwzględnego błędu procentowego (ang. *mean absolute percentage error*, *MAPE*) wyznaczanego następująco:

$$MAPE = \frac{1}{N} \sum_{i=1}^N |\delta_i|, \quad (4.72)$$

4.9 Podsumowanie

W niniejszym rozdziale przedstawiono propozycję autorskich narzędzi przetwarzania dwuwymiarowych obrazów ultrasonograficznych w kontekście detekcji i segmentacji struktur tkankowych w postaci zmian ogniskowych, jak i sztucznych, których przykładem jest igła biopsyjna.

Przedstawiono protokół przetwarzania wstępnego danych po akwizycji z wykorzystaniem systemu rejestracji danych (Rozdz. 3) oraz wskazano występujące w obrazach artefakty.

Następnie przedstawiono teoretyczne wprowadzenie przedstawiające typy topologii stosowanych architektur opartych o konwolucyjne sieci neuronowe wraz z opublikowanymi przykładami działających rozwiązań, opisano działanie jednostek ukrytych oraz stosowanych powszechnie warstw tworzących rozwiązania oparte o CNN. Porównano warstwy stosowane w przypadku klasyfikacji z warstwami wdrażanymi celem realizacji zadań z zakresu segmentacji semantycznej z uwzględnieniem typów implementowanych funkcji straty. Opisano mechanizmy regularyzacji stosowane celem uniknięcia przeuczenia trenowanych modeli.

W kolejnych dwóch podrozdziałach przedstawiono autorską architekturę klasyfikacji treści obrazów ultrasonograficznych w dwóch ujęciach: (*i*) klasyfikacji binarnej (obecność igły lub jej brak) oraz (*ii*) klasyfikacji wieloklasowej (zmiana echodatnia, zmiana echoujemna oraz brak zmiany–zdrowa tkanka oraz obraz pozbawiony treści diagnostycznej). Dla każdego z rozwiązywanych zagadnień przedstawiono opis proponowanej architektury opartej o CNN, podział danych na 15 zbiorów treningowych, walidacyjnych oraz testowych na poziomie unikalnego identyfikatora pacjenta metodą walidacji krzyżowej, augmentację danych za pomocą przekształceń geometrycznych oraz filtracji celem rozszerzenia zbioru danych, i finalnie proces treningu modeli z wykorzystaniem techniki dostrajania.

W podrozdziale poświęconym detekcji igły biopsyjnej przedstawiono autorską metodę szacowania trajektorii igły opartą o propozycję architektury sieci dedykowanej segmentacji semantycznej, pozwalającej na wyodrębnienie jasnej, liniowej struktury narzędzia względem tła, a następnie szacowaniu trajektorii z wykorzystaniem transformaty Radona. Opis, podobnie jak w przypadku architektury klasyfikującej, uwzględnia szczegóły architektury, technikę podziału danych na zbiory treningowe, walidacyjne oraz testowe z wykorzystaniem metody walidacji krzyżowej, opis augmentacji geometrycznej danych i proces uczenia modeli z uwzględnieniem mechanizmów minimalizujących ryzyko nadmiernego dopasowania modeli do zbiorów treningowego i walidacyjnego.

Przedstawiona została również metoda segmentacji zmiany ogniskowej za pomocą klasycznych metod: rozrostu obszaru, aktywnego konturu, aktywnego konturu uwzględniającego rozszerzenie zasięgu gradientu, progowanie oraz al-

gorytm wododziałowy. Wykorzystanie metod punktowych, lokalnych i krawędziowych umożliwia częściowe uniezależnienie metody od heterogenicznego charakteru obrazów ultrasonograficznych. Binarna maska wynikowa wyznaczana jest poprzez agregację wyników poszczególnych metod za pomocą komitetu klasyfikującego (ang. *ensemble*). Proces segmentacji poprzedzony został analizą porównawczą stosowanych w przypadku obrazów ultrasonograficznych metod filtracji: uśredniającej, medianowej, gaussowskiej, dyfuzji anizotropowej, filtra *shock* oraz średnich nielokalnych. Dla każdego z przefiltrowanych obrazów wyznaczone zostały powszechnie stosowane miary jakości filtracji, które stanowiły podstawę w podejściu konkursowym celem wskazania rozwiązania korzystnego dla obrazów zmian ogniskowych obecnych w gruczole sutkowym.

W niniejszym rozdziale zawarto również opis narzędzi dedykowanych ewaluacji efektywności działania algorytmów przetwarzania obrazów. Miary zostały podzielone na trzy kategorie: *(i)* ocena jakości klasyfikacji, *(ii)* ocena segmentacji oraz *(iii)* dokładności szacowania trajektorii. Wyniki otrzymane za pomocą przywołanych metryk przedstawione zostały w rozdziale 5.

5. Eksperymenty i wyniki

Poniższy rozdział prezentuje wyniki eksperymentów szacujących dokładność elementów składowych niniejszej rozprawy. W pierwszej kolejności prezentowane są dokładności systemu rejestracji danych, następnie przedstawione zostały eksperymenty związane z oszacowaniem wartości hiperparametrów sieci konwolucyjnych, efektywność autorskiej architektury klasyfikacji binarnej i wieloklasowej. Kolejne podrozdziały poświęcone zostały wynikom uzyskanym przez propozycję algorytmów segmentacji struktur z wykorzystaniem metod głębokich w przypadku igły biopsyjnej oraz klasycznych dla zmian ogniskowych.

5.1 System rejestracji danych

Na dokładność działania każdego systemu nawigacji mają wpływ bieżące dokładności lokalizacji znaczników (markerów) oraz wcześniejsze błędy wynikające z kolejnych etapów kalibracji. Wpływ lokalizacji znacznika w scenie, jak również procesu kalibracji na dokładność, został przedstawiony w niniejszym podrozdziale.

5.1.1 Pomiary dokładności systemów śledzenia

Dokładność lokalizacji markerów każdego systemu śledzenia została określona przez producenta Northern Digital Inc. [116, 117] poprzez wyznaczenie wariancji statycznego położenia markera w pomiarze trwającym kilkadziesiąt minut [37]. Błąd lokalizacji znacznika, zgodnie z notą producenta, wynosi 0,25 mm dla systemu optycznego i 0,70 mm dla systemu elektromagnetycznego (Tab. 3.1). Dokładność śledzenia położenia obiektów, na powierzchni których umieszczone są znaczniki, uzależniona jest od dokładności uzyskanej podczas protokołu kalibracji opisanego przez Ungi i wsp. [169]. Niedokładność kalibracji jednego z etapów powoduje spadek dokładności następujących etapów, dlatego kontrola pomiaru dokładności powinna być realizowana po każdym z kroków przewidzianych protokołem i w przypadku zbyt dużej akumulacji błędu, krok ten powinien zostać powtórzony. W przypadku systemu elektromagnetycznego lokalizacja zaburzona jest również przez metalowe elementy występujące w otoczeniu.

Tab. 5.1: Średnie błędy lokalizacji znacznika położonego centralnie w różnej odległości od urządzenia śledzącego.

lp.	NDI Polaris		NDI Aurora	
	odległość [cm]	średni błąd [mm]	odległość [cm]	średni błąd [mm]
1	67	0,145±0,003	16	0,011±0,001
2	83	0,163±0,004	32	0,032±0,006
3	99	0,172±0,008	48	0,064±0,022
4	115	0,207±0,012	64	0,169±0,067

Podczas badania dokładności lokalizacji znaczników zmierzono wartości błędów w czterech różnych odległościach od stereokamery (albo generatora pola) w centralnej części pola widzenia, celem oszacowania wpływu odległości znacznika na dokładność pomiaru. Zebrane i uśrednione wartości błędów dla każdego z systemów śledzenia przedstawiono w tabeli 5.1. Wykonano również pomiary dokładności gdy znacznik zlokalizowany był w skrajnym obszarze pola widzenia. Te wyniki zebrano i przedstawiono w tabeli 5.2. Dla każdej z pozycji zapisano co najmniej 425 pomiarów.

Wartości pomiarów błędów lokalizacji znaczników w centrum pola widzenia porównano za pomocą testu t-Studenta dla prób niezależnych celem weryfikacji, czy uzyskane wyniki pomiarów cechują się różnicami błędów w zależności od odległości od urządzenia śledzącego. Otrzymane wartości wykazują istotne statystycznie różnice dla błędów lokalizacji znaczników ($p < 0,0001$) zarówno w przypadku systemu optycznego Polaris Vica, jak również systemu elektromagnetycznego Aurora.

Porównanie pozycji centralnej (cechującej się najmniejszym średnim błędem lokalizacji znaczników) z wartościami wyznaczonymi dla skrajnych punktów pola widzenia za pomocą testu t-Studenta dla prób niezależnych, odpowiednio dla każdego z systemów śledzenia wykazało, że wyznaczone wartości porównania centralnej pozycji ze skrajnymi, jak również wartości skrajnych między sobą, wskazują na występowanie istotnych statystycznie różnic ($p < 0,0001$).

Późniejsze kroki kalibracji realizowano w centrum pola widzenia, w najmniejszej możliwej odległości pomiędzy znacznikiem a urządzeniem śledzącym ze względu na największą dokładność w tym obszarze.

5.1.2 Kalibracja wskaźnika

Pierwszy etap protokołu kalibracji systemu śledzącego to kalibracja końcówki wskaźnika wykonana zgodnie z metodologią opisaną w podrozdziale 3.3.2. Współrzędne znacznika zataczającego ruch po sferze stanowią podstawę do wyznaczenia przekształcenia geometrycznego umożliwiającego lokalizację koń-

Tab. 5.2: Średnie błędy lokalizacji znacznika położonego w brzeżnym obszarze pola widzenia urządzenia śledzącego w odległości 67 cm i 16 cm odpowiednio dla systemu Polaris i Aurora.

lp.	położenie	NDI Polaris średni błąd	NDI Aurora średni błąd
1	centrum	0,145±0,003	0,011±0,001
2	góra	0,213±0,010	0,340± 0,086
3	dół	0,187±0,015	0,137±0,049
4	prawo	0,201±0,007	0,268±0,104
5	lewo	0,134±0,010	0,244±0,094

Tab. 5.3: Średnie błędy kalibracji wskaźników stosowanych w systemach nawigacji.

system śledzenia	średni błąd [mm]
NDI Polaris	0,146±0,032
NDI Aurora	0,104±0,028

cówki wskaźnika.

Podczas badania dokładności procesu kalibracji wskaźnika obu systemów śledzenia wykonano po trzydzieści cztery powtórzenia kalibracji. Wyznaczone średnie wartości błędów dla wskaźników obu modalności umieszczono w tabeli 5.3. Wartości mieszczą się w zakresie wartości wskazywanym dla typowych błędów kalibracji wskaźnika szacowanych na poziomie <0,5 mm [86]. Otrzymane wyniki wskazują na mniejsze wartości błędów w przypadku systemu elektromagnetycznego.

Porównanie błędów kalibracji wskaźników dedykowanych wykorzystanym systemom śledzenia przeprowadzono za pomocą testu t-Studenta dla prób niezależnych. Wyznaczone wartości porównania błędów kalibracji wskazują na występowanie istotnych statystycznie różnic ($p < 0,0001$).

5.1.3 Kalibracja pistoletu biopsyjnego

Pistolet biopsyjny jest obiektem, na którego powierzchni jest umieszczony nieruchomo pasywny znacznik systemu optycznego. Pistolet z igłą jest rodzajem wskaźnika, który można skalibrować zgodnie z założeniami protokołu kalibracji [169], gdzie lokalizacja końcówki igły wyznaczana jest w oparciu o ruch po sferze sfer budujących znacznik. Igła, ze względu na dwuelementową budowę i małą średnicę każdego z elementów, jest obiektem giętkim, co skutkuje wysokimi błędami kalibracji. Z tego powodu przekształcenie geometryczne umożliwiające wskazanie końcówki igły, zostało zrealizowane z użyciem zmodyfikowanej metody kalibracji igły, która została opisana w podrozdziale 3.3.3.

Tab. 5.4: Średnie błędy kalibracji pistoletu biopsyjnego [mm].

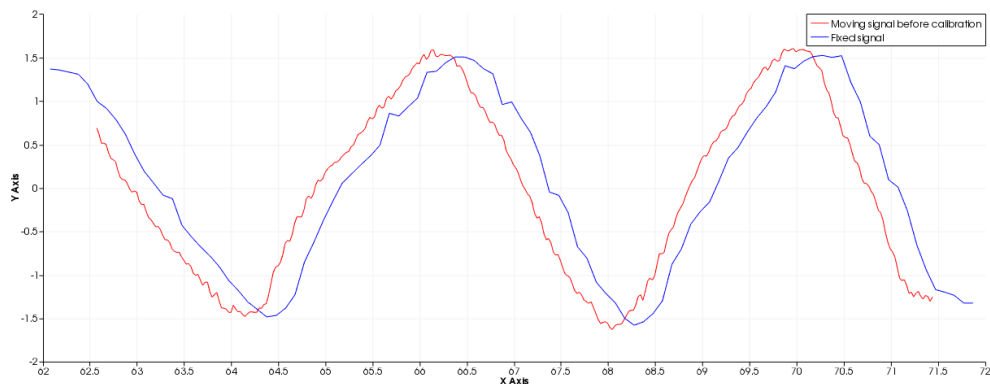
metoda kalibracji	średni błąd
metoda opisana w rozdz. 3.3.2	1,392±0,379
metoda opisana w rozdz. 3.3.3	0,186±0,041

W tabeli 5.4 przedstawiono średnie wartości błędów wyznaczonych dla 34 powtórzeń kalibracji dwoma metodami, które realizowane były w tej samej odległości 67 cm od kamery systemu optycznego. Wartości te porównano za pomocą testu t-Studenta dla prób niezależnych. Wykazano, że wartości błędów kalibracji dla każdej z metod wskazują na występowanie istotnych statystycznie różnic pomiędzy nimi ($p < 0,0001$).

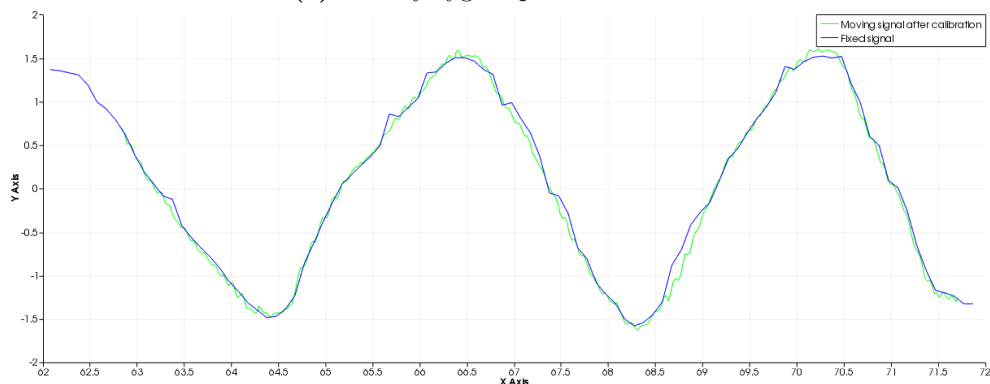
5.1.4 Dokładność rejestracji fantomu

Kolejny etap kalibracji systemu pomiarowego stanowi wyznaczenie przekształcenia geometrycznego pomiędzy fantomem pomiarowym a znacznikiem umieszczonym na nim nieruchomo. Fantom pomiarowy fCal stanowi integralny element pełnej kalibracji głowicy ultrasonograficznej (Podrozdz. 3.3.3), której celem jest wyznaczenie transformacji z układu współrzędnych dwuwymiarowego obrazu ultrasonograficznego do układu współrzędnych znacznika, który jest nieruchomo umieszczony na powierzchni głowicy liniowej [19]. Geometria oraz wymiary fantomu kalibracyjnego wraz z punktami charakterystycznymi są znane, co pozwala na zastosowanie go, celem określenia dokładności systemu śledzenia (Rys. 3.11).

Kalibracja fantomu pomiarowego zakłada wskazanie za pomocą wskaźnika współrzędnych każdego z ośmiu charakterystycznych punktów umieszczonych na ścianach fantomu. Proces wskazywania punktów powtórzono pięciokrotnie w krótkich odstępach czasu z uwzględnieniem zmiany położenia wskaźnika względem generatora pola elektromagnetycznego. Położenie fantomu względem FG nie ulegała w trakcie serii pomiarowej. Wartości współrzędnych zebranych w trakcie serii pomiarowej zostały uśrednione i wykorzystane jako właściwe współrzędne punktu do wyznaczenia macierzy transformacji pomiędzy układami współrzędnych. Wartości błędów dokładności wyznaczono na podstawie rzeczywistych, zapisanych wartości współrzędnych. Serie pomiarowe wykonano trzykrotnie: w odległości 16 cm, 32 cm i 48 cm w linii prostej od monitora lokalizacji, co zostało podyktowane polem widzenia systemu śledzenia. Wykonywane pomiary przeprowadzane były zarówno w układzie współrzędnych elektromagnetycznego systemu śledzenia jak również w układzie współrzędnych znacznika referencyjnego umieszczonego nieruchomo na fantomie.



(a) Surowy sygnał położenia obrazu



(b) Sygnał położenia obrazu uwzględniający przesunięcie względem czasu

Rys. 5.1: Wykresy przebiegów znormalizowanych sygnałów położenia obrazu: 5.1a przed i 5.1b po eliminacji opóźnienia w czasie akwizycji obrazu USG. Kolor niebieski odpowiada sygnałowi zapisanemu dla obrazu USG, kolor czerwony położeniu głowicy przed eliminacją opóźnienia, natomiast kolor zielony prezentuje sygnał głowicy po synchronizacji.

5.1.5 Kalibracja głowicy ultrasonograficznej

Pomiar kalibracji czasowej obrazu USG

Kalibracja czasowa obrazu ultrasonograficznego realizowana jest celem eliminacji opóźnienia akwizycji obrazu względem informacji o położeniu głowicy liniowej. Kalibracja czasowa została przeprowadzona zgodnie z przyjętym protokołem [169], który został opisany w podrozdziale 3.3.4. Przykładowy wykres prezentujący sygnał oryginalny i przesunięty w czasie został przedstawiony na rysunku 5.1. Podczas wyznaczania wartości opóźnienia wykonano dziesięciokrotne powtórzenie tego kroku. Na podstawie zebranych wyników, oszacowano średnią wartość równą $268,830 \pm 5,927$ ms.

Pomiar kalibracji przestrzennej obrazu USG

Ogólna ocena dokładności kalibracji przestrzennej głowicy ultrasonograficznej realizowana jest za pomocą średniego błędu bezwzględnego MAE [86], zgodnie ze wzorem 4.71. Protokół kalibracji przestrzennej głowicy liniowej śledzonej z użyciem elektromagnetycznego systemu Aurora powtórzono 34 krotnie, uzyskując średnią wartość błędu MAE wynoszącą $0,682 \pm 0,152$ mm.

Ocena dokładności kalibracji przestrzennej głowicy ultrasonograficznej zrealizowana została również za pomocą miary powtarzalności kalibracji (ang. *Calibration Reproducibility, CR*), ze względu na wpływ oddziaływania grubości wiązki ultrasonograficznej na lokalizację punktów w przestrzeni trójwymiarowej. Miara ta porównuje N niezależnie zrealizowanych kalibracji przestrzennych głowicy z uwzględnieniem nieruchomo umieszczonego znacznika, w tej samej pozycji dla wszystkich prób [19, 132].

Efekt kalibracji stanowi zbiór N macierzy rotacji i translacji $(M_U^P)_i$, określających zależność pomiędzy obrazem a znacznikiem głowicy. Dla każdego obrazu wytypowano 5 pikseli odpowiadających środkowi i narożnikom. Każdy z pikseli został poddany N -transformacjom do układów współrzędnych znacznika za pomocą wyznaczonych protokołem kalibracji macierzy. Następnie obliczono dystans pomiędzy współrzędnymi transformowanych wokseli a wartościami współrzędnych odpowiadających im punktów w przestrzeni. Powtarzalność kalibracji CR_i definiowana jest wzorem:

$$CR_i = \frac{1}{N} \sum_{j=1}^N \|(M_U^P)_i p_{i,j}^U - p_i^{\bar{P}}\|, \quad (5.1)$$

gdzie $p_{i,j}^U$, $i \in 1, \dots, 5$, $j \in 1, \dots, N$ oznacza i -ty piksel odpowiadający narożnikom i środkom obrazów na j -tym obrazie ultrasonograficznym, natomiast $p_i^{\bar{P}}$ jest centroidem przekształconych pikseli:

$$p_i^{\bar{P}} = \frac{1}{N} \sum_{j=1}^N (M_U^P)_i p_{i,j}^U \quad (5.2)$$

Na podstawie przeprowadzonego protokołu kalibracji przestrzennej głowicy z nieruchomo przymocowanym znacznikiem kompatybilnym z systemem EMTS wyznaczono średnie wartości powtarzalności kalibracji, które przedstawiono w tabeli 5.5.

Typowy błąd kalibracji przestrzennej mieści się w zakresie $[0, 8; 1, 5]$ mm [86]. Otrzymane wartości są porównywalne z sugerowanym zakresem oraz wartościami kalibracji przedstawianymi w literaturze [19, 132]. Większe wartości miary CR w przypadku narożników w porównaniu do punktu środkowego powodowane są odległością od osi rotacji. Woksel środkowy jest bliżej niej, co

Tab. 5.5: Powtarzalność kalibracji (CR) głowicy ultrasonograficznej [mm].

punkt	CR
lewy górny	1,174
prawy górny	1,118
lewy dolny	0,544
prawy dolny	0,710
środek	0,476
M	0,804

skutkuje mniejszym wpływem rotacji na jego położenie i zwiększa powtarzalność wyników.

5.2 Dane źródłowe

W części eksperymentalnej niniejszej pracy wykorzystywano obrazy ultrasonograficzne pochodzące z różnych źródeł. Do oceny metod klasyfikacji igły biopsyjnej i wyznaczania jej trajektorii zakwalifikowano tylko te zbiory źródłowe, w których wykonywana była procedura biopsji gruboigłowej. Informacje na temat pochodzenia źródłowych danych obrazowych przedstawiono w następujących dwóch podrozdziałach, 5.2.1 oraz 5.2.2.

5.2.1 Obrazy USG z zapisem biopsji gruboigłowej

Ultrasonograficzne dane obrazowe, które w niniejszej pracy wykorzystano w algorytmach detekcji igły biopsyjnej oraz wyznaczania jej trajektorii, pochodzą z dwóch źródeł danych.

Pierwsze z nich, zawierające 268 obrazów USG wykonanych u 58 pacjentów, jest zbiorem danych, które zostało udostępnione publicznie przez autorów na stronie internetowej www.ultrasoundcases.info [154]. Strona internetowa <https://www.ultrasoundcases.info/> została stworzona we współpracy z Hitachi Medical Systems European i Wydziałem Radiologii szpitala Gelderse Vallei w Ede w Holandii. Obrazy pochodzące z tego zbioru mają różną rozdzielczość i zostały wykonane z nieokreślonymi parametrami akwizycji. Brak jest również informacji o aparatach USG, którymi wykonano obrazy, oraz o sposobie konwersji obrazu video do plików JPEG. Pomimo tych niedogodności, zbiór ten został wykorzystany w innych pracach badawczych związanych z segmentacją struktur w obrazach USG. Publicznie dostępne obrazy ultrasonograficzne zmian ogniskowych zostały wykorzystane w propozycjach metod automatycznej segmentacji zmian ogniskowych przedstawionych przez Zhuang i wsp. [205], Kriti i wsp. [82] oraz Wang i wsp. [173], jak również metodach klasyfikacji cha-

Tab. 5.6: Obrazy USG wykorzystane w niniejszej rozprawie w zakresie oceny metod detekcji igły biopsyjnej.

klasa	źródło 1 [154]	źródło 2	łącznie
igła obecna	81	538	619
brak igły	187	690	877
łącznie	268	1228	1496

rakteru zmiany ogniskowej przez Zhuang et al. [204]. Witryna ta jest również źródłem obrazów ultrasonograficznych wykorzystywanych do proponowania algorytmów wspomagających diagnostykę choroby tarczycy [106], choroby wątroby [74] lub zmian skórnych [35]. W tym zbiorze, na 81 spośród wszystkich 268 obrazów widoczna jest igła biopsyjna.

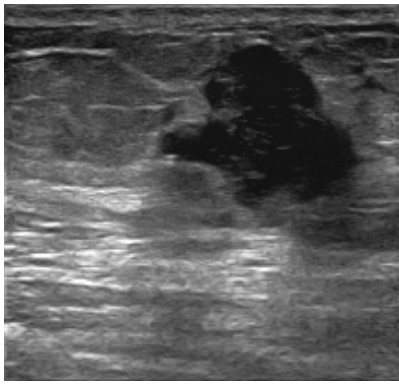
Drugi zbiór danych obrazowych został zebrany z wykorzystaniem stanowiska pomiarowego opisanego w rozdziale 3.1 niniejszej pracy podczas 204 procedur biopsji gruboigłowych wykonanych u 33 pacjentek przez lekarzy specjalistów radiologii w Oddziale w Gliwicach Narodowego Instytutu Onkologii im. Marii Skłodowskiej-Curie w trakcie realizacji projektu naukowo-badawczego pt. MILESTONE (projekt w ramach programu NCBiR Strategmed II [133]). Biopsja była przeprowadzana pod kontrolą ultrasonografu Hitachi EZU-MT25-S1 z głowicą liniową EUP-L6, na której umieszczony był czujnik elektromagnetyczny, uprzednio skalibrowany zgodnie z protokołem przygotowanym przez Lasso i wsp. [86]. Zapis video monitora ultrasonografu przechwytywano za pomocą konwertera S-Video na USB (Imaging Source DFG/USB2pro), który był elementem stanowiska pomiarowego.

Maksymalna głębokość obrazowania ultradźwiękowego pozostawała w przedziale od 25 do 40 mm w zależności od umiejscowienia zmiany ogniskowej, a częstotliwość pracy głowicy wynosiła od 6 do 14 MHz. Obrazy pobierano w rozdzielczości 768×576 z częstotliwością 25 ramek na sekundę.

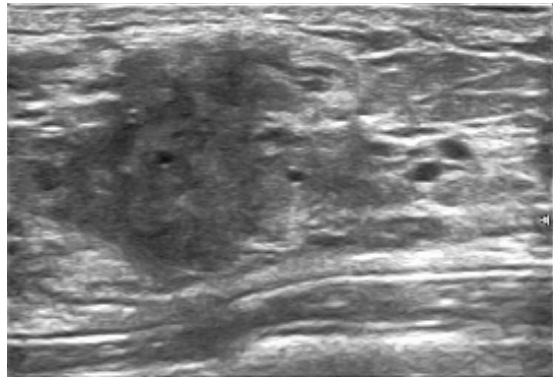
Ze wszystkich pobranych sekwencji video wybrano losowo 1228 obrazów, a na 538 z nich igła biopsyjna jest widoczna całkowicie lub częściowo. Przykładowe obrazy przedstawiono na rysunku 5.2.

Obrysy eksperckie igły biopsyjnej

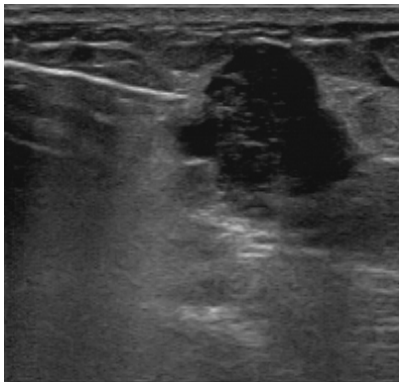
Oba przedstawione zbiory danych finalnie zostały połączone tworząc wspólny zbiór zawierający 1496 obrazów USG gruczołu sutkowego. Wszystkie obrazy ze zbioru zostały ocenione przez lekarza (zwanego dalej ekspertem) mającego doświadczenie kliniczne z tego typu obrazami. Stwierdził on, że na 619 obrazach widoczna jest igła biopsyjna. Informacje o obrazach przedstawiono zbiorczo w tabeli 5.6.



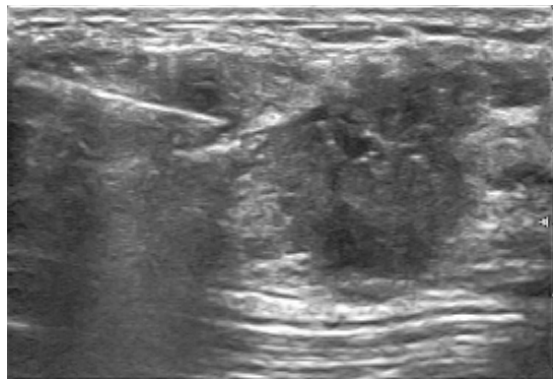
(a)



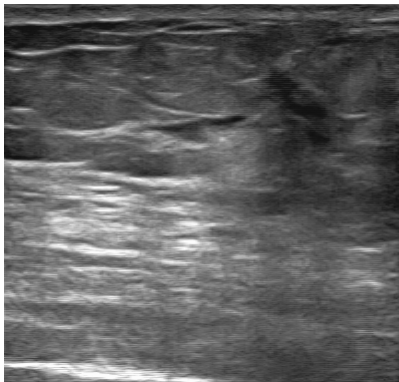
(b)



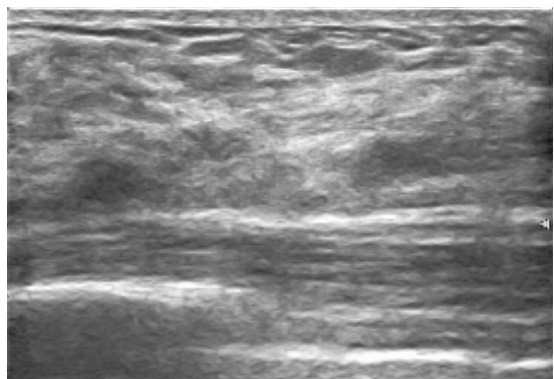
(c)



(d)



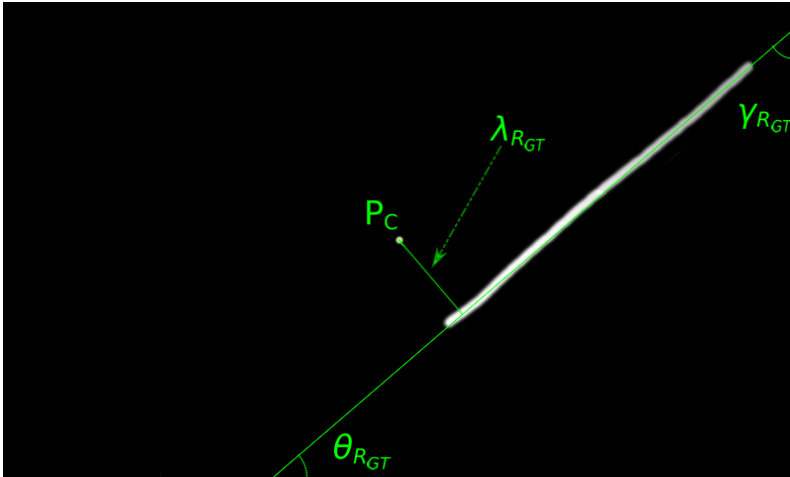
(e)



(f)

Rys. 5.2: Zmiana ogniskowa echoujemna (a) i echodatnia (b). Te same zmiany z widoczną igłą biopsyjną (odpowiednio (c) i (d)). Obraz zdrowych tkanek gruczołu sutkowego sąsiadujących ze zmianami (odpowiednio (e) i (f)).

Na rysunku 5.3 przedstawiono finalną maskę obrysu eksperckiego igły, wi-



Rys. 5.3: Obrys ekspercki igły (biała linia). Wyznaczona trajektoria igły (zielona prosta). P_C – centralny punkt obrazu. λ_{RGT} odległość od P_C do trajektorii. $\angle\gamma_{RGT}$, $\angle\theta_{RGT}$ – kąty między trajektorią a osiami obrazu odpowiednio Y i X .

doczną jako białą, względnie prostą, linię. Wszystkie obrysy igły zostały utworzone przez jej ręczne oznaczenie na komputerze PC w edytorze graficznym KolourPaint z użyciem myszy komputerowej narzędziem „pędzel” o arbitralnie wybranej grubości. Obrysy te stanowią złoty standard (ang. *ground truth, GT*) obszaru igły.

Rzeczywista trajektoria igły obliczona została na podstawie obrysu eksperckiego poprzez zastosowanie transformaty Radona [166], i przedstawiona została na rysunku 5.3 zieloną linią prostą. Następnie wyznaczono wartość λ_{RGT} będącą odległością od punktu centralnego obrazu P_C do wyznaczonej trajektorii. Zmierzony został również kąt wprowadzenia igły względem czoła głowicy USG opisany jako kąt $\angle\gamma_{RGT}$ będący kątem pomiędzy trajektorią a osią Y obrazu, oraz kąt $\angle\theta_{RGT}$ czyli kąt między trajektorią a osią X . Pomiędzy tymi kątami zachodzi zależność [38]:

$$\angle\theta_{RGT} = 90^\circ - \angle\gamma_{RGT} \quad (5.3)$$

W następnej kolejności wyznaczono procentowy obszar igły zaznaczonej przez eksperta względem całego obrazu, P_{wzg} .

W pierwszym zbiorze danych igła jest widoczna w 81 obrazach i została wprowadzona pod średnim kątem $\angle\gamma_{RGT}$ wynoszącym $76,07^\circ \pm 8,54^\circ$. Odległość λ_{RGT} wynosiła średnio $44,23 \pm 25,29$ pikseli. Średnia wartość S_{wzg} wynosi $0,72\% \pm 0,37\%$.

Igła widoczna w zapisanych obrazach drugiego zbioru ma średnią wartość kąta $\angle\gamma_{RGT}$ równą $73,22^\circ \pm 7,75^\circ$. Odległość λ_{RGT} wynosi średnio $72,06 \pm 31,79$

Tab. 5.7: Podsumowanie parametrów igły widocznej w dwuwymiarowych obrazach w połączonej bazie danych. $\angle\gamma_{RGT}$ stanowi kąt pomiędzy prostą odpowiadającą trajektorii igły biopsyjnej a osią wertykalną obrazu. $\angle\theta_{RGT}$ jest kątem pomiędzy prostą odpowiadającą trajektorii igły biopsyjnej a osią horyzontalną obrazu. λ_{RGT} jest najmniejszą odległością pomiędzy punktem centralnym obrazu P_C a trajektorią igły (Rys. 5.3). Udział procentowy P_{wzg} odzwierciedla obszar igły do całego analizowanego obrazu.

stat.	$\angle\gamma_{RGT}$ [°]	$\angle\theta_{RGT}$ [°]	λ_{RGT} [pix]	P_{wzg} [%]
M	73,60	16,40	68,43	0,81
Me	73,90	16,10	65,94	0,71
SD	7,92	7,92	32,41	0,39
min	49,56	0,00	4,22	0,16
max	90,00	40,44	165,35	2,94

pikseli, natomiast średnia wartość S_{wzg} to $0,82\% \pm 0,39\%$.

Uzyskane wyniki w obu zbiorach danych nie są istotnie statystycznie różne. Należy zwrócić uwagę, że finalna średnia wartość $\angle\gamma_{RGT}$ równa $73,60^\circ \pm 7,92^\circ$ nie odbiega od średnich wartości wskazywanych w pracy Mwikirize i wsp. [111]. Na 276 analizowanych obrazach igła wprowadzana była z lewej strony, a na 343 obrazach – ze strony prawej. Pełne wyniki statystyki opisowej połączonych zbiorów danych zawarto w tabeli 5.7.

Przetwarzanie wstępne obrazów

Obrazy wraz z maskami poddane zostały przetwarzaniu wstępnemu w celu ujednoczenia liczby kanałów oraz szerokości i wysokości, ze względu na wymagania dalszych etapów przetwarzania. W pierwszej kolejności usunięto przeplot metodą najbliższego sąsiada, a następnie dokonano zamiany przestrzeni kolorów RGB do ośmiobitowej skali szarości. Następnie przeskalowano obraz za pomocą interpolacji metodą najbliższego sąsiada do wymiarów 256×256 pikseli (*vide* podrozdział 4.1).

5.2.2 Obrazy USG ze zmianami ogniskowymi

Zbiór danych obrazowych wykorzystanych w klasyfikacji zmian ogniskowych gruczołu stukowego pochodzi z pięciu źródeł. W tabeli 5.8 przedstawiono liczebności zbiorów źródłowych, z rozróżnieniem zmian echoujemnych, echodatnych oraz obrazów bez zmian ogniskowych.

Pierwsze źródło jest bazą danych zebraną i udostępnioną publicznie przez szpital Baheya Hospital for Early Detection & Treatment of Women's Cancer (Kair, Egipt), zawierającą 779 obrazów USG wykonanych u 561 pacjentek w wieku od 25 do 75 lat [34]. Obrazy pozyskiwane były za pomocą aparatów

Tab. 5.8: Liczebności pięciu źródłowych zbiorów obrazów USG wykorzystywanych do klasyfikacji i detekcji zmiany ogniskowej.

stat.	źródło 1 [34]	źródło 2 [197]	źródło 3 [136]	źródło 4 [154]	źródło 5	łącznie
zmiana echoujemna	410	135	185	130	53	913
zmiana echodatnia	236	28	65	135	219	683
brak zmiany	133	0	0	0	956	1089
łącznie	779	163	250	265	1228	2685

USG LOGIQ E9 i LOGIQ E9 Agile wyposażonych w głowicę liniową ML6-15-D Matrix. Autorzy nie podali dokładniejszych parametrów akwizycji. Udostępnione obrazy mają różne wymiary. Do każdego z obrazów wykonano ręczny obrys ekspercki.

Baza danych wykorzystywana była dotąd w wielu pracach związanych z klasyfikacją typu zmiany (łagodna *vs* złośliwa), na przykład przez Moon i wsp. [104], Sadad i wsp. [140] albo Wan i wsp. [172], a także w pracach związanych z segmentacją zmiany ogniskowej metodami klasycznymi przedstawionymi np. przez Sadad i wsp. [140] albo Ma i wsp. [96] oraz metodami opartymi o głębokie uczenie zaprezentowanymi przez Byra i wsp. [17] albo Gomez-Flores i wsp. [47].

Drugim wykorzystanym zestawem danych jest baza udostępniona przez zespół Yap i wsp. [197] zawierająca 163 obrazy zmian łagodnych i złośliwych (odpowiednio 110 i 53 obrazy) pochodzące od 142 pacjentek. Badania USG wykonywano w UDIAT Diagnostic Centre of the Parc Taulí Corporation, Sabadell (Hiszpania) za pomocą aparatu ultrasonograficznego Siemens ACUSON Sequoia C512 wyposażonego w głowicę 17L5 HD (8,5 MHz). Do każdego obrazu ze zmianą ogniskową dołączono również obrys ekspercki. Zbiór ten został wykorzystany w pracach poświęconych klasyfikacji [196] oraz segmentacji semantycznej [17, 47, 196].

Następny zbiór to publiczna baza danych udostępniona przez Rodriguesa i wsp. w 2017 roku [136]. Liczy on 250 obrazów łagodnych i złośliwych zmian ogniskowych, które zostały wykonane u 50 pacjentek. Obrazy w tej bazie cechują się różnymi wymiarami i parametrami akwizycji. Brak jest informacji o aparatach USG, za pomocą których zostały zebrane dane, jak również nie udostępniono obrysów eksperckich. Baza danych została wykorzystana w pracach związanych z klasyfikacją typu zmian ogniskowych (łagodna *vs* złośliwa), wśród których wymienić można m.in. Ramachandran i wsp. [87, 134, 172].

Czwarty zestaw danych stanowi publicznie dostępna baza obrazów SonoSkills [154], którą w niniejszej pracy wykorzystano również w klasyfikacji i de-

tekcji igły biopsyjnej. Z tego zbioru wykorzystano 265 obrazów z widoczną zmianą ogniskową. Dane pochodzą od 58 pacjentek, nie ma jednak informacji o aparatach ultrasonograficznych oraz parametrach akwizycji. Do każdego z obrazów zmiany ogniskowej utworzono ręcznie jej obrys zgodnie z metodą przygotowywania obrysów igły biopsyjnej, który został przedstawiony w podrozdziale 5.2.1. Zbiór ten wykorzystywano m.in. w pracy dotyczącej segmentacji zmiany ogniskowej opracowanej przez Kriti i wsp. [82] oraz Wang i wsp. [173].

Ostatni zbiór danych obrazowych jest tożsamy z ze zbiorem 1228 obrazów wykorzystywanych w klasyfikacji i detekcji igły biopsyjnej w niniejszej pracy, opisanym w poprzednim podrozdziale. Wszystkie zmiany ogniskowe zostały obrysowane przez eksperta w podobny do igły biopsyjnej sposób (Podrozdz. 5.2.1). Przykładowe obrazy przedstawiono na rysunku 5.2.

Opisane bazy danych zostały połączone tworząc zbiór zawierający 2685 dwuwymiarowych obrazów ultrasonograficznych gruczołu sutkowego (Tab. 5.8). Dane zawarte w połączonej bazie podzielono na trzy kategorie w zależności od prezentowanej treści: zmiana echoujemna (Rys. 5.2a), zmiana echodatnia (Rys. 5.2b) oraz brak zmiany (Rys. 5.2e, Rys. 5.2f). Obraz ultrasonograficzny zmiany ogniskowej echoujemnej cechuje się znacznie obniżoną echogenicznością, aż do całkowitego braku echa. Do zmian ogniskowych echodatnich zaklasyfikowano obrazy o wzmożonej echogeniczności większej niż otaczające tło [99, 179].

Zbiór obrazów ze zmianami echoujemnymi zawiera 913 obrazów. Średnia wartość pola powierzchni P zmiany wynosi 21792,96 piksela, co stanowi średnio 8,75% całego obrazu USG. Wartości długości osi głównych L_1 i L_2 wynoszą średnio 192,85 i 116,57 piksela. Kierunek R_θ zmiany przyjmuje w tym przypadku wartość 2,28°.

Średnie wartości parametrów 683 obrazów zmian echodatnich mieszczą się w podobnych zakresach. Pole powierzchni P to średnio 27633,54 piksela. Pole powierzchni zmiany średnio stanowi 13,03% całego obrazu ultrasonograficznego. Wartości długości osi głównych L_1 i L_2 przyjmują średnio 227,54 i 204,10 piksela. Wartość kierunku obiektu R_θ przyjmuje wartość 0,16°. Wartości te zebrano i przedstawiono w tabeli 5.9).

Przetwarzanie wstępne obrazów

Obrazy wykorzystywane do oceny segmentacji zmian ogniskowych zostały poddane temu samemu przetwarzaniu wstępnemu, co obrazy wykorzystywane w klasyfikacji igły. Przetwarzanie to ma na celu usunięcie przeplotu video, konwersję do skali szarości oraz przeskalowanie do wymiarów 256×256 (*vide* podrozdział 4.1).

Tab. 5.9: Parametry geometryczne zmian ogniskowych echoujemnych i echodatnich. P – pole powierzchni zmiany ogniskowej, L_1, L_2 – długości osi głównych. R_θ – kierunek obiektu. P_{wzg} – procentowy obszar pola powierzchni zmiany ogniskowej względem całego obrazu.

stat.	P [pix]	L_1 [pix]	L_2 [pix]	R_θ [°]	P_{wzg} [%]
zmiany ogniskowe echoujemne					
M	21792,96	192,85	116,57	2,28	8,75
Me	13241,00	174,78	101,21	1,63	5,38
SD	24277,47	109,86	69,24	24,24	9,39
min	500,00	29,67	16,80	-89,53	0,24
max	176271,00	597,03	391,87	89,84	55,62
zmiany ogniskowe echodatnie					
M	27633,54	227,54	131,58	0,16	13,03
Me	17306,00	204,10	112,90	-0,14	9,09
SD	29869,94	117,12	74,86	27,76	12,87
min	569,00	37,97	21,61	0,-89,54	0,21
max	209121,00	687,76	410,10	89,93	67,31

5.3 Klasyfikacja i segmentacja struktur w obrazie USG

W poniższym podrozdziale przedstawiono rezultaty części eksperymentalnej poświęconej określeniu dokładności klasyfikacji i segmentacji struktur w obrazach ultrasonograficznych za pomocą autorskich metod opartych o metody klasyczne oraz głębokie.

5.3.1 Klasyfikacja obrazów USG z igłą biopsyjną

Baza danych zawierająca 1496 dwuwymiarowych obrazów ultrasonograficznych (Tab. 5.6) została losowo podzielona na trzy zbiory na podstawie identyfikatorów pacjentów: zbiór treningowy, zbiór walidacyjny oraz zbiór testowy, za pomocą mechanizmu k-krotnej walidacji krzyżowej zgodnie z opisem zawartym w podrozdziale 4.4.2. Przy pomocy zbiorów treningowego i walidacyjnego przeprowadzono naukę autorskiego modelu klasyfikacji obecności igły w obrazie. Efektywność proponowanego rozwiązania została zweryfikowana za pomocą zbioru testowego. Proces treningu i testowania został powtórzony 15-krotnie z użyciem 15 losowych zestawów składających się z trzech zbiorów każdy.

Implementacja proponowanego klasyfikatora została wdrożona w środowisku MATLAB 2021a. Eksperymenty zostały przeprowadzone na komputerze PC pod kontrolą systemu operacyjnego Windows 10 z procesorem AMD Ryzen 9 3900 (CPU), pamięcią operacyjną DDR4 32 GB (RAM) 3 GHz, z kartą graficzną NVIDIA RTX Titan 24 GB (GPU).

Tab. 5.10: Podsumowanie testowanych hiperparametrów autorskiego klasyfikatora obecności igły.

parametr	testowany zakres	wybrana wartość
liczba epok	1–50	–
wymiary obrazu	$16 \times 16 \times 1$ – $256 \times 256 \times 1$	$256 \times 256 \times 1$
optymalizator	ADAM, SGDM	ADAM
współczynnik tempa uczenia	0–1	0,15
wagowanie funkcji straty	0,05–0,95	–
regularyzator typu <i>dropout</i>	0–0,99	0,1

Autorski model klasyfikacji obecności igły został oceniony pod względem liczby epok realizowanych w procesie treningu oraz pod kątem wymiarów obrazów wejściowych. Przetestowano wartości liczby epok z zakresu od 1 do 50 oraz następujące wymiary obrazu podawanego na wejściu: 16×16 , 32×32 , 64×64 , 128×128 , 224×224 i 256×256 . Ocenie poddano również typ optymalizatora i współczynnik tempa uczenia w zakresie od 0 do 1 z krokiem 0,05. Ze względu na niezbalansowany zbiór danych przetestowano parametr ważenia funkcji straty oraz współczynnik prawdopodobieństwa p zerowania wag ukrytych jednostek za pomocą regularyzatora typu *dropout*. Zakresy testowanych hiperparametrów zebrano w tabeli 5.10.

Wyniki klasyfikacji autorską siecią dla poszczególnych hiperparametrów oceniono i porównano za pomocą, szczegółowo opisanych w podrozdziale 4.8.1, miar: dokładności (ACC), miary F_1 , współczynnika wyników prawdziwie pozytywnych TPR (czułości SEN), współczynnika wyników prawdziwie negatywnych TNR (swoistości SPE), wartości predykcyjnej dodatniego wyniku (PPV) i wartości predykcyjnej ujemnego wyniku (NPV). Oszacowany również został średni czas t klasyfikacji jednego obrazu przez proponowany model.

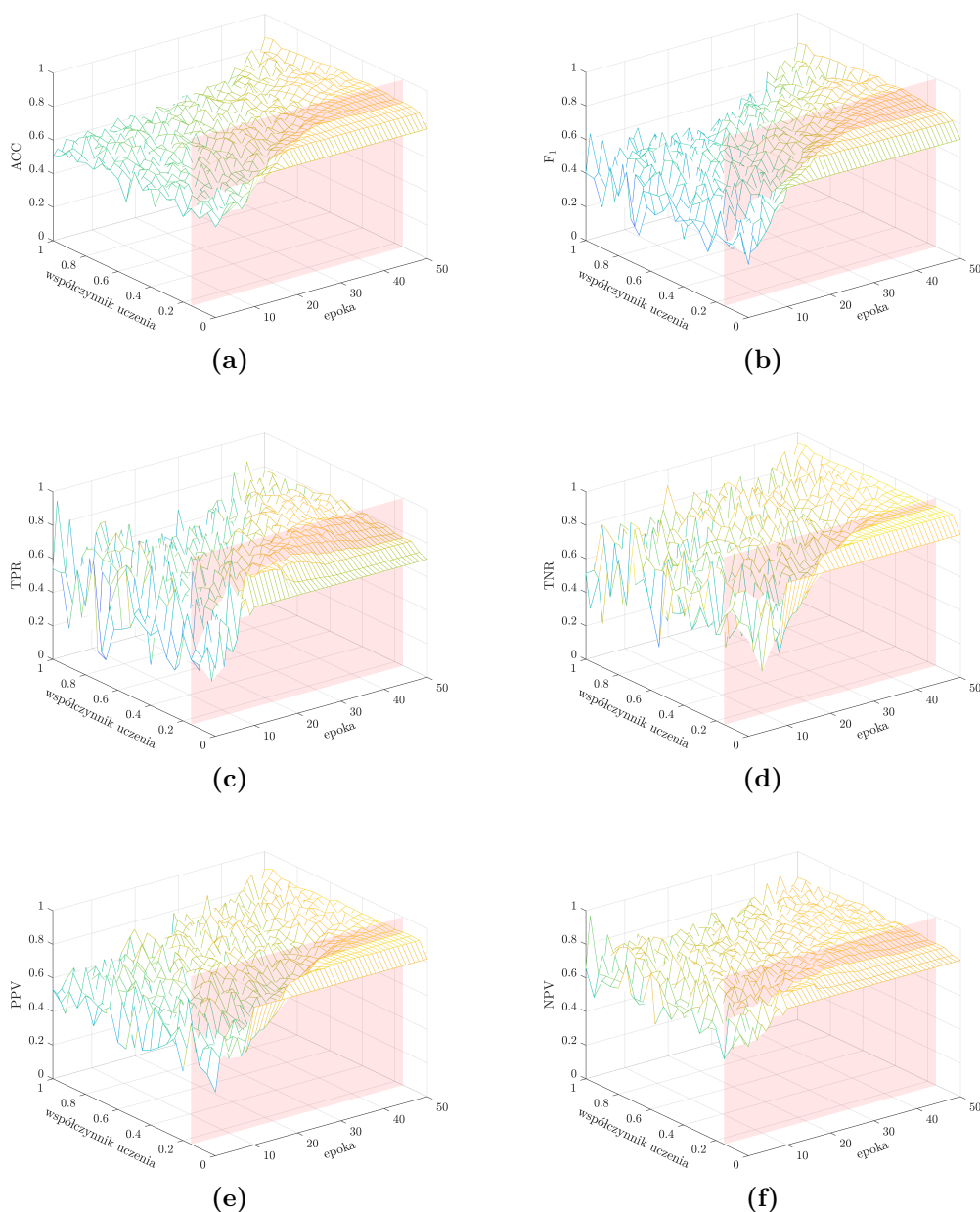
Współczynnik uczenia

Proces treningu związany jest z doбором szeregu hiperparametrów, które bezpośrednio wpływają na finalną efektywność klasyfikatora. Współczynnik spadku tempa uczenia (ang. *learn rate drop factor*) określane również mianem współczynnika uczenia, kontroluje zmianę modelu w odpowiedzi na wartość szacowanego błędu z określoną czułością. Dobór wartości współczynnika uczenia jest istotny, ponieważ zbyt mała jego wartość może wydłużyć proces treningu a w przeciwnym przypadku skutkować zbyt szybkim treningiem, efektem którego może być nieoptymalny zestaw wag dla zbioru danych. Celem

dobrania najkorzystniejszej wartości współczynnika uczenia, przeprowadzony został eksperyment, którego efektem były modele uwzględniające różne wartości współczynnika uczenia z zakresu $[0, 1]$ z krokiem 0,05. Otrzymane wyniki odpowiadające poszczególnym miarom zebrano na rysunku 5.4. Wartości miar sugerują zależność pomiędzy wartością współczynnika uczenia a liczbą epok. W przypadku miar ACC (Rys. 5.4a), F_1 (Rys. 5.4b), TNR (Rys. 5.4d) i PPV (Rys. 5.4e) widoczny jest wpływ optymalnej wartości na stabilizację wyników. Jednocześnie możliwa jest do zaobserwowania zależność, iż niższe wartości współczynnika uczenia w tym przypadku szybciej pozwalają na uzyskanie optymalnego zestawu wag dla zbioru danych. Współczynnik wyników prawdziwie pozytywnych, podobnie jak w przypadku początkowych wymiarów reprezentacji obrazu, sugeruje, iż wydłużony proces treningu może skutkować pogorszeniem wyników (Rys. 5.4c). W szczególności miara TPR , ale również pozostałe, sugerują, iż zbyt niska wartość współczynnika uczenia, nie pozwala w przewidzianym czasie, uzyskać efektywnego zestawu wag. Najmniejszy brak wrażliwości na wartość współczynnika uczenia wykazuje miara NPV , dla której nieznacznie niższe wartości obserwuje się w początkowych epokach procesu treningu, a następnie wyniki są porównywalne. Najmniejszy rozrzut wartości obserwowany jest dla niższych wartości i wydłużonego treningu (Rys. 5.4f). Ostatecznie wartość współczynnika uczenia została oszacowana na podstawie wskazania najwyższych wyników przez poszczególne miary w możliwie najkrótszym procesie nauki. Ocena poszczególnych miar umożliwiła wytypowanie wartości 0,15, która została zaznaczona na wykresie czerwonym kolorem 5.4.

Prawdopodobieństwo regularyzatora typu dropout

Kolejny parametr poddawany ocenie to wartość prawdopodobieństwa p regularyzatora typu *dropout*, który został wykorzystany w autorskiej architekturze klasyfikacji obecności igły biopsyjnej w obrazach ultrasonograficznych. Prawdopodobieństwo p przekłada się na szanse zerowania wag ukrytych jednostek. Zabieg ten stosuje się celem wymuszenia treningu wszystkich jednostek ukrytych, co pozwala na zmniejszenie prawdopodobieństwa nadmiernego dopasowania modelu do zbioru danych treningowych. W celu wskazania optymalnej wartości prawdopodobieństwa p we wdrożonych w architekturze regularyzatorach typu *dropout*, zrealizowany został eksperyment, w trakcie którego wygenerowane zostały modele cechujące się różną wartością p . Testowano zakres $[0, 1]$ z krokiem 0,05. Wytrenowane modele poddane zostały ocenie za pomocą miar efektywności z uwzględnieniem liczby epok treningu (Rys. 5.5). Wartości miar ACC (Rys. 5.5a), F_1 (Rys. 5.5b) i TPR (Rys. 5.5c) sugerują znaczący wpływ regularyzatora typu *dropout* na wyniki. Im większa wartość prawdopodobieństwa p , tym wartości miar niższe. Otrzymane wyniki sugerują również znaczący wpływ liczby epok, ponieważ przy zbyt dużej ich liczbie przy wyższych warto-



Rys. 5.4: Podsumowanie wpływu zmiany wartości współczynnika tempa uczenia i liczby epok na efektywność autorskiego klasyfikatora obecności igły w obrazie. Na rysunkach przedstawiono kolejno wartości miar: 5.4a dokładność, 5.4b miara F_1 , 5.4c współczynnik wyników prawdziwie pozytywnych, 5.4d współczynnik wyników prawdziwie negatywnych, 5.4e wartość predykcyjna dodatniego wyniku i 5.4f wartość predykcyjna ujemnego wyniku.

ściach p miary są znacząco gorsze. Wartości prawdopodobieństwa p korzystnie wpływają na ocenę miarami TNR i PPV , co zostało przedstawione odpowiednio na wykresach 5.5d i 5.5e. Prawdopodobieństwo p z zakresu $[0, 1; 0, 6]$ i wydłużonego procesu treningu (powyżej 20 epok) wskazuje najwyższe wyniki. W odróżnieniu od współczynnika uczenia, miara NPV , ze względu na rozważanie elementów ocenionych prawdziwie i fałszywie jako negatywne, wykazuje wysoką wrażliwość na wyższe wartości prawdopodobieństwa p i wydłużony proces treningu (Rys. 5.5f). Finalna wartość prawdopodobieństwa p została wybrana na podstawie oceny poszczególnych miar i wytypowania wartości, która we wszystkich przypadkach pozwala uzyskać najwyższe wartości przy możliwie najkrótszym procesie treningu. Wytypowana wartość prawdopodobieństwa p dla regularyzatora typu *dropout* dla analizowanego zbioru danych wynosi 0,1.

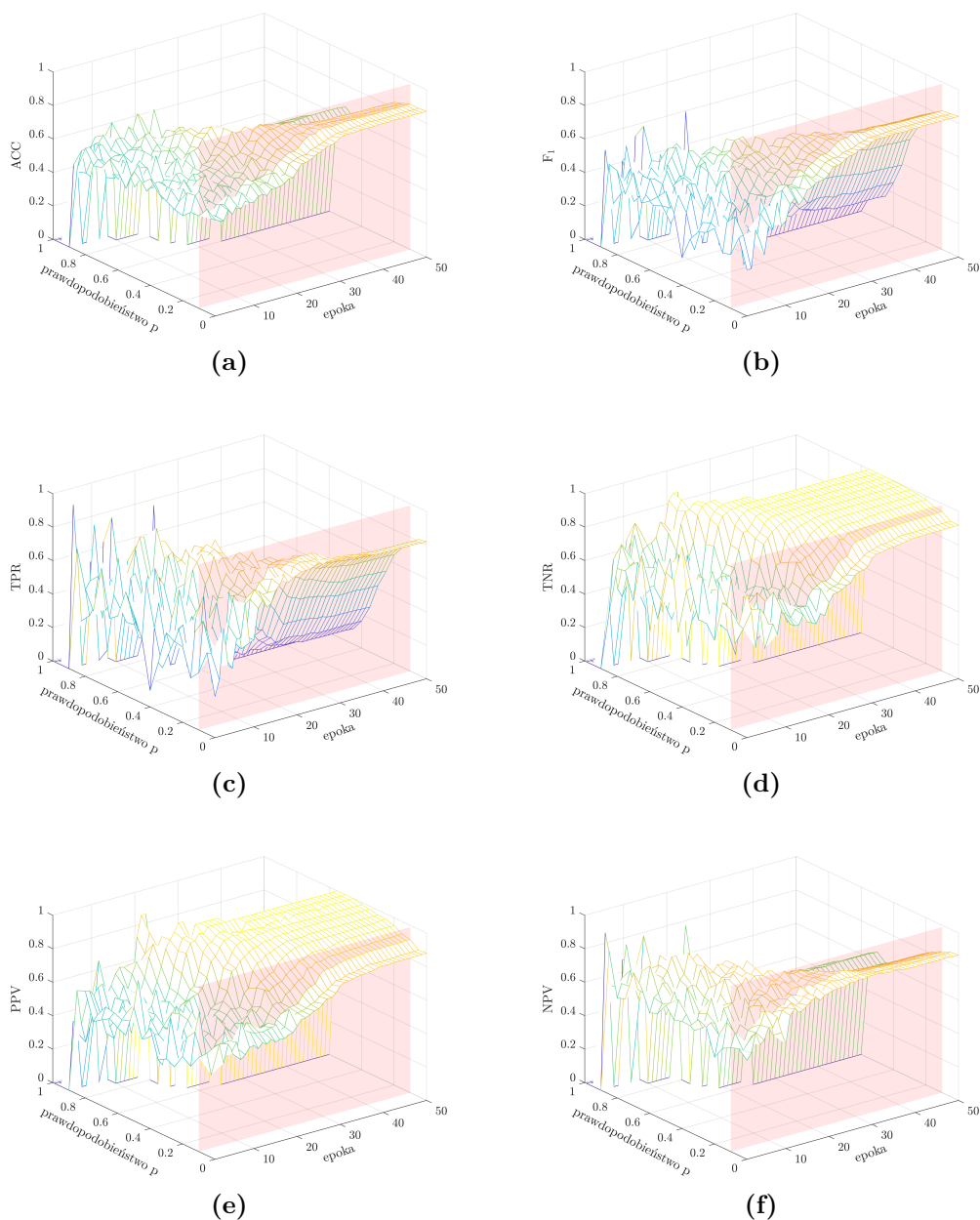
Dobór wagi funkcji straty

Zabieg ważenia funkcji straty stanowi narzędzie stosowane w przypadku niezbalansowanego zbioru danych, kiedy liczebność jednej z klas przewyższa liczebność drugiej. Im zbiór danych jest mniejszy tym wpływ braku równowagi powoduje większe promowanie liczniejszego zbioru. Przykładem eliminacji wpływu braku równowagi zbioru na efektywność klasyfikatora, jest zastosowanie ważonej entropii krzyżowej jako funkcji straty. Podejście to pozwala na promowanie wyższą wagą klasy mniej licznej. Wpływ ważenia klasy igły i braku igły został poddany weryfikacji z punktu odniesienia klasy igły oraz klasy braku igły. Uzyskane wyniki zebrano i przedstawiono w formie trójwymiarowych wykresów zebranych na rysunkach 5.6 i 5.7. W wierszu górnym każdego z rysunków przedstawione są wyniki miar dla klasy igła, w dolnym dla klasy brak igły.

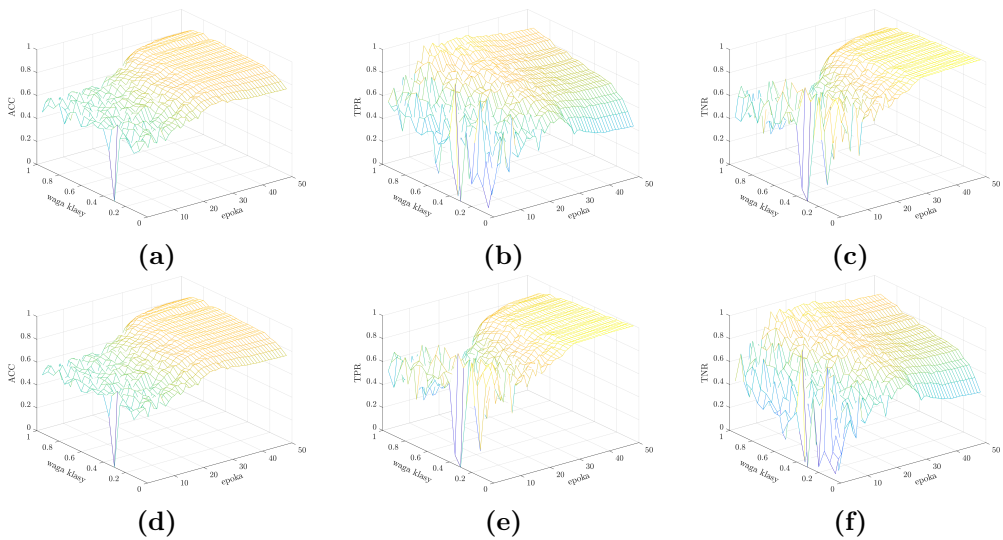
Stosunek liczebności obrazów w klasach wykorzystanych w poszczególnych eksperymentach wykazuje zależność odpowiadającą 0,4 dla klasy igły i 0,6 dla klasy braku igły w poszczególnych zestawach danych wyznaczonych metodą k -krotnej walidacji krzyżowej. W przypadku każdej z miar możliwe jest wnioskowanie, iż liczba około 25 epok jest minimalnym czasem treningu w rozważanym przypadku klasyfikacji igły. Na podstawie analizy wyników i liczebności klas w poszczególnych eksperymentach oszacowano wagę 0,6 dla klasy igła oraz 0,4 dla klasy brak igły.

Klasyfikator trenowany oryginalnym zbiorem danych

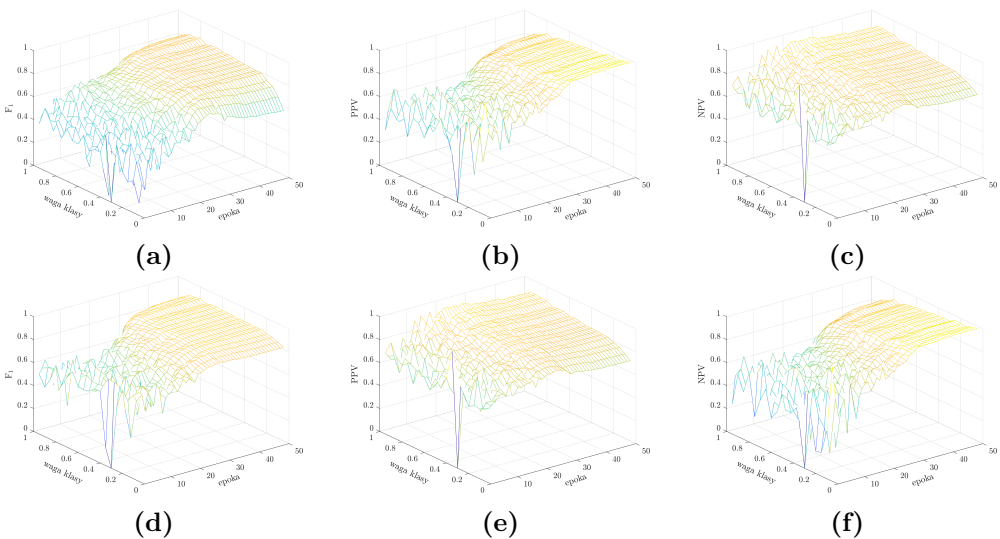
Dokładność proponowanej metody jest uzależniona od wymiarów obrazu wejściowego i liczby epok przypadającej na proces treningu. Zależność tę przedstawiono na wykresach zebranych na rysunku 5.8. W przypadku każdej z ocenianych miar (ACC , F_1 , TPR , TNR , PPV , NPV) zaobserwować można po-



Rys. 5.5: Podsumowanie wpływu zmiany wartości prawdopodobieństwa p regularyzatora typu *dropout* na efektywność autorskiego klasyfikatora obecności igły w obrazie. Na rysunkach przedstawiono kolejno wartości miar: 5.4a dokładność, 5.4b miara F_1 , 5.4c współczynnik wyników prawdziwie pozytywnych, 5.4d współczynnik wyników prawdziwie negatywnych, 5.4e wartość predykcyjna dodatniego wyniku i 5.4f wartość predykcyjna ujemnego wyniku.



Rys. 5.6: Podsumowanie wpływu ważenia klas podczas procesu treningu na efektywność autorskiego klasyfikatora obecności igły w obrazie. Na rysunkach przedstawiono kolejno wartości miar: 5.4a dokładność, 5.4b miara F_1 , 5.4c współczynnik wyników prawdziwie pozytywnych, 5.4d współczynnik wyników prawdziwie negatywnych, 5.4e wartość predykcyjna dodatniego wyniku i 5.4f wartość predykcyjna ujemnego wyniku.



Rys. 5.7: Podsumowanie wpływu ważenia klas podczas procesu treningu na efektywność autorskiego klasyfikatora obecności igły w obrazie. Na rysunkach przedstawiono kolejno wartości miar: 5.4a dokładność, 5.4b miara F_1 , 5.4c współczynnik wyników prawdziwie pozytywnych, 5.4d współczynnik wyników prawdziwie negatywnych, 5.4e wartość predykcyjna dodatniego wyniku i 5.4f wartość predykcyjna ujemnego wyniku.

prawę wyników wraz ze wzrostem zarówno rozdzielczości obrazów wejściowych jak i liczby epok.

Średnia wartość dokładności ACC w 5. epoce treningu dla modelu trenowanego obrazami o rozmiarach 16×16 przyjmuje wynik 0,479, dla rozdzielczości $128 \times 128 - 0,546$, natomiast dla $256 \times 256 - 0,565$. Dla epoki 15., odpowiednio dla każdej z rozdzielczości, średnia wartość ACC wynosi 0,677, 0,685 i 0,609, natomiast dla 25. dokładność to odpowiednio 0,715, 0,756 i 0,775 (Rys. 5.8a). Wartości dokładności przy niższych rozdzielczościach obrazu wzrastają szybciej i przy mniejszej liczbie epok osiągają wartość maksymalną. Modele trenowane większymi obrazami wymagają trenowania większą liczbą epok, ale finalnie cechują się wyższymi wartościami miary dokładności. Zależność tę potwierdzają obserwację pozostałych miar stosowanych do oceny efektywności klasyfikatora.

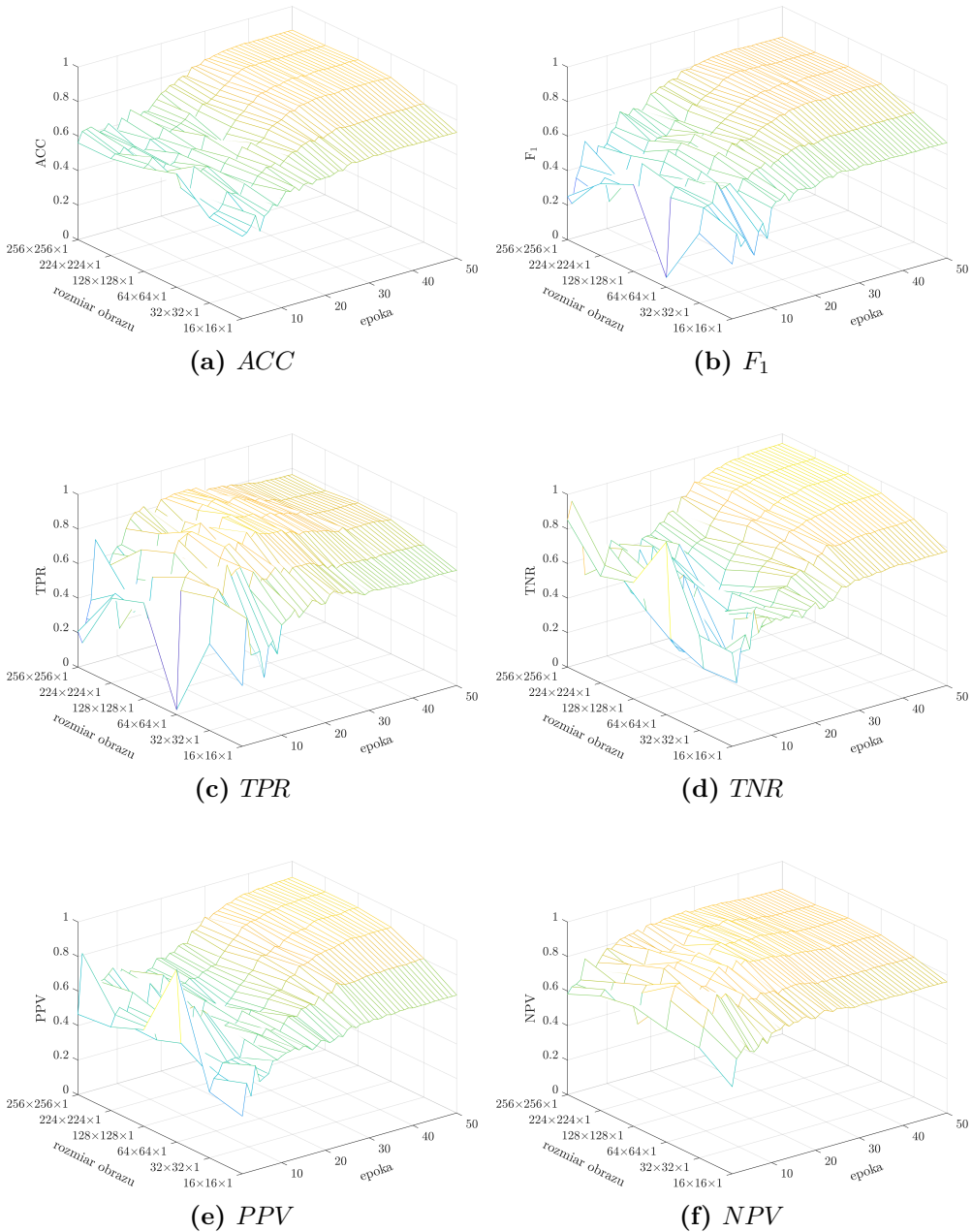
W przypadku miary F_1 obrazu o wyższej rozdzielczości cechują się wyższymi wynikami wraz ze wzrostem liczby epok zastosowanych podczas treningu. Dla modeli trenowanych obrazami o wymiarach 16×16 i 32×32 stabilizacja wyników następuje około 20. epoki, a dla pozostałych wymiarów około 30. (Rys. 5.8b). Wartości te sugerują, iż wraz ze wzrostem rozmiarów analizowanych map cech wzrasta ich złożoność. Ocena przestrzennego rozłożenia średnich wartości miary F_1 umożliwia również oszacowanie wartości granicznej liczby epok stanowiąc mechanizm zmniejszający prawdopodobieństwo przeczenia trenowanych modeli.

Miary współczynników TPR i TNR wykazują, iż zbyt krótki trening – niezależnie od rozdzielczości obrazu – prowadzi do zbyt wielu fałszywie dodatnich wyników (Rys. 5.8c i Rys. 5.8d). TPR przyjmuje najwyższe średnie wartości dla liczby epok z przedziału między 20 a 30. W kontekście detekcji igły biopsyjnej, ze względu na dalsze etapy przetwarzania, ważniejsze jest ograniczenie wyników fałszywie dodatnich nawet kosztem wzrostu wartości fałszywie ujemnych.

Wartości miary PPV oraz NPV również wzrastają wraz ze wzrostem rozdzielczości obrazu i liczby epok (Rys. 5.8e i Rys. 5.8f). Stabilizacja wartości następuje po około 20 – 30 epokach treningu. W przeciwieństwie do pozostałych metryk, NPV wykazuje małą wrażliwość na rozdzielczość obrazów wykorzystywanych w treningu.

Klasyfikator trenowany techniką dostrajania

Ze względu na małą bazę obrazów, przeprowadzono trening modeli o zaproponowanej architekturze sieci klasyfikującej igłę biopsyjną z wykorzystaniem techniki dostrajania (ang. *fine-tuning*) [147]. W tym podejściu trening jest dwuetapowy, zgodnie z opisem przedstawionym w podrozdziale 4.4.4 i na schemacie 4.11. Podobnie do trenowania opisanego wyżej, także w modelach wytrenowanych techniką dostrajania dokonano oceny wpływu rozdzielczości obrazu wejściowego oraz liczby epok treningu. Wyniki zostały zaprezentowane



Rys. 5.8: Wpływ wymiarów obrazu wejściowego i liczby epok na efektywność autorskiego klasyfikatora obecności igły w obrazie, trenowanego oryginalnym zbiorem danych.

na rysunku 5.9.

Podobnie do poprzednich wyników, również tym razem widoczny jest wzrost wartości metryk wraz ze zwiększeniem rozdzielczości obrazów i wydłużeniem treningu. W przypadku każdej miary występuje mniejszy rozrzut wartości w początkowych epokach procesu treningu w porównaniu do poprzedniego podejścia. Miara dokładności przyjmuje wartości 0,674, 0,607 i 0,574 dla rozdzielczości odpowiednio 16×16 , 128×128 i 256×256 w epoce 5. Wartości te w epoce 15. wzrastają odpowiednio do 0,737, 0,780 i 0,795. W przypadku epoki 25., średnie wartości dokładności oszacowano na poziomie odpowiednio 0,765, 0,876 i 0,876, co pozwala na wyciągnięcie wniosków, że dla niższych rozdzielczości wartości te początkowo wzrastają szybciej, lecz osiągają niższe wartości maksymalne. Średnie wartości dokładności i pozostałych miar oceny są wyższe po zastosowaniu treningu metodą *fine-tuning*. Z analizy rozkładu miar ACC , F_1 i szczególnie TPR wynika, że wskazane jest zmniejszenie liczby epok w procesie treningu. Maksymalne wartości wskaźników osiąmane są przy mniejszej liczbie epok.

Mechanizm wcześniejszego zatrzymania, oryginalne obrazy

Zgodnie z opisem przedstawionym w podrozdziale 4.4.4, przeprowadzono pomiary działania sieci z zastosowaniem mechanizmu wcześniejszego zatrzymania (ang. *early stopping*) [200].

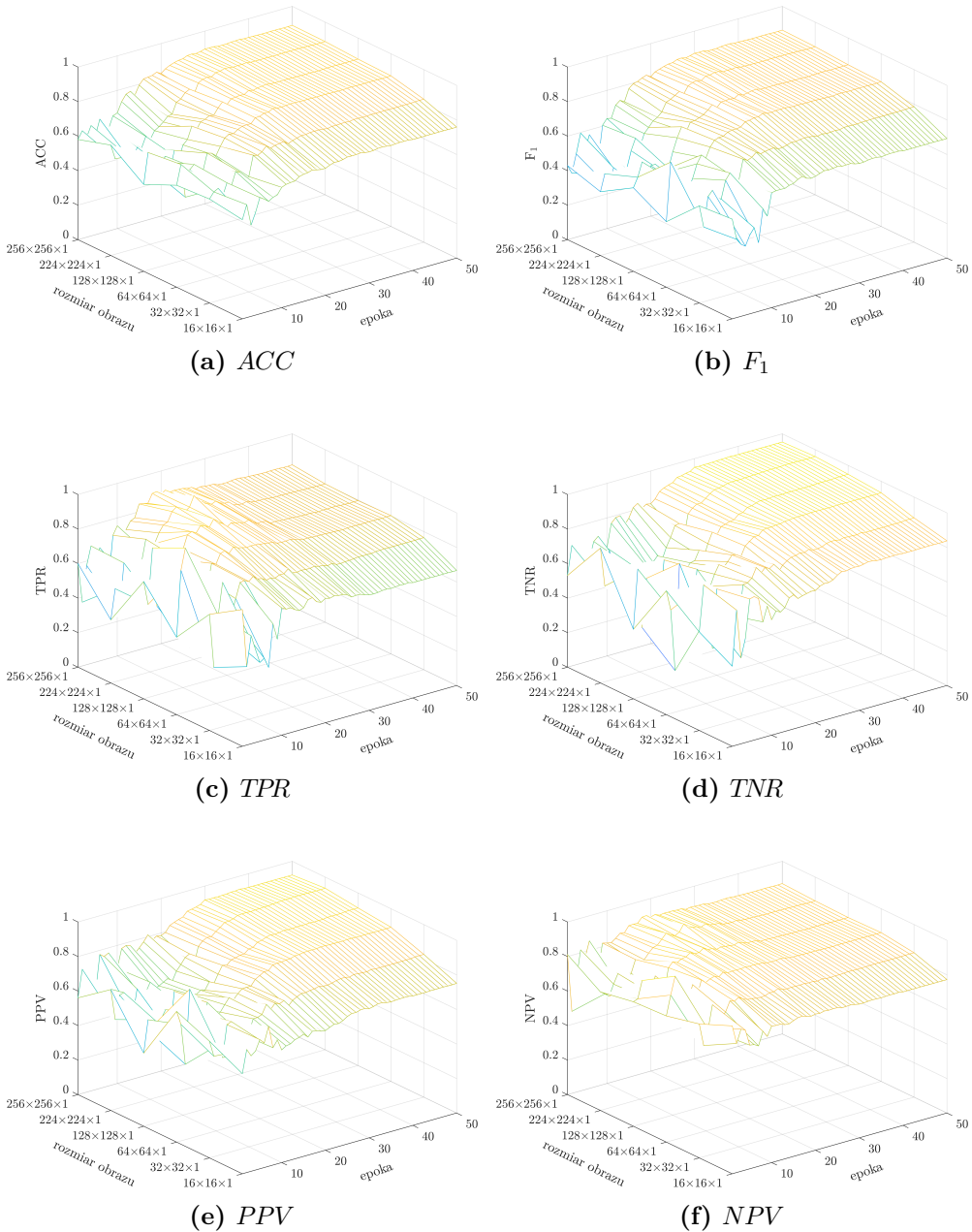
Mediana liczby epok treningu w tych eksperymentach przyjmowała wartości 13, 19, 19, 24, 25 i 26 dla rosnących wymiarów obrazów przyjmowanych przez warstwę wejściową. Im większa reprezentacja obrazu, tym większa liczba ekstrahowanych cech, dla których trenowane są wartości wag, dlatego też liczba niezbędnych epok rosła.

W treningu z mechanizmem wcześniejszego zatrzymania na oryginalnych danych obrazowych mediana wartości ACC mieściła się – dla różnych wielkości obrazów wejściowych – w zakresie od $0,694 \pm 0,142$ do $0,793 \pm 0,114$ (Rys. 5.10a).

Ocena efektywności miarą F_1 wykazała najniższą wartość mediany klasyfikacji wynoszącą $0,623 \pm 0,231$ w przypadku najmniejszych obrazów. Najwyższą medianę otrzymano w eksperymentach na obrazach 128×128 i 224×224 , dla których otrzymano mediany wyników odpowiednio $0,756 \pm 0,132$ i $0,733 \pm 0,086$ (Rys. 5.10b).

Współczynnik wyników prawdziwie dodatnich TPR przyjmuje najniższą medianę z wartości wyników klasyfikacji dla obrazów o najmniejszych wymiarach: 16×16 ($0,592 \pm 0,286$) i 32×32 ($0,787 \pm 0,175$). Najwyższą wartości mediany w przypadku miary TPR zaobserwowano dla wariantów 128×128 oraz 256×256 , które wynosiły odpowiednio $0,914 \pm 0,161$ i $0,914 \pm 0,215$ (Rys. 5.10c).

Wartości miar TPR , TNR , PPV i NPV przedstawiono w kolejnych wykresach rysunku 5.10. Dla wszystkich wyników dostrzegalna jest tendencja wzrostu



Rys. 5.9: Wpływ wymiarów obrazu wejściowego i liczby epok na efektywność autorского klasyfikatora obecności igły w obrazie, trenowanego metodą dostrajania.

wartości wraz ze zwiększaniem się rozdzielczości obrazów wejściowych. Zależność ta nie jest bardzo mocna i w większości przypadków wyniki przetwarzania

najmniejszych i największych rozdzielczości nie są znamienne statystycznie, jednak można wyciągnąć wniosek, że zwiększenie wymiarów obrazów wejściowych globalnie prowadzi do poprawy jakości działania sieci.

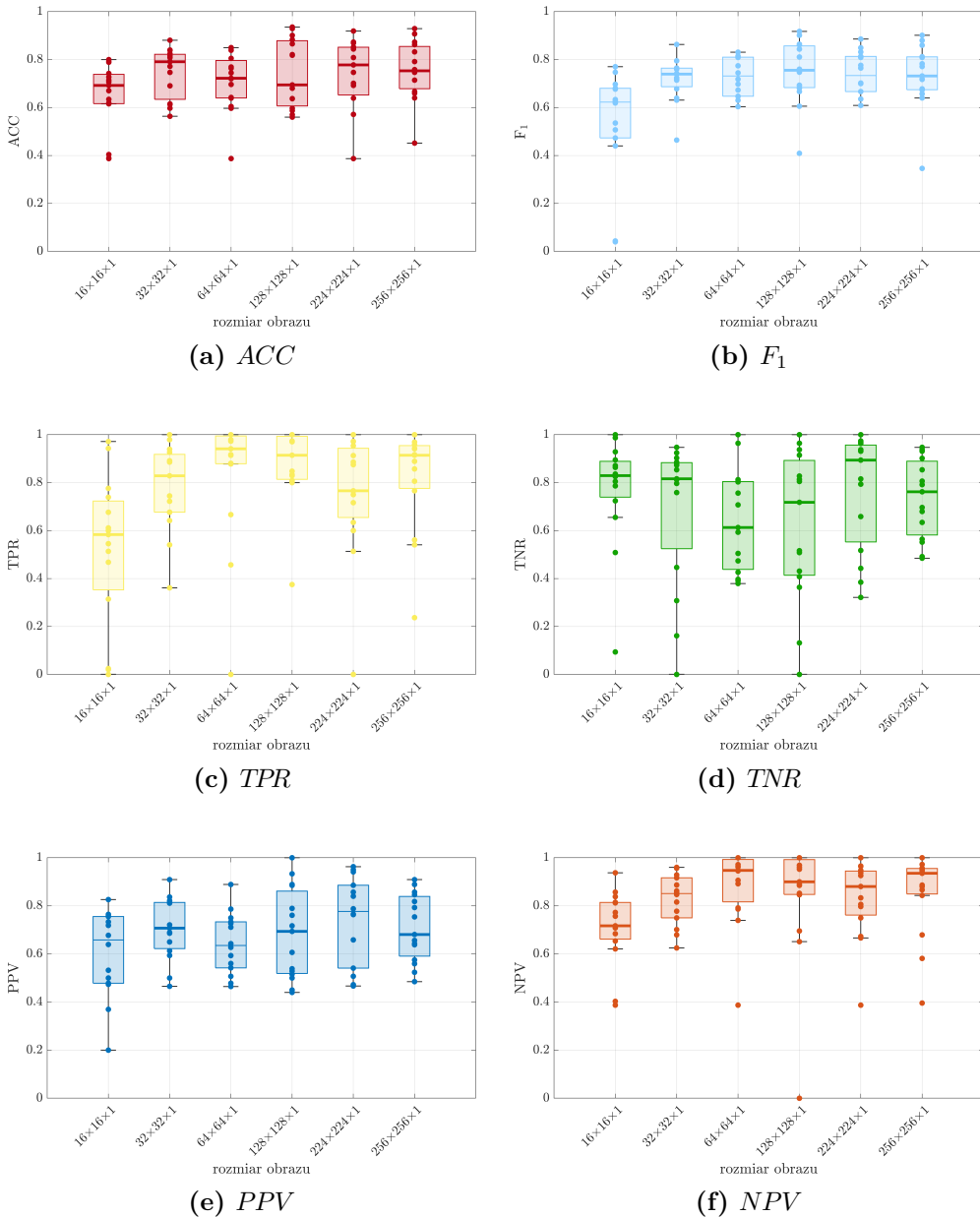
Test Wilcozona wykazał brak istotnych statystycznie różnic pomiędzy wartościami poszczególnych miar wyznaczonych dla różnych wymiarów reprezentacji obrazu ($p > 0,05$) z wyłączeniem porównania obrazu wejściowego o wymiarach $16 \times 16 \times 1$ z pozostałymi testowanymi ($p < 0,05$).

Mechanizm wcześniejszego zatrzymania, trening metodą dostrajania

Autorska architektura sieci klasyfikującej igłę biopsyjną w obrazach ultrasonograficznych została również poddana treningowi techniką dostrajania (ang. *fine-tuning*)[147] ze względu na małą bazę danych obrazowych zebraną w tabeli 5.6. Otrzymane wartości miar dla treningu metodą dostrajania z uwzględnieniem mechanizmu wcześniejszego zatrzymania przedstawiono za pomocą wykresów pudełkowych zebranych na rysunku 5.11. Mediany liczby epok w treningach realizowanych dla poszczególnych wymiarów obrazów wynosiły wraz ze wzrostem rozmiaru odpowiednio 14, 14, 14, 22, 27 i 28. Podobnie jak w przypadku treningu opartego jedynie o zbiór danych oryginalnych, widoczne jest zwiększenie liczby epok wraz ze zwiększeniem wymiarów obrazów przekazywanych do warstwy wejściowej.

Mediana wartości miary dokładności ACC w przypadku treningu metodą dostrajania z uwzględnieniem mechanizmu wcześniejszego zatrzymania, wzrasta wraz ze zwiększeniem wymiarów obrazów przyjmując wartości z zakresu $0,734 \pm 0,078 - 0,883 \pm 0,072$ (Rys. 5.11a). Podobnie w przypadku miary F_1 , sugerowanej w ocenie metod opartych o ręcznie etykietowane dane, najniższe mediany odnotowano w przypadku zastosowania obrazów o najmniejszych wymiarach a najwyższe dla największych. W pierwszym przypadku dla wymiarów $16 \times 16 \times 1$ i $32 \times 32 \times 1$ mediana wartości miary F_1 wynosiła odpowiednio $0,696 \pm 0,234$ i $0,745 \pm 0,178$, w drugim natomiast dla $224 \times 224 \times 1$ i $256 \times 256 \times 1$ odnotowano odpowiednio $0,850 \pm 0,074$ i $0,840 \pm 0,073$ (Rys. 5.11b).

Współczynnik wyników prawdziwie dodatnich TPR przyjmuje najniższe wartości mediany w przypadku klasyfikatorów opartych o najniższe rozdzielczości obrazów (Rys. 5.11c), podobnie jak w przypadku jednoetapowego treningu danymi oryginalnymi (Rys. 5.10c). Odnotowane mediany dla miary TPR mieszczą się w zakresie $0,598 \pm 0,293 - 0,912 \pm 0,241$. Zakres wartości median w przypadku współczynnika wyników prawdziwie dodatnich TNR wynosił od $0,862 \pm 0,093$ do $0,948 \pm 0,092$ (Rys. 5.11d). Zarówno w przypadku miary TPR jak i TNR obserwowana jest zależność pomiędzy wymiarowością obrazu a wartością mediany. Pozwala to wnioskować o korzystnym wpływie obrazów o wyższej rozdzielczości na wyniki klasyfikacji proponowaną architekturą. Wykresy pudełkowe wyznaczone dla miar wartości predykcyjnej dodatniego wyniku oraz



Rys. 5.10: Wykresy pudełkowe efektywności autorskiej metody klasyfikacji obecności igły biopsyjnej w obrazie ultrasonograficznym wyznaczone dla modeli trenowanych z uwzględnieniem mechanizmu wcześniejszego zatrzymania oryginalną bazą obrazów.

wartości predykcyjnej ujemnego wyniku przedstawione odpowiednio na rysunkach 5.11e i 5.11f również wskazują, jak w przypadku pozostałych miar, po-

prawę wyników klasyfikacji metodą uwzględniającą technikę dostrajania. Obserwuje się również, podobnie do pozostałych miar, korzystną relację pomiędzy zwiększającą się rozdzielczością obrazu a otrzymywanymi wynikami.

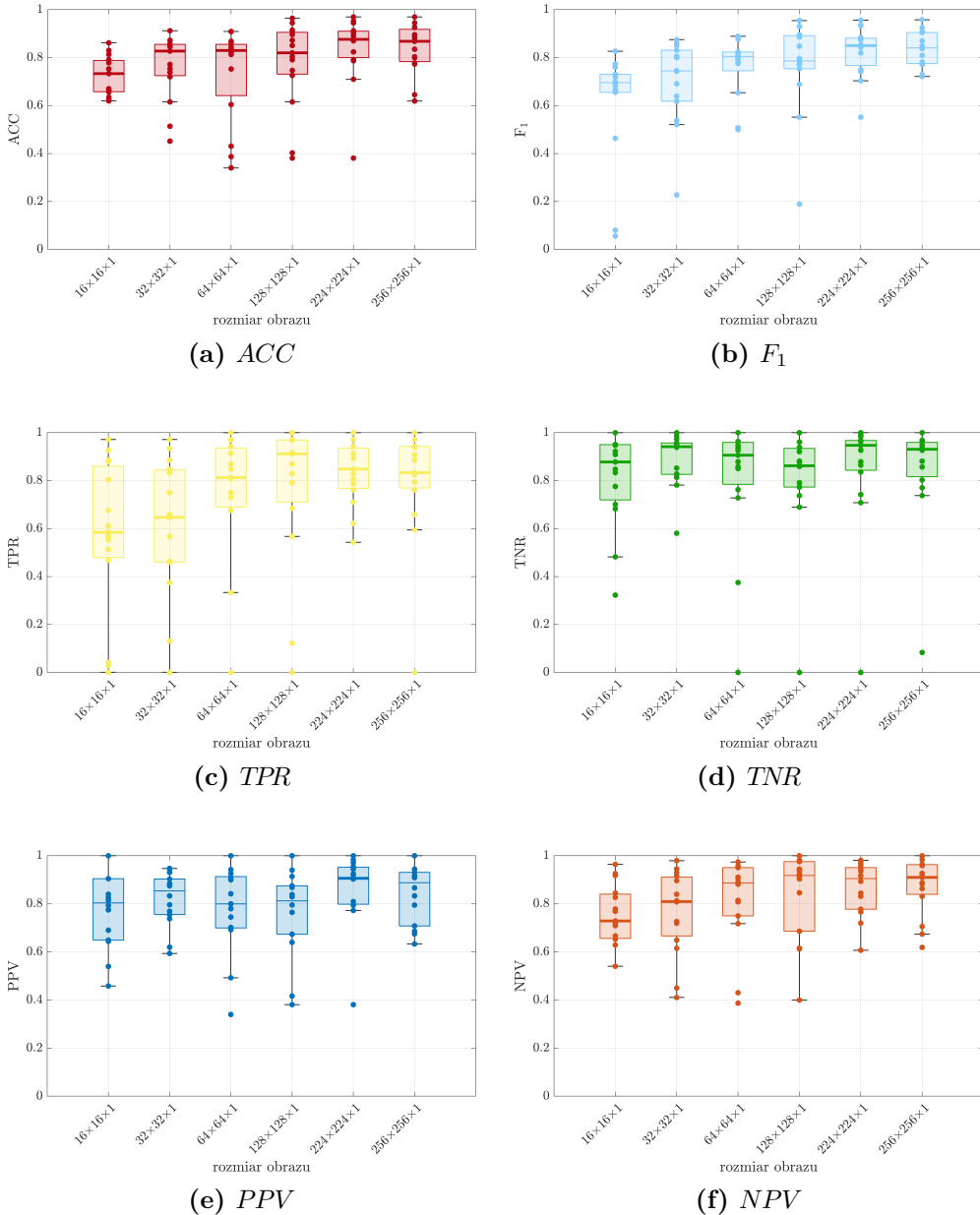
Ocena obliczonych wartości median dla poszczególnych miar pozwala wnioskować, że zastosowanie procesu treningu z uwzględnieniem techniki dostrajania [147] w przypadku małych baz danych korzystnie wpływa na efektywność klasyfikacji opisywaną architekturą. Dla każdej z miar odnotowano wzrost wartości median dla modeli trenowanych metodą dostrajania względem modeli trenowanych jednym etapem uwzględniającym tylko oryginalną bazę obrazów. W przypadku miary F_1 , która sugerowana jest w ocenie klasyfikatorów trenowanych danymi przygotowywanych przez człowieka, odnotowano średni wzrost wyników o 7,1% (mniejszy dla obrazów o niższej rozdzielczości, większy dla o większej). Skuteczność metody predykcji uzależniona jest również od wymiarów obrazu wejściowego. Treść obrazów o wyższej rozdzielczości częściej jest poprawnie klasyfikowane w kontekście małych obiektów, do których należy igła biopsyjna stanowiąca średnio 0,81% wszystkich pikseli obrazu. Obserwacje korzystnego wpływu treningu metodą dostrajania w przypadku obrazów o wymiarach $224 \times 224 \times 1$ i $256 \times 256 \times 1$ potwierdza statystyczny test Wilcozona w przypadku metryk ACC , F_1 , TNR i PPV ($p < 0,05$), co pozwala na wnioskowanie o zasadności stosowania dwuetapowego procesu uczenia w przypadku danych ultrasonograficznych.

Czas klasyfikacji jednego obrazu zarówno w przypadku modelu trenowanego danymi oryginalnymi, jak również techniką dostrajania przyjmuje najniższą średnią wartość dla obrazów o najmniejszych wymiarach: $16 \times 16 \times 1$ i wynosi 2,6 ms. Wraz ze wzrostem rozdzielczości obrazu wejściowego do trenowanych klasyfikatorów, nieznacznie rośnie średnia wartość czasu przetwarzania, niemniej jednak wartości te są porównywalne (Rys. 5.12a i 5.12b). Porównanie czasów przetwarzania jednego obrazu dla klasyfikatorów trenowanych jedno- i dwuetapowo za pomocą testu Wilcozona wskazuje istotne statystycznie różnice pomiędzy różnymi wymiarami reprezentacji obrazu ($p < 0,05$).

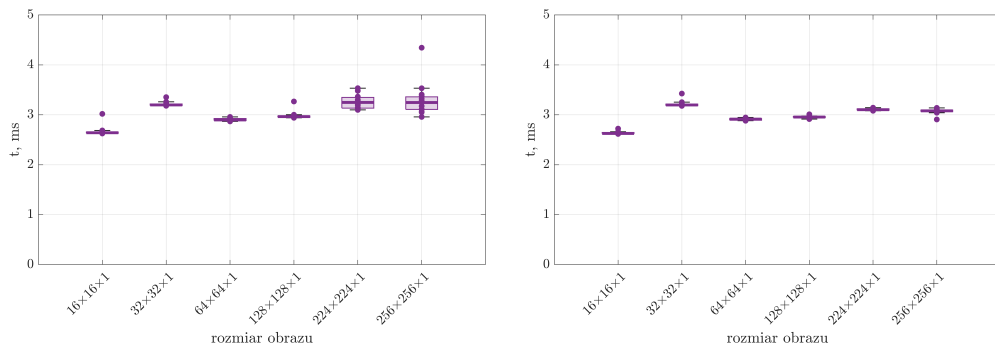
Finalny dobór hiperparametrów i porównanie z istniejącymi sieciami

Na podstawie analizy otrzymanych wyników, jako finalny model predykcji wytypowano model z początkową reprezentacją obrazu o wymiarach $256 \times 256 \times 1$ trenowany techniką dostrajania. Wyboru dokonano na podstawie oceny miar oceny efektywności klasyfikatora, gdzie dla każdej z nich przyjmował jedną z najwyższych średnich wartości.

Wybraną metodę porównano z publicznymi, opisanymi w literaturze, wstępnie trenowanymi architekturami: AlexNet [83], VGG16 [151], VGG19 [151], SqueezeNet [70], GoogLeNet [162], ResNet101 [56] i InceptionResNetV2 [161]. Każda z wymienionych sieci cechuje się inną liczbą warstw ukrytych, paramet-



Rys. 5.11: Wykresy pudełkowe efektywności autorskiej metody klasyfikacji obecności igły biopsyjnej w obrazie ultrasonograficznym wyznaczone dla modeli trenowanych z uwzględnieniem mechanizmu wcześniejszego zatrzymania oraz z wykorzystaniem techniki dostrajania.



(a) Trening z mechanizmem wcześniejszego zatrzymania na oryginalnym zbiorze danych.

(b) Trening z mechanizmem wcześniejszego zatrzymania z wykorzystaniem techniki dostrajania.

Rys. 5.12: Rozkład czasów klasyfikacji obrazów USG o różnej wielkości za pomocą autorskiej metody klasyfikacji obecności igły biopsyjnej.

trów oraz organizacją połączeń (Tab. 4.1). Różnica również występuje w rozmiarze obrazu podawanego na wejściu warstwie wejściowej. Wstępny trening każdej z tych sieci został wykonany za pomocą bazy obrazów ImageNet[139]. Następnie każda z architektur została dostrojona za pomocą bazy obrazów ultrasonograficznych uwzględniających obecność igły oraz jej brak (Tab. 5.6). Ocena efektywności testowanych rozwiązań została przeprowadzona za pomocą miar ACC , F_1 , TPR , TNR , PPV , NPV oraz średniego czasu t klasyfikacji jednego obrazu.

Ocena efektywności autorskiej metody klasyfikacji obecności igły biopsyjnej w obrazach ultrasonograficznych oraz wstępnie trenowanych architektur po dostrojeniu wykazała najniższą wartość mediany dla miary dokładności ACC w przypadku użycia sieci typu VGG16 oraz AlexNet. Odnotowane dla nich dokładności wynosiły odpowiednio $0,706 \pm 0,094$ oraz $0,721 \pm 0,087$ (Rys. 5.13a). Podobne zależności pomiędzy wynikami uzyskanymi dla poszczególnych typów klasyfikatorów są widoczne w przypadku miary F_1 (Rys. 5.13b). Najniższe wartości median obserwowane są w przypadku architektur AlexNet ($0,629 \pm 0,184$) i SqueezeNet ($0,692 \pm 0,160$). Najkorzystniejsze wyniki wykazują autorski model klasyfikacji ($0,840 \pm 0,073$) oraz ResNet101 ($0,828 \pm 0,088$).

Współczynnik wyników prawdziwie dodatnich TPR przyjmuje najniższe wartości median w przypadku klasyfikatorów opartych o AlexNet i GoogLeNet z wynikami na poziomie $0,553 \pm 0,244$ oraz $0,686 \pm 0,242$. Najwyższe wartości mediany w przypadku miary TPR odnotowano natomiast za pomocą modeli InceptionResNetV2 i ResNet101 z wynikami odpowiednio $0,972 \pm 0,107$ i $0,903 \pm 0,088$ (Rys. 5.13c). Najniższe wartości mediany otrzymane dla współczynnika wyników prawdziwie dodatnich TNR obserwowane są w przypadku

architektur SqueezeNet ($0,714 \pm 0,269$) i InceptionResNetV2 ($0,758 \pm 0,359$). Najkorzystniejsze wartości miary *TNR* cechują architektury autorską oraz GoogLeNet, dla których otrzymane wartości odpowiednio wynoszą $0,929 \pm 0,232$ i $0,879 \pm 0,110$ (Rys. 5.13d). Ocena wartością predykcijną dodatniego *PPV* wykazała najniższe rezultaty w przypadku architektur SqueezeNet i InceptionResNetV2 z medianami wynoszącymi $0,623 \pm 0,205$ i $0,677 \pm 0,183$. Najwyższe rezultaty zostały osiągnięte za pomocą autorskiej metody oraz architektury ResNet101, gdzie wyznaczone wartości mediany były równe $0,888 \pm 0,119$ i $0,795 \pm 0,158$ (Rys. 5.13e).

Wartość predykcjna ujemnego wyniku jest najniższa w przypadku modeli AlexNet ($0,709 \pm 0,119$) i GoogLeNet ($0,793 \pm 0,151$). Najwyższe wyniki miary *NPV* obserwuje się w przypadkach architektur InceptionResNetV2 i autorskiej metody klasyfikacji obecności igły biopsyjnej, dla których odnotowano odpowiednio wartości $0,982 \pm 0,087$ i $0,917 \pm 0,100$ (Rys. 5.13f).

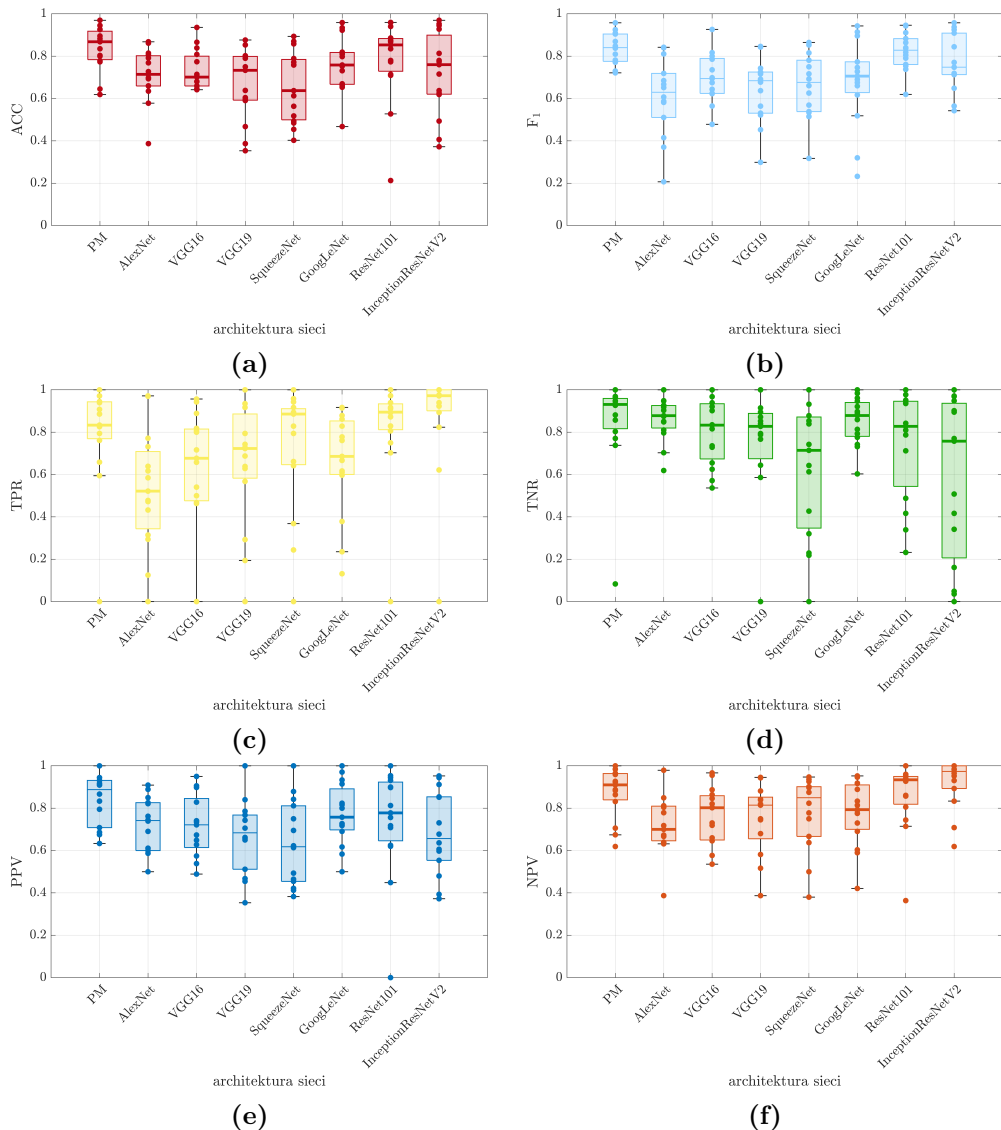
Czas *t* klasyfikacji jednego obrazu ultrasonograficznego jest najniższy dla autorskiej metody klasyfikacji i wynosi 3,09 ms. Wraz ze wzrostem liczby warstw ukrytych w architekturze rośnie czas przetwarzania jednego przypadku, co zostało przedstawione na rysunku 5.14. Najwyższe czasy *t* obserwowane są dla najbardziej złożonych modeli, jakimi są InceptionResNetV2 (52,70 ms) i ResNet101 (20,37 ms). Różnice te sięgają rzędu wielkości w porównaniu do autorskiej metody i AlexNet, które cechują się najmniejszą liczbą warstw.

5.3.2 Klasyfikacja obrazów USG ze zmianą ogniskową

Baza obrazów zawierająca 2685 dwuwymiarowych obrazów ultrasonograficznych (Tab. 5.8) została podzielona losowo na trzy zbiory: treningowy, walidacyjny i testowy na poziomie unikatowych identyfikatorów pacjentów z wykorzystaniem mechanizmu *k*-krotnej walidacji krzyżowej zgodnie z opisem przedstawionym w podrozdziale 4.5.2. Podejście to wdrożono celem eliminacji wycieku informacji, która mogłaby wpłynąć na jakościową ocenę działania modelu poprzez nadmierne dopasowanie do zbioru testowego. Proces uczenia modelu wykorzystującego autorską architekturę klasyfikującą został przeprowadzony za pomocą zbiorów testowego i walidacyjnego, ewaluacja klasyfikacji została natomiast oparta o zbiór testowy. Eksperyment wykonano 15-krotnie w oparciu o 15 losowych podziałów bazy danych na trzy części.

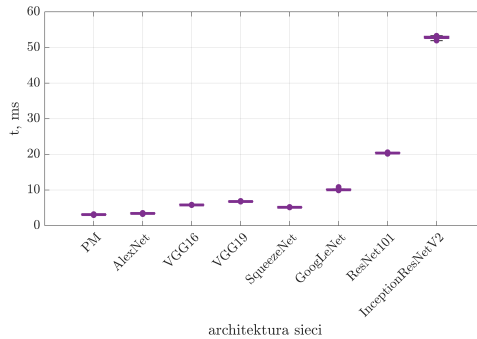
Implementacja architektury oraz proces uczenia zostały przygotowane z wykorzystaniem środowiska MATLAB 20221a zainstalowanego na komputerze PC z systemem operacyjnym Windows 10 wyposażonym w procesor AMD Ryzen 9 3900 (CPU), pamięć operacyjną DDR4 32 GB (RAM) i kartę graficzną NVIDIA RTX Titan 24 GB (GPU).

Model klasyfikacji obecności zmiany ogniskowej został poddany testowaniu w oparciu o hiperparametry (Tab. 5.10) określone na podstawie eksperymentów



Rys. 5.13: Wykresy pudełkowe efektywności autorskiej metody klasyfikacji obecności igły biopsyjnej w obrazie ultrasonograficznym wyznaczone dla modeli trenowanych z uwzględnieniem mechanizmu wcześniejszego zatrzymania oraz z wykorzystaniem techniki dostrajania.

poświęconych klasyfikacji obecności igły biopsyjnej przedstawionych w podrozdziale 5.3.1. Testowano podejścia oparte o różne wymiary początkowe obrazu z zakresu od $16 \times 16 \times 1$ do $256 \times 256 \times 1$, jak również różny czas trwania treningu modeli. Przedstawione wyniki odpowiadają modelom przyjmującym początkową reprezentację obrazów w skali szarości o wymiarach 256×256 z czasem



Rys. 5.14: Wykresy pudełkowe czasu klasyfikacji obrazu ultrasonograficznego za pomocą autorskiej metody klasyfikacji obecności igły biopsyjnej i metod referencyjnych, którymi są rozwiązania pretrenowane: AlexNet, VGG16, VGG19, SqueezeNet, GoogLeNet, ResNet101 oraz InceptionResNetV2. Wyniki sortowane są na podstawie liczby warstw ukrytych.

nauki uwzględniającym mechanizm wcześniejszego zatrzymania w przypadku wzrostu wartości funkcji straty, która stosowana jest celem minimalizacji błędów predykcji modelem.

Wyniki klasyfikacji obecności zmiany ogniskowej autorskimi oraz opublikowanymi modelami poddano ewaluacji za pomocą miar efektywności, wśród których znajdują się: dokładność (ACC), miara F_1 , współczynnik wyników prawdziwie pozytywnych TPR (czułość SEN), współczynnik wyników prawdziwie negatywnych TNR (swoistość SPE), wartość predykcyjna dodatniego wyniku (PPV) oraz wartość predykcyjna ujemnego wyniku (NPV), których definicje przedstawiono w podrozdziale 4.8.1. Oszacowany również został średni czas t klasyfikacji jednego obrazu przez proponowany model.

Analogicznie do eksperymentów dedykowanych ewaluacji klasyfikacji obecności igły biopsyjnej i zgodnie z opisem zawartym w podrozdziale 4.5.4, przeprowadzone zostały eksperymenty oceniające efektywność klasyfikacji obecności zmiany ogniskowej w dwuwymiarowych obrazach ultrasonograficznych. Proces uczenia uwzględniał mechanizm wcześniejszego zatrzymania (ang. *early stopping*) [200] oraz technikę dostrajania (ang. *fine-tuning*) [147]. Etap pierwszy treningu zakładał wykorzystanie danych przetworzonych za pomocą filtracji entropią, filtrem *shock* oraz metodą średnich nielokalnych zgodnie z opisem przedstawionym w podrozdziale 4.5.3. Drugi etap treningu techniką dostrajania zrealizowany został za pomocą oryginalnej bazy obrazów z zachowaniem podziału danych na poziomie pacjenta.

Proponowaną architekturę klasyfikującą zestawiono z opisanymi w literaturze i publicznie dostępnymi architekturami (Tab. 4.1): AlexNet [83], VGG16 [151], VGG19 [151], SqueezeNet [70], GoogLeNet [162], ResNet101 [56] oraz InceptionResNetV2 [161], które cechują się różnicami w liczbie warstw ukrytych, pa-

rametrami oraz wymiarami reprezentacji obrazu w warstwie wejściowej. Wstępny trening każdego z wymienionych modeli realizowany był za pomocą bazy obrazów ImageNet [139]. Wyznaczone w ten sposób wagi początkowe zostały dostrójone w procesie uczenia zbiorem obrazów ultrasonograficznych opisanych w podrozdziale 5.2.2 z zachowaniem 15-krotnego podziału metodą k-krotnej walidacji krzyżowej.

Mediana liczby epok, której wymagał trening modeli przyjmowała wartość 21 dla reprezentacji początkowej obrazu o wymiarach $256 \times 256 \times 1$. Ewaluacja wyników oszacowanych w oparciu o zbiory testowe dla modeli trenowanych oryginalnymi obrazami zmian ogniskowych i tkanek wskazuje medianę dokładności rozwiązania na poziomie $0,866 \pm 0,113$ (Rys. 5.15a). Wartość ta jest jedną z najwyższych w porównywanych testowanych rozwiązaniach, dla których zakres wartości kształtował się od $0,406 \pm 0,181$ w przypadku modeli InceptionResNetV2, do $0,878 \pm 0,062$ dla AlexNet. Najmniejsze rozrzuty wyników obserwuje się dla architektur SqueezeNet i GoogLeNet.

Wartości median w przypadku metryk F_1 , TPR i TNR wynosiły odpowiednio $0,857 \pm 0,121$, $0,860 \pm 0,107$ i $0,935 \pm 0,053$ dla autorskiej architektury klasyfikującej obecność zmiany ogniskowej w obrazach ultrasonograficznych. Wyniki te wskazują precyzyjniejszą klasyfikację w porównaniu do architektur ResNet101 oraz InceptionResNetV2. Te dwa rodzaje modeli były konkurencyjne do autorskiej metody na etapie wykorzystania jej celem szacowania prawdopodobieństwa obecności igły w obrazie USG (Rys. 5.13). W przypadku oceny obecności zmiany ogniskowej miara F_1 wykazała najniższą wartość mediany wynoszącą $0,601 \pm 0,2161$ w przypadku modeli InceptionResNetV2. Najwyższe wartości mediany dla metryki F_1 odnotowano dla modeli AlexNet na poziomie $0,858 \pm 0,068$ i autorskiej metody klasyfikacji. (Rys. 5.15b).

Współczynniki wyników prawdziwie dodatnich TPR i wyników prawdziwie negatywnych TNR przyjmują najniższe wartości mediany dla architektur VGG16, ResNet101 oraz InceptionResNetV2 (odpowiednio $0,333 \pm 0,233$, $0,333 \pm 0,182$ i $0,371 \pm 0,177$), natomiast najwyższe dla autorskiej metody oraz AlexNet ($0,852 \pm 0,066$), co zostało przedstawione za pomocą wykresów pudełkowych na rysunkach 5.15c i 5.15d. Wpływ na korzystniejsze wyniki dla autorskiej architektury może mieć fakt zastosowania zwiększającego się pola receptywnego w porównaniu do architektur VGG16, które oparte są o pola receptywne o wymiarach 3×3 , co w przypadku klasyfikacji złożonej i zaszumionej treści, jaką jest obraz USG, może negatywnie wpływać na wyniki. Wartości miar PPV i NPV dla wszystkich testowanych rozwiązań przedstawiono za pomocą wykresów na rysunkach 5.15e i 5.15f. Wyniki te również wskazują, że w przypadku klasyfikacji struktur o nieregularnych kształtach, jakimi są zmiany ogniskowe, korzystniejsze jest wykorzystanie architektur cechujących się mniejszą liczbą warstw ukrytych lub zróżnicowaniem wielkości wymiarów i liczby filtrów w poszczególnych warstwach. W przypadku oceny obecności

igły biopsyjnej, zabieg ten pozwala na korzystniejsze uogólnienie ekstrahowanych cech z uwzględnieniem lokalnych odkształceń tkanek patologicznych i zdrowych, co w przypadku zmian ogniskowych nie jest wyszczególnione ostrą granicą.

Analiza statystyczna oparta o nieparametryczny test Wilcoxona dla par obserwacji wskazuje brak istotnych różnic pomiędzy parą najkorzystniejszych rozwiązań – autorskiej metody klasyfikacji oraz architektury AlexNet ($p < 0,05$), jak również z dwoma kolejnymi – SqueezeNet i GoogLeNet ($p < 0,05$). Istotne statystycznie różnice istnieją pomiędzy dwoma najkorzystniejszymi rozwiązaniami a modelami ResNet101 oraz InceptionResNetV2 ($p < 0,05$).

Czas klasyfikacji jednego obrazu przyjmuje najniższą średnią wartość dla obrazów poddawanych przetwarzaniu autorską metodą klasyfikacji – 3 ms 5.16. Statyczne porównanie czasu przetwarzania jednego obrazu wykazuje istotne różnice dla każdej z analizowanych par ($p < 0,05$).

5.3.3 Detekcja igły biopsyjnej w obrazach USG

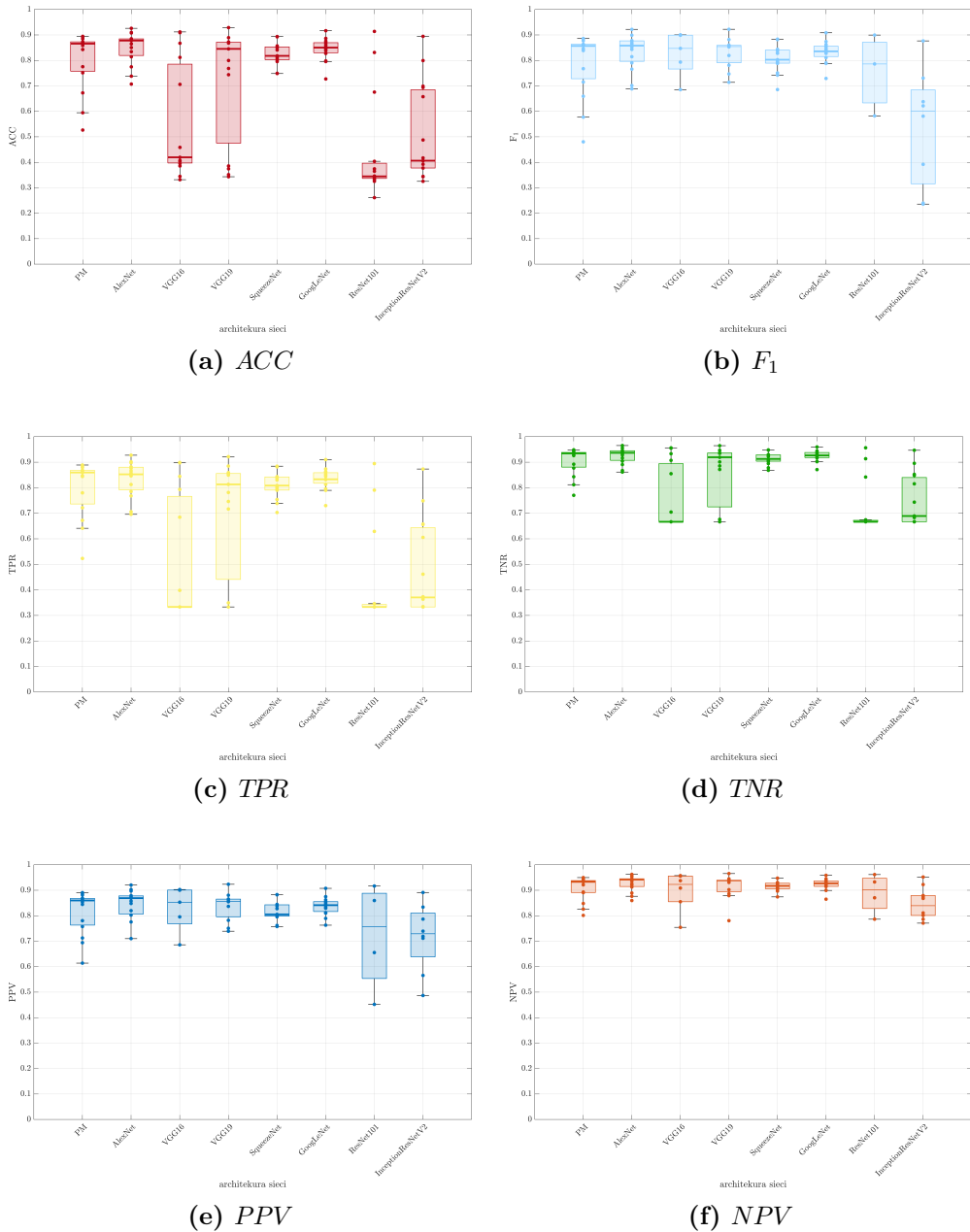
Obrazy z połączonej bazy danych, na których jest widoczna igła biopsyjna w liczbie 619 (Tab. 5.6), stanowiły podstawę do weryfikacji proponowanej metody detekcji igły biopsyjnej. Zbiór danych został losowo podzielony na trzy zestawy na poziomie identyfikatorów pacjentów: zbiór uczący, zbiór walidacyjny i zbiór testowy, w relacji 0,8 : 0,1 : 0,1. Przy pomocy zestawów uczącego i walidacyjnego przeprowadzono uczenie autorskiego modelu, a korzystając ze zbioru testowego sprawdzono jego efektywność. Podział danych został przeprowadzony 15 razy, tak samo jak etap uczenia z walidacją i etap testowania.

Przy tych samych założeniach przeprowadzono proces uczenia architektury zaproponowanej przez Lee i wsp. [89] oraz klasyfikator statystyczny opisany przez Hatt i wsp. [55]. Metody te wybrano do porównania, ponieważ w tych pracach wykorzystano bazy danych klinicznych zebranych podczas zabiegu biopsji gruboigłowej wykonywanej pod kontrolą ultrasonograficzną. Lee i wsp. zebrali 996 obrazów klinicznych od 8 pacjentów, podczas gdy Hatt i wsp. pracował z bazą danych składającą się z 577 obrazów klinicznych i 45 obrazów fantomów zwierzęcych.

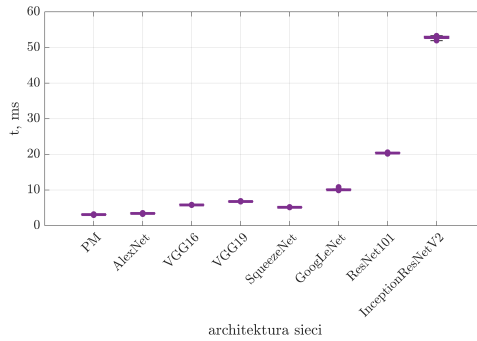
Autorska metoda oraz metoda Lee zostały wdrożone w środowisku MATLAB 2020b. Metoda Hatt została zaimplementowana w Pythonie 3.8.5. Eksperymenty zostały przeprowadzone na Windows 10 z platformą sprzętową zawierającą AMD Ryzen 9 3900 (CPU), 32 GB RAM 3GHz, NVIDIA RTX Titan 24 GB (GPU).

Wyniki segmentacji i detekcji igły za pomocą metod porównawczych skonfrontowano z proponowaną metodą.

Przeprowadzenie pojedynczego eksperymentu modelem zajęło średnio 22 ms dla proponowanej metody, 35 ms dla metody Lee i 7,64 s dla metody Hatt.



Rys. 5.15: Wykresy pudełkowe efektywności autorskiej metody klasyfikacji obecności zmiany ogniskowej w obrazie ultrasonograficznym wyznaczone dla modeli trenowanych z uwzględnieniem mechanizmu wcześniejszego zatrzymania oryginalną bazą obrazów.

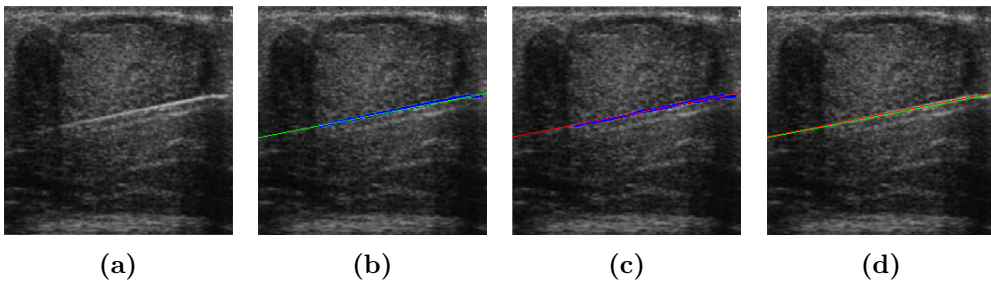


Rys. 5.16: Wykresy pudełkowe czasu klasyfikacji obrazu ultrasonograficznego za pomocą autorskiej metody klasyfikacji obecności zmiany ogniskowej i metod referencyjnych, którymi są rozwiązania pretrenowane: AlexNet, VGG16, VGG19, SqueezeNet, GoogLeNet, ResNet101 oraz InceptionResNetV2. Wyniki sortowane są na podstawie liczby warstw ukrytych.

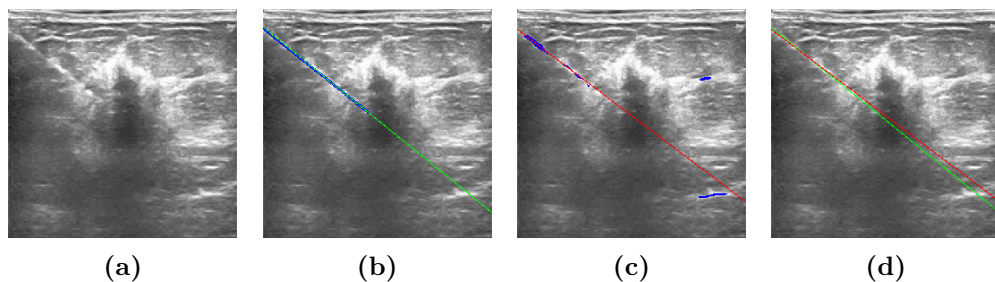
Wyniki segmentacji igły

Wyniki segmentacji semantycznej dla najlepszego, średniego i błędnego przypadku przedstawione zostały odpowiednio na rysunku 5.17, rysunku 5.18 oraz rysunku 5.19. Nawet dla średnich wyników segmentacji ocenianych miarą IoU , przedstawiona trajektoria leży bardzo blisko złotego standardu.

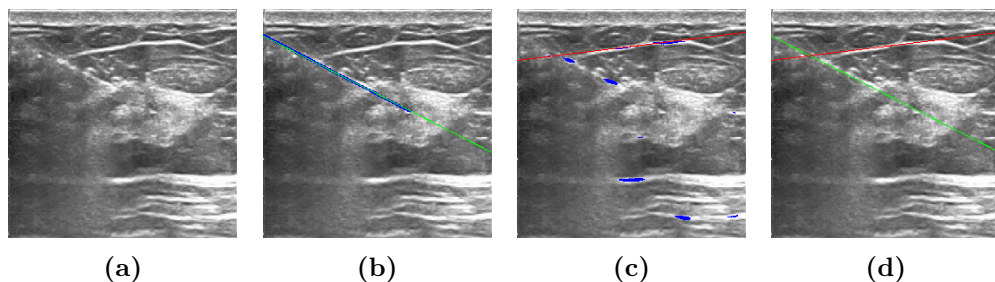
Segmentacja igły widocznej w obrazach ultrasonograficznych może być zakłócona obecnością długich, jasnych elementów, które są anatomicznie obecne w tkance gruczołu sutkowego (Rys. 5.20). Przykładami takich obiektów są granice płątów oraz tkanka podskórna. Struktury te mogą zostać nieprawidłowo rozpoznane jako igła, np. jak na rysunku 5.19.



Rys. 5.17: Najlepszy wynik segmentacji zmiany ogniskowej echoujemnej autorską metodą w ocenie za pomocą miary F_1 . (a) oryginalny obraz USG poddany przetwarzaniu, (b) złoty standard wyznaczony przez eksperta (kolor niebieski) z obliczoną trajektorią igły (kolor zielony), (c) wynik segmentacji proponowaną metodą (kolor niebieski) z wyznaczoną trajektorią igły (kolor czerwony), (d) obliczone trajektorie z obrazów (b) i (c) przedstawione razem.



Rys. 5.18: Wynik autorskiej metody segmentacji odpowiadający średnim wartościom $\Delta\lambda$ i $\Delta\theta$. (a) oryginalny obraz USG poddany przetwarzaniu, (b) złoty standard wyznaczony przez eksperta (kolor niebieski) z obliczoną trajektorią igły (kolor zielony), (c) wynik segmentacji proponowaną metodą (kolor niebieski) z wyznaczoną trajektorią igły (kolor czerwony), (d) obliczone trajektorie z obrazów (b) i (c) przedstawione razem.



Rys. 5.19: Nieprawidłowy wynik segmentacji igły autorską metodą. (a) oryginalny obraz USG poddany przetwarzaniu, (a) oryginalny obraz USG poddany przetwarzaniu, (b) złoty standard wyznaczony przez eksperta (kolor niebieski) z obliczoną trajektorią igły (kolor zielony), (c) wynik segmentacji proponowaną metodą (kolor niebieski) z wyznaczoną trajektorią igły (kolor czerwony), (d) obliczone trajektorie z obrazów (b) i (c) przedstawione razem.

Ocenę wyników segmentacji przeprowadzono dla metod opartych na konwulcyjnych sieciach neuronowych, tj. autorskiej metodzie (PM) i metodzie Lee i in. [89]. Obie metody zostały porównane również z klasyfikatorem statystycznym zaproponowanym przez Hatt i in. [55]. Obliczono następujące metryki: globalna dokładność (ACC_{glob}), średnia dokładność w klasach (ACC_{mean}), indeks Jaccarda (IoU), ważony indeks Jaccarda (IoU_w) i miara F_1 . Wyniki eksperymentów podsumowano w tabeli 5.11.

Globalne dokładności proponowanej metody i metody Hatt są do siebie podobne ($0,992\pm 0,003$ vs. $0,990\pm 0,004$). Globalna dokładność tych metod jest znacznie wyższa niż wynik otrzymany dla segmentacji za pomocą metody Lee ($0,603\pm 0,151$). Ponieważ obszar igły jest bardzo mały w porównaniu z tłem, wynik ten odzwierciedla przede wszystkim klasyfikację tła. Średnio obszar

Tab. 5.11: Podsumowanie statystyk dla wyników segmentacji regionu igły dla proponowanej metody (PM) oraz metod porównawczych: Lee i wsp. [89] i Hatt i wsp. [55].

stat.	ACC_{glob}	ACC_{mean}	IoU	IoU_w	F_1
	PM				
M	0,99	0,78	0,69	0,99	0,83
Me	0,99	0,79	0,69	0,99	0,85
SD	0,00	0,11	0,07	0,00	0,10
min	0,98	0,52	0,52	0,97	0,50
max	1,00	0,98	0,86	1,00	1,00
Lee i wsp.					
M	0,60	0,74	0,31	0,60	0,12
Me	0,58	0,74	0,30	0,58	0,11
SD	0,15	0,11	0,08	0,15	0,05
min	0,20	0,48	0,10	0,19	0,05
max	0,90	0,95	0,52	0,89	0,31
Hatt i wsp.					
M	0,99	0,51	0,50	0,98	0,50
Me	0,99	0,50	0,50	0,98	0,44
SD	0,00	0,01	0,01	0,01	0,14
min	0,98	0,50	0,49	0,96	0,38
max	1,00	0,58	0,57	0,99	0,91

segmentowanej igły stanowi $0,76\% \pm 0,38\%$ obrazu ultrasonograficznego. Stosunkowo niska wartość metody Lee spowodowana jest nadsegmentacją igły. W przypadku igły korzystniejszą sytuacją wydaje się jej niedosegmentowanie, ponieważ pojedyncze piksele zidentyfikowane poprawnie jako igła są wystarczające do wyznaczenia jej trajektorii.

Średnia dokładność w klasach nie różni się znacząco pomiędzy proponowaną metodą a metodą Lee ($0,778 \pm 0,105$ vs. $0,737 \pm 0,109$). W przypadku metody Hatt wartość miary ACC_{mean} jest niższa w porównaniu do pozostałych przypadków ($0,505 \pm 0,013$). Ten parametr jest miarą określającą wszystkie poprawnie sklasyfikowanych pikseli w poszczególnych klasach.

Miara IoU przyjmuje znacząco wyższe wartości dla PM niż dla metody Hatt i Lee ($0,686 \pm 0,068$ vs. $0,500 \pm 0,012$ i vs. $0,313 \pm 0,084$). Wynik ten dowodzi niższej liczby fałszywie dodatnich i fałszywie ujemnych wyników uzyskanych przez proponowane rozwiązanie.

Proponowana architektura sieci skutkuje podobnymi ważonymi wynikami segmentacji (IoU_w) w porównaniu do metody Hatt ($0,987 \pm 0,005$ vs. $0,982 \pm 0,007$) ale cechuje się wyższymi wartościami miary F_1 niż metoda Hatt ($0,831 \pm 0,102$ vs. $0,496 \pm 0,138$). Obie metody przyjmują dużo wyższe wartości IoU_w oraz F_1

niż średnie wartości tych miar wyznaczone na podstawie rezultatów segmentacji metodą Lee ($0,595 \pm 0,150$ i $0,116 \pm 0,047$).

Autorska metoda segmentacji igły biopsyjnej od początku testowana była na zbiorze danych klinicznych – zebranych podczas zabiegu biopsji gruboigłowej wykonywanej pod kontrolą USG. Zbiór jest różnorodny i liczebnością porównywalny z pracami, które również bazują na danych klinicznych. W przypadku metod porównawczych Lee i wsp. oraz Hatt i wsp. również były wykorzystane dane kliniczne ale od znacząco mniejszej liczby pacjentów (odpowiednio 8 i 6), ale brak różnorodności zarówno na poziomie pacjentów, jak również aparatury ultrasonograficznej, powoduje brak odporności tych metod na bardziej zróżnicowane dane.

Zaproponowany zbiór danych obrazowych igły biopsyjnej jest mniej liczny (Tab. 5.6), niż zbiory fantomowe ale w przypadku pracy na fantomie jest mniej ew. źródeł błędów (brak skóry w obrazie usg, w której występują białe liniowe struktury identyfikowane często na etapie segmentacji jako igła. Zbiory danych fantomowych charakteryzują się znacznie wyższą rozdzielczością obrazów, mniejszą liczbą artefaktów i bardziej jednorodnym tłem). Metody segmentacji cechują się brakiem odporności w przypadku treningu sieci tylko danymi fantomowymi lub jednorodnym zbiorem danych pochodzącym od bardzo małej grupy pacjentów [184]. Wynik segmentacji takim rozwiązaniem będzie obarczony większymi błędami w konfrontacji z szerszymi danymi klinicznymi.

Normalność rozkładów analizowanych miar została zbadana za pomocą testu Shapiro Wilka. Wynikowe wartości p wyznaczone dla poszczególnych metod przedstawione są w tabeli A.1. Hipoteza o rozkładzie normalnym badanej cechy została potwierdzona tylko w przypadku miary IoU wyznaczonej dla proponowanej architektury sieci ($p=0,796$). W pozostałych przypadkach badane cechy nie mają rozkładu normalnego ($p < 0,05$).

Ze względu na brak rozkładu normalnego większości analizowanych cech w kolejnym kroku analizy statystycznej przeprowadzony został test Wilcozona dla par obserwacji, aby porównać miary wyznaczone dla wyników segmentacji otrzymanych za pomocą proponowanej metody, metody Lee i metody Hatt. Określona została również wielkość efektu za pomocą współczynnika korelacji dwuseryjnej dla par dopasowanych (r). Wyniki zostały zebrane i podsumowane w tabelach A.2 (globalna dokładność ACC_{glob}), A.3 (średnia dokładność w klasach ACC_{mean}), A.4 (indeks Jaccarda IoU), A.5 (ważony indeks Jaccarda IoU_w) i A.6 (miara F_1).

Analiza statystyczna globalnej dokładności dla każdej z metod wykazała istotne statystycznie różnice pomiędzy uzyskanymi wartościami ($p < 0,05$). Wyniki przedstawiono w tabeli A.2. Zaobserwowano również duży efekt między metodą Lee a pozostałymi dwoma metodami ($0,7 \leq |r| < 0,9$); wielkość efektu jest średnia w przypadku porównania proponowanej architektury z metodą Hatt ($0,4 \leq |r| < 0,7$).

Tab. 5.12: Błędy detekcji trajektorii igły oraz średni czas analizy obrazu dla autor-
skiej metody (PM), metody Lee i Hatt. Symbole są wyjaśnione w sekcji 4.8.3.

Metoda	$NLSR_i$ [%]	$\Delta\lambda_{rms}$ [pix]	$\Delta\theta_{rms}$ [°]	t [ms]
PM	90,02	2,80	1,81	22,75
Lee i wsp.	73,15	18,62	10,65	35,08
Hatt i wsp.	29,94	30,06	12,92	76400

Średnia dokładność w klasach przyjmuje p -wartości $< 0,05$ dla wszystkich testowanych metod (Tab. A.3). Porównanie proponowanej architektury z metodą Lee wskazuje małą wielkość efektu pomiędzy nimi ($0,1 \leq |r| < 0,4$). W przypadku porównania proponowanej metody z metodą Hatt oraz metody Lee z metodą Hatt obserwowany jest duży efekt ($0,7 \leq |r| < 0,9$).

Indeks Jaccarda (IoU) wyznaczony dla każdej z metod został również poddany analizie statystycznej (Tab. A.4). Odnotowano istotne statystycznie różnice ($p < 0,05$) oraz dużą wielkość efektu ($0,7 \leq |r| < 0,9$) we wszystkich przypadkach. Podobne wyniki obserwuje się dla ważonego indeksu Jaccarda (IoU_w), które są przedstawione w tabeli A.5. Powodem dużego zróżnicowania wyników może być niski poziom szczegółowości igły w stosunku do tła (Rys. 5.20).

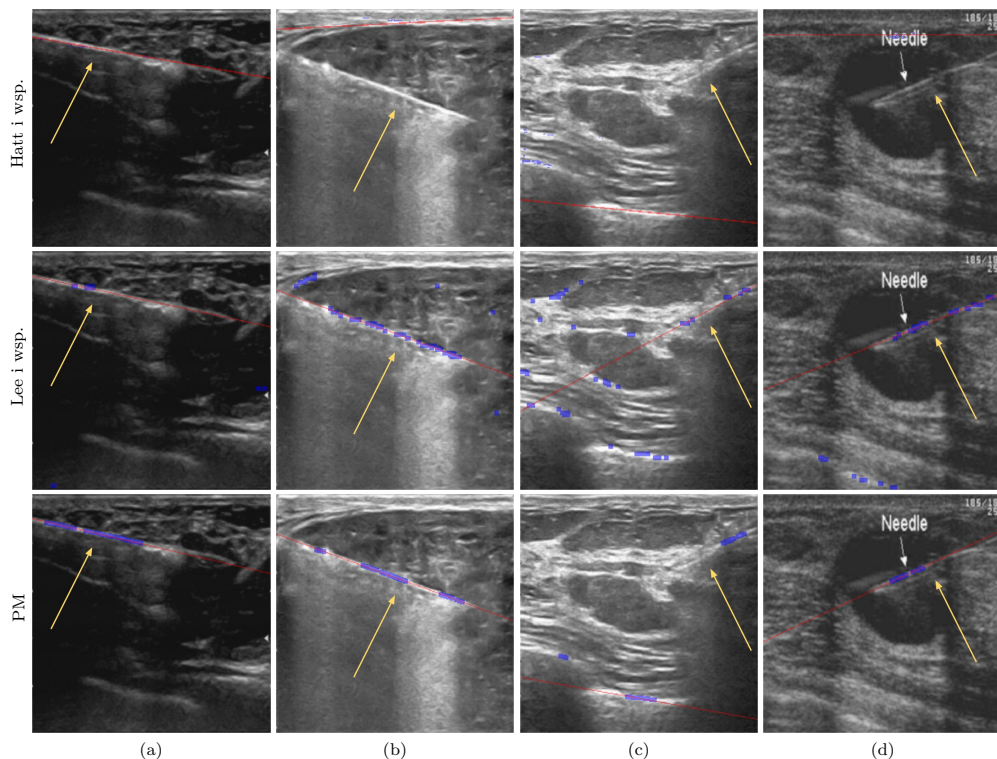
Analiza statystyczna porównująca proponowaną metodę segmentacji igły z metodami Lee i Hatt wykazała istotne statystycznie różnice ($p < 0,05$) i dużą wielkość efektu ($0,7 \leq |r| < 0,9$) w przypadku miary F_1 (Tab. A.6).

Porównanie statystyczne wartości metryk segmentacji za pomocą PM i pozostałych metod prawie w każdym przypadku wskazuje duży rozmiar efektu A. Wskazuje to, na duży rozrzut wyników na etapie segmentacji, co pozwala na wnioskowanie, że tylko detekcja igły oparta tylko o etap segmentacji nie jest wystarczająca do poprawnego wskazania igły i wymagane są dalsze etapy przetwarzania.

Wyniki detekcji trajektorii igły

Wskaźnik sukcesu lokalizacji igły ($NLSR$), $\Delta\lambda_{rms}$, $\Delta\theta_{rms}$ i średni czas przetwarzania obrazu (t) zostały obliczone dla proponowanej metody, jak również dla metod Lee i Hatt. Otrzymane wyniki podsumowano w tabeli 5.12. $NLSR$ ma najwyższą wartość dla proponowanej metody (90,02 %). Wartości pozostałych metryk podsumowujących trajektorię igły określono dla przypadków, w których $NLSR_i$ wynosiło 1. Odnotowano również wystąpienie najmniejszych wartości błędów średniokwadratowych $\Delta\lambda_{rms}$, $\Delta\theta_{rms}$ dla proponowanej metody w porównaniu do pozostałych. Średni czas przetwarzania obrazu był znacznie krótszy w przypadku proponowanej metody i metody Lee niż wymagany przez metodę Hatt (odpowiednio 23 ms vs. 35 ms vs. 76400 ms).

Średnie błędy detekcji trajektorii igły obliczone dla proponowanej metody



Rys. 5.20: Przykładowe wyniki jakościowe i ilościowe ($\Delta\lambda$, $\Delta\theta$) dla testowanych metod detekcji igły biopsyjnej. Wynik segmentacji oraz szacowana na jego podstawie trajektoria igły zaznaczone są odpowiednio niebieskim i czerwonym kolorem. Rzeczywiste położenie igły wskazywane jest żółtą strzałką. (a) poprawne wyniki uzyskane wszystkimi metodami: Hatt i wsp., Lee i wsp. i autorską metodą (PM), (b) igła jest wyraźnie widoczna w obrazie, ale nie przekłada się to na łatwość jej detekcji, (c) obecność jasnych, liniowych struktur oddziałują na proces przetwarzania obrazu w kontekście detekcji igły biopsyjnej, (d) przykładowe artefakty (adnotacja *Needle*) obecne w obrazach ultrasonograficznych.

to 1,60 pikseli i $1,04^\circ$ odpowiednio dla $\Delta\lambda$ i $\Delta\theta$. Te parametry przyjmują wartości 10,34 pikseli i $6,20^\circ$ dla metody Lee oraz 20,19 pikseli i $9,97^\circ$ dla metody Hatt (Tab. 5.13).

Średnie wartości parametrów λ i θ wyznaczone dla trajektorii igły w oparciu o złoty standard ($GT(x, y)$) oraz wyniki segmentacji dla każdej z metod zebrano w tabeli 5.14. Najmniejsze różnice między wartościami $GT(x, y)$ a metodami obliczeniowymi występują dla proponowanej metody (PM). Największe różnice natomiast obserwowane są pomiędzy $GT(x, y)$ a metodą Hatt.

Rozkład normalny wyznaczonych wektorów cech dla parametrów λ oraz θ oraz błędów trajektorii igły $\Delta\lambda$ i $\Delta\theta$ otrzymanych na podstawie wyników detekcji poszczególnymi metodami został zweryfikowany za pomocą testu Shapiro

Tab. 5.13: Podsumowanie statystyk błędów detekcji trajektorii igły na wyniku segmentacji autorską architekturą sieci (PM), metodzie Lee i wsp. i metodzie Hatt i wsp.

stat.	PM		Lee i wsp.		Hatt i wsp.	
	$\Delta\lambda$ [pix]	$\Delta\theta$ [°]	$\Delta\lambda$ [pix]	$\Delta\theta$ [°]	$\Delta\lambda$ [pix]	$\Delta\theta$ [°]
M	1,60	1,04	10,34	6,20	20,19	9,97
Me	1,11	0,91	2,85	2,09	13,96	8,59
SD	2,30	1,49	15,52	8,68	22,33	8,23
min	0,00	0,00	0,02	0,00	0,01	0,00
max	20,33	11,01	80,05	59,80	98,99	36,98

Tab. 5.14: Podsumowanie statystyk parametrów detekcji trajektorii igły λ i θ dla złotego standardu (GT), autorskiej metody (PM) oraz metod porównawczych: Lee i wsp. i Hatt i wsp.

stat.	GT		PM		Lee i wsp.		Hatt i wsp.	
	λ [pix]	θ [°]	λ [pix]	θ [°]	λ [pix]	θ [°]	λ [pix]	θ [°]
M	39,20	20,96	39,51	20,92	45,93	18,30	53,89	15,86
Me	37,44	22,00	38,13	22,00	46,92	18,84	53,18	14,93
SD	18,51	6,78	17,99	6,27	21,46	9,74	24,01	9,76
min	3,54	2,02	0,30	0,00	5,83	0,00	2,16	0,00
max	86,60	36,98	84,93	36,11	106,07	79,84	125,98	45,00

Wilka (Tab. A.7). Otrzymane p -wartości dla każdej z badanych miar wykazały brak rozkładu normalnego ($p < 0,05$) z wyłączeniem parametru θ wyznaczonego dla proponowanej metody.

Ze względu na brak rozkładu normalnego w parach porównywanych cech analiza statystyczna została przeprowadzona z użyciem testu Wilcozona dla par obserwacji. Aby porównać wartości parametrów i błędów trajektorii wyznaczonych na podstawie wyników detekcji igły autorską metodą detekcji igły, metodą Lee i metodą Hatt określona została wielkość efektu za pomocą współczynnika korelacji dwuseryjnej dla par dopasowanych (r). Wyniki zostały zebrane i podsumowane w tabelach A.8 w przypadku parametru λ , A.9 dla parametru θ oraz odpowiednio A.10 i A.11 dla miar błędów detekcji $\Delta\lambda$ i $\Delta\theta$.

Otrzymane wyniki wskazują brak istotnych statystycznie różnic w porównaniu pomiędzy $\lambda_{R_{GT}}$ a trajektorią wyznaczoną za pomocą proponowanej metody λ_R ($p > 0,05$). Istotne statystycznie różnice występują we wszystkich pozostałych badanych parach ($p < 0,05$). Wyniki przedstawiono w tabeli A.8. Porównanie parametru λ_R wyznaczonego na podstawie wyniku metody Lee z pozostałymi metodami, każdorazowo wskazuje na małą wielkość efektu ($0,1 \leq |r| < 0,4$). Metoda Hatt porównana ze złotym standardem oraz PM cechuje się średnią wielkością efektu ($0,4 \leq |r| < 0,7$).

Porównanie wartości kąta θ_R trajektorii igły wyznaczonej proponowaną metodą z kątem θ_{RGT} otrzymanym dla złotego standardu nie wykazało istotnych statystycznie różnic pomiędzy uzyskanymi wartościami ($p > 0,05$). Wartości kąta otrzymane na podstawie wyniku segmentacji metodami porównawczymi porównane z kątem rzeczywistym określonym na podstawie obrysu eksperckiego różnią się istotnie statystycznie ($p < 0,05$). Otrzymana p -wartość ($< 0,05$) w analizie podobieństwa pomiędzy metodami obliczeniowymi wykazała również istotne statystycznie różnice (Tab. A.11). Wielkość efektu jest mała dla każdej z analizowanych par ($0,1 \leq |r| < 0,4$) z wyłączeniem porównania proponowanej metody z metodą Hatt, gdzie współczynnik r wskazuje średnią wielkość efektu ($0,4 \leq |r| < 0,7$).

Błędy detekcji trajektorii igły $\Delta\lambda_R$ i $\Delta\theta_R$ określone dla proponowanej metody i metod porównawczych zostały również ocenione statystycznie (Tab. A.10 i A.11). Uzyskane wyniki wskazują istotne statystycznie różnice między badanymi parami dla obu miar błędów ($p < 0,05$) i średnią wielkość efektu w porównaniu proponowanej metody z metodą Lee ($0,4 \leq |r| < 0,7$) oraz duży efekt w porównaniu PM z metodą Hatt ($0,7 \leq |r| < 0,9$). Wielkość efektu jest średnia w przypadku $\Delta\lambda_R$ ($0,4 \leq |r| < 0,7$) i mała dla $\Delta\theta_R$ ($0,1 \leq |r| < 0,4$) w analizie pary metod Lee i Hatt.

Analiza statystyczna wykazała, iż w przypadku oceny trajektorii igły autorska metoda cechuje się małymi istotnymi różnicami w odniesieniu parametrów złotego standardu (λ_{GT} , θ_{GT}), co pozwala stwierdzić, że PM jest porównywalna z ekspertem. Proces detekcji igły autorską metodą cechuje się dużą dokładnością, niemniej jednak jest obarczony pewnymi ograniczeniami. Obecność strzałek czy tekstu, a także bardzo intensywne granice sąsiadujących tkanek miękkich lub kości wzdłuż trajektorii igły mogą zakłócać widoczność igły (Rys. 5.20), jak również odmienne od standardowych głębokie wklęcia pod kątem ostrym, co potwierdza istotność zastosowania złożonego zbioru danych klinicznych w procesie uczenia.

Igła nie jest obiektem całkowicie sztywnym i może przybrać łukowaty kształt w przypadku zastosowania zbyt dużej siły podczas jej wprowadzania w tkanki. W tej sytuacji każda z rozważanych metod szacuje trajektorię z większym błędem zależnym od stopnia krzywizny igły. To samo zdanie dotyczy pozostałych metod. Mimo wskazanych wyżej trudności autorska metoda detekcji cechuje się wartościowymi wynikami, przewyższającymi inne rozwiązania. Korzyścią dla pacjenta w przypadku praktycznego zastosowania autorskiej metody jest zapis przemieszczenia całej igły (jej trajektorii), ponieważ istnieją doniesienia, że biopsje mogą być czynnikami ryzyka rozsiewu nowotworu po biopsji igłowej piersi [143].

Histogramy błędów detekcji trajektorii igły wyznaczonych dla proponowanej metody oraz metod porównawczych Lee i Hatt przedstawiono na rysunku 5.21. W przypadku wszystkich błędów prezentowane są wartości średnie,

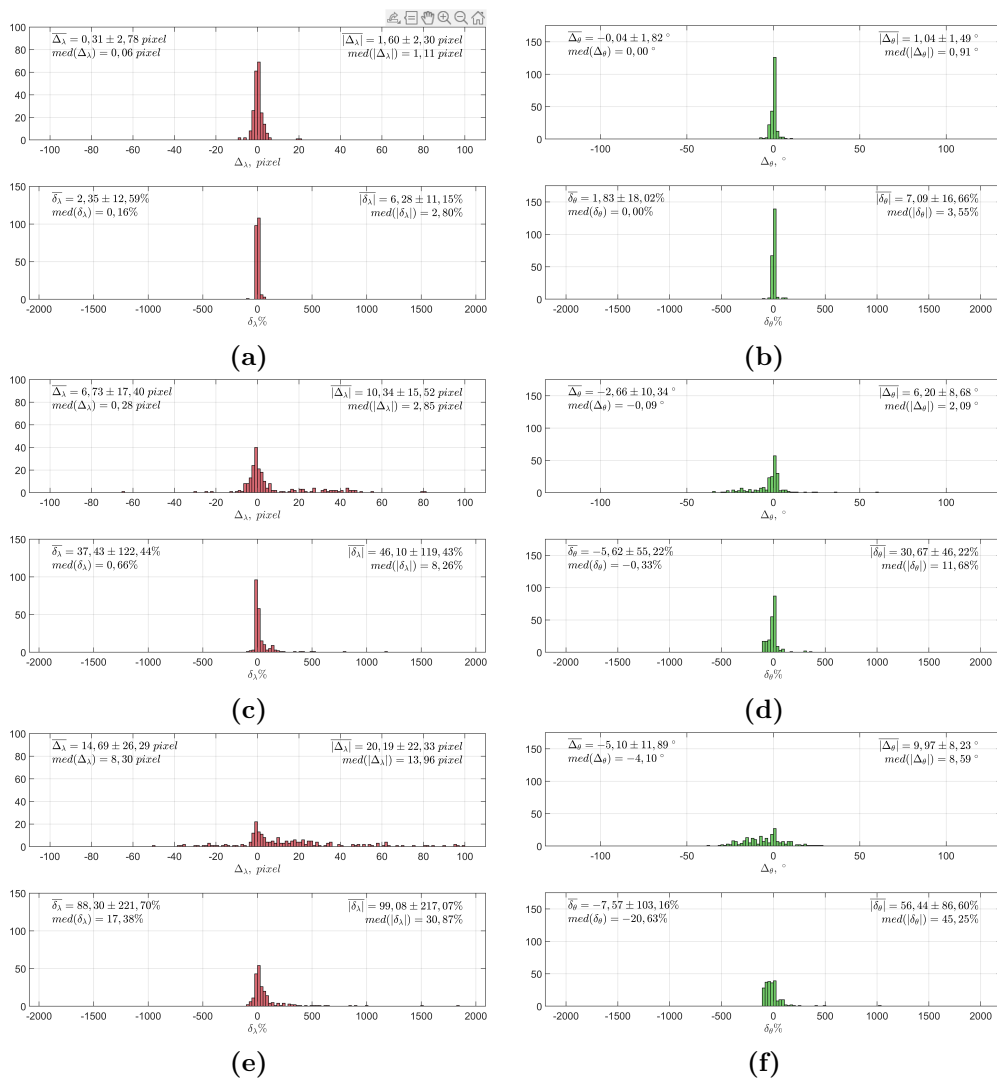
odchylenia standardowe i mediany zarówno dla błędów bezwzględnych i względnych. Średni błąd bezwzględny (MAE) dla błędu dystansu $|\Delta_\lambda|$ wynosi 1,6 piksela (Rys. 5.21a), natomiast w przypadku błędu kąta $|\Delta_\theta|$ wartość MAE wynosi $1,04^\circ$ (Rys. 5.21b). Wartość średniego bezwzględnego błędu procentowego ($MAPE$) w przypadku oceny błędów dystansu $|\delta_\lambda|$ wynosi 6,28% przy wartości mediany tego błędu na poziomie 2,80%. Ocena błędu kąta $|\delta_\theta|$ za pomocą $MAPE$ dla proponowanej metody szacuje się na poziomie 7,09% przy wartości mediany na poziomie 3,55%. W przypadku metody Lee wartość MAE odpowiednio dla błędów dystansu i kąta wynosi 10,34 pikseli i $6,20^\circ$ (Rys. 5.21c i Rys. 5.21d), natomiast błąd $MAPE$ kształtuje się na poziomie 46,10% dla $|\delta_\lambda|$ i 30,67% dla $|\delta_\theta|$ w odniesieniu do wartości rzeczywistych, za które uznaje się parametry wyznaczone z wykorzystaniem złotego standardu. Metoda Hatt cechuje się największymi wartościami średnimi błędów bezwzględnych spośród porównywanych podejść. Błąd MAE metody Hatt wynosi 20,19 pikseli dla $|\delta_\lambda|$ i $9,97^\circ$ w przypadku $|\delta_\theta|$ przy średnich wartościach bezwzględnego błędu procentowego 99,08% i 56,44% odpowiednio dla dystansu i kąta (Rys. 5.21e i Rys. 5.21f).

5.3.4 Segmentacja zmiany ogniskowej w obrazach USG

Połączone cztery źródła zbiorów dwuwymiarowych obrazów ultrasonograficznych, stanowiły podstawę do weryfikacji proponowanej metody segmentacji zmiany ogniskowej. W tym celu wykorzystano 913 obrazów USG zmiany ogniskowej echoujemnej oraz 683 obrazy zmiany ogniskowej echodatniej (Tab. 5.8). Dla każdego obrazu zawierającego zmianę ogniskową przeprowadzone zostały segmentacje, poprzedzone filtracją metodą średnich nielokalnych, za pomocą metod: aktywnego konturu (ang. *active contour*, AC) [43, 60], aktywnego konturu wykorzystującego obraz gradientowy (ang. *gradient vector flow*, GVF) [137], rozrostu obszaru (ang. *region growing*, RG) [90] oraz algorytmu wododziałowego (ang. *watershed segmentation*, WS) [67]. Każda z metod była testowana na obrazach klinicznych odpowiednio w liczbie 6 i 46 dla AC, 48 dla GVF, 30 w przypadku RG i 60 dla WS. Raportowane wyniki należały do zakresu 67,56% – 88,80% miary F_1 . Maski wygenerowane za pomocą poszczególnych metod następnie zostały zagregowane i ocenione metodą komitetu klasyfikującego (ang. *ensemble*).

Proponowane podejście zostało zaimplementowane w środowisku MATLAB 2020b. Eksperymenty zostały przeprowadzone z wykorzystaniem systemu operacyjnego Windows 10 z platformą sprzętową zawierającą procesor AMD Ryzen 9 3900 (CPU), pamięć operacyjną 32 GB RAM 3GHz i kartę graficzną NVIDIA RTX Titan 24 GB (GPU).

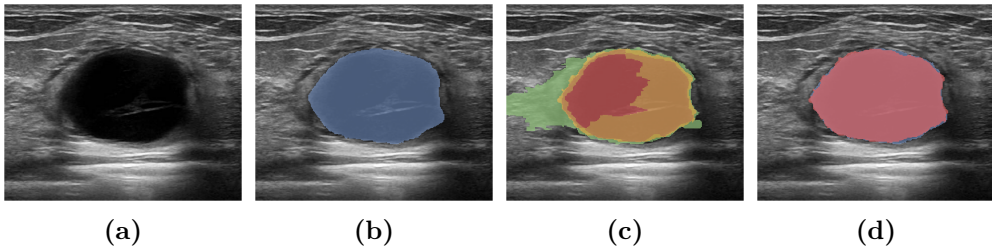
Przeprowadzenie segmentacji metodami składowymi odpowiednio dla każdej z nich trwało średnio: AC – 16 ms, GVF – 20 ms, RG – 5 ms, WS – 6 ms.



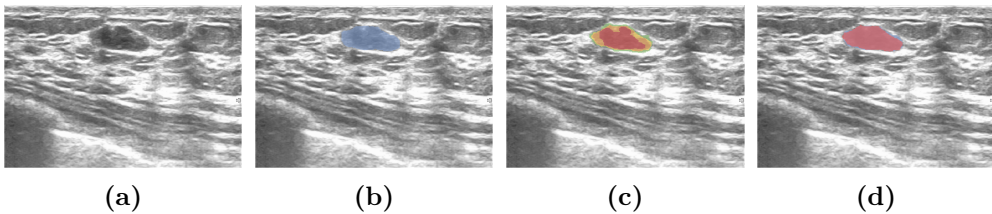
Rys. 5.21: Histogramy efektywności automatycznej detekcji trajektorii igły Δ_λ i Δ_θ dla autorskiej metody (Rys. 5.21a i 5.21b) oraz metod Lee (Rys. 5.21c i 5.21d) i Hatt (Rys. 5.21e i 5.21f): wartości błędów bezwzględnych Δ (górny histogram każdego z rysunków) i względnych δ (dolny histogram każdego z rysunków). Wartość rzeczywista odpowiada parametrom $\lambda_{R_{GT}}$ i $\theta_{R_{GT}}$ wyznaczonych na podstawie złotego standardu. W prawym górnym rogu przedstawiono wartości średniej z odchyleniem standardowym oraz mediany analizowanego błędu.

Czas trwania całego procesu przetwarzania obrazu z uwzględnieniem głosowania komitetem klasyfikującym wyniósł średnio 49 ms.

Wyniki segmentacji dla najlepszego oraz średniego przypadku przedstawione zostały odpowiednio dla zmian ogniskowych echoujemnych na rysun-



Rys. 5.22: Najlepszy wynik segmentacji zmiany ogniskowej echoujemnej autorską metodą w ocenie za pomocą miary F_1 . (a) oryginalny obraz USG poddany przetwarzaniu, (b) złoty standard wyznaczony przez eksperta zaznaczony kolorem niebieskim, (c) wynik komitetu klasyfikującego dla kolejnych progów odcięcia zgodnie z gradacją: region wskazany wszystkimi metodami zaznaczono na czerwono, trzema na pomarańczowo, dwoma na żółto i jedną na zielono, (d) porównanie wyniku segmentacji z wykorzystaniem komitetu klasyfikującego z progiem odcięcia dwa (kolor czerwony) z obrysem eksperckim (kolor niebieski).



Rys. 5.23: Najlepszy wynik segmentacji zmiany ogniskowej echodatniej autorską metodą w ocenie za pomocą miary F_1 . Opisy są takie same, jak na rysunku 5.22.

kach 5.22 oraz 5.24, natomiast dla zmian ogniskowych echodatnich na rysunkach 5.23 i 5.25.

Segmentacja zmian ogniskowych widocznych w obrazach USG może być zakłócona obecnością innych kulistych elementów o zbliżonej wartości intensywności pikseli, których przykładem mogą być struktury anatomicznie obecne w tkance gruczołu sutkowego lub środek znieczulający wprowadzony przed wykonaniem biopsji. Cechą charakterystyczną wpływającą na proces segmentacji jest brak jednoznacznej granicy pomiędzy tkanką zmiany ogniskowej a tkanką zdrową, co przedstawiono na rysunku 5.24.

Ocenę wyników segmentacji proponowaną metodą, jak również kroków pośrednich, przeprowadzono wykorzystując klasyczne metryki oceny jakości segmentacji: dokładność (ACC), współczynnik wyników prawdziwie pozytywnych (TPR), współczynnik wyników prawdziwie negatywnych (TNR), przecięcia sumy (indeksu Jaccarda, IoU) oraz miary F_1 (indeks Dice'a), które zostały zdefiniowane w podrozdziale 4.8.

Wyniki segmentacji zmiany ogniskowej metodami klasycznymi

Poszczególne dokładności dla klasycznych metod segmentacji rozważane dla wszystkich zmian ogniskowych w analizowanym zbiorze obrazów zebrane zostały w tabeli 5.15. Najwyższe wartości dokładności odnotowano dla metody aktywnego konturu wykorzystującego obraz gradientu ($0,930 \pm 0,096$). Najniższa wartość średnia miary ACC cechuje metodę rozrostu obszaru ($0,882 \pm 0,183$).

W przypadku oceny miarą TPR najwyższy wynik obserwuje się w przypadku metody WS ($0,670 \pm 0,250$), natomiast najniższy odnotowano w segmentacji metodą AC ($0,528 \pm 0,291$). Wartości metryki TNR wskazują jednak niższy średni wynik WS ($0,922 \pm 0,126$) niż w przypadku AC ($0,993 \pm 0,014$), w których to wynikach znajduje odwzorowanie nadsegmentacja zmian ogniskowych w przypadku wykorzystania WS i niedoszacowanie jej powierzchni za pomocą AC. Szacowanie regionu zainteresowania odpowiadającego zmianie ogniskowej jest istotne w przypadku zabiegu biopsji gruczołowej celem wskazania obszarów właściwych do pobrania przez lekarza wycinków. Proces nadsegmentacji może powodować pobranie zdrowej tkanki, co będzie skutkowało wynikiem fałszywie negatywnym. Pobranie wycinka z środkowej części zmiany, co może być skutkiem niedosegmentowanej powierzchni zmiany, również może być obciążone ryzykiem braku możliwości szerszej oceny w przypadku zmian heterogenicznych.

Większe zróżnicowanie wyników obserwowane jest dla miar przecięcia sumy oraz F_1 . Segmentacja klasycznymi metodami wskazuje najkorzystniejsze rezultaty w przypadku algorytmu GVF, dla którego średnia wartość wynosi $0,490 \pm 0,259$ w przypadku IoU oraz $0,612 \pm 0,265$ dla miary F_1 . Najniższe średnie wartości miar IoU i F_1 cechują ocenę wyników wyznaczonych metodą rozrostu obszaru, które odpowiednio wynoszą $0,360 \pm 0,265$ oraz $0,475 \pm 0,287$. Wpływ na rezultaty działania poszczególnych metod ma niejednorodny, zaszumiony charakter obrazu, co w przypadku prostszych metod, których przykładem jest RG, odzwierciedla się mniejszą liczbą pikseli spełniających zadane kryterium.

Wyniki segmentacji zmiany ogniskowej za pomocą komitetu klasyfikującego

Suma masek składowych i wybranie za pomocą komitetu klasyfikującego tych pikseli, które zostały wskazane przez co najmniej dwie z czterech metod, skutkowało średnią dokładnością wyników wynoszącą $0,962 \pm 0,095$ (Tab. 5.16). Właśnie to kryterium przyjęto za najkorzystniejsze, ze względu na eliminację ryzyka nadsegmentacji zmiany ogniskowej, co wskazane zostało kolorem zielonym na rysunku 5.25c. Zastosowanie komitetu z odcięciem na poziomie dwóch metod skutkuje również wzrostem średnich wartości wyników prawdziwie pozytywnych i prawdziwie negatywnych w porównaniu do wszystkich metod składowych z wyłączeniem algorytmu wododziałowego.

Komitet R_2 cechuje się średnimi wartościami na poziomie $0,663 \pm 0,257$ dla

Tab. 5.15: Podsumowanie statystyk dla wyników segmentacji regionu zmiany ogniskowej niezależnie od jej typu znanymi metodami klasycznej segmentacji obiektów w obrazach.

stat.	<i>ACC</i>	<i>TPR</i>	<i>TNR</i>	<i>IoU</i>	F_1
aktywny kontur – AC					
M	0,93	0,53	0,99	0,47	0,59
Me	0,96	0,55	1,00	0,50	0,67
SD	0,10	0,29	0,01	0,27	0,27
min	0,35	0,00	0,79	0,00	0,00
max	1,00	1,00	1,00	0,95	0,97
aktywny kontur oparty o obraz gradientowy – GVF					
M	0,93	0,58	0,99	0,49	0,61
Me	0,96	0,62	1,00	0,53	0,69
SD	0,10	0,31	0,02	0,26	0,26
min	0,36	0,00	0,78	0,00	0,00
max	1,00	1,00	1,00	0,95	0,97
rozrost obszaru – RG					
M	0,82	0,61	0,84	0,36	0,47
Me	0,88	0,65	0,95	0,31	0,47
SD	0,18	0,28	0,20	0,27	0,29
min	0,09	0,00	0,02	0,00	0,00
max	1,00	1,00	1,00	0,97	0,98
algorytm wododziałowy – WS					
M	0,89	0,67	0,92	0,45	0,58
Me	0,93	0,72	0,97	0,43	0,60
SD	0,13	0,26	0,13	0,24	0,24
min	0,04	0,00	0,00	0,00	0,00
max	1,00	1,00	1,00	0,97	0,98

miary *TPR*, co jest wynikiem nieznacznie niższym w porównaniu do metody WS ($0,670 \pm 0,250$). Algorytm wododziałowy cechuje się jednak niższym wynikiem w przypadku miary *TNR*, która w tym przypadku wynosi $0,922 \pm 0,126$ a dla zaproponowanego komitetu jest równa $0,972 \pm 0,058$. Porównanie w tym przypadku median pomiędzy metodami WS a R_2 , które dla każdej z nich wynoszą kolejno 0,973 i 0,994, wskazuje na możliwość ograniczenia nadsegmentacji zmiany ogniskowej, która jest obecna w przypadku metody opartej o wododział.

Dla zabiegu biopsji gruboigłowej istotniejszym jest wskazanie obszaru, dla którego prawdopodobieństwo występowania zmiany jest wysokie, nawet kosztem jej niedosegmentowania, niż włączenie w rozważany obszar regionu będą-

cego zdrową tkanką. Podejście to jest obarczone ryzykiem wyniku fałszywie negatywnego na poziomie histopatologicznej oceny pobranych wycinków.

Większe zróżnicowanie wyników obserwowane jest dla miar przecięcia sumy oraz F_1 . W obydwóch przypadkach wyniki wyznaczone dla komitetu R_2 są wyższe w porównaniu do indywidualnie ocenianych metod segmentacji. Segmentacja klasycznymi metodami wskazuje najkorzystniejsze rezultaty w przypadku algorytmu GVF, dla którego średnia wartość wynosi $0,490 \pm 0,259$ w przypadku IoU oraz $0,612 \pm 0,265$ dla miary F_1 . Zastosowanie komitetu klasyfikującego pozwoliło podwyższyć te wyniki do poziomu $0,540 \pm 0,243$ i $0,665 \pm 0,235$. Otrzymane wartości są niższe od raportowanych w literaturze dla wykorzystanych metod klasycznych, niemniej jednak w niniejszej pracy poprawność ich działania weryfikowana była na znacząco większym i bardziej zróżnicowanym zbiorze obrazów ultrasonograficznych.

Analiza statystyczna wyników segmentacji zmiany ogniskowej za pomocą metod klasycznych agregowanych komitetem klasyfikującym

Analiza statystyczna dokładności dla każdej z rozważanych metod wykazała istotne statystycznie różnice pomiędzy uzyskanymi wartościami dla każdej z badanych par ($p < 0,05$), co zostało przedstawione w tabeli A.13. Zaobserwowano również średni efekt między klasycznymi metodami RG i WS a uznanym za najkorzystniejszy komitetem R_2 ($0,4 \leq |r| < 0,7$). Komitet R_2 cechuje się również średnim efektem w porównaniu z pozostałymi komitetami ($0,4 \leq |r| < 0,7$) z wyłączeniem R_3 , gdzie występuje mały efekt ($0,1 \leq |r| < 0,4$). Wielkość efektu na poziomie małym obserwowana jest dla pozostałych porównywanych par ($0,1 \leq |r| < 0,4$).

Porównanie statystyczne par dla miary TPR wykazało brak istotnych różnic pomiędzy metodami GVF i RG ($p > 0,05$). Pozostałe wyniki są istotnie statystycznie różne ($p < 0,05$). Wyznaczone dla nich wielkości efektu wskazują średni efekt pomiędzy algorytmem wododziałowym a pozostałymi rozważanymi przypadkami ($0,4 \leq |r| < 0,7$), z wyłączeniem porównania z komitetem R_2 , gdzie wartość efektu sugeruje jego brak ($|r| < 0,1$). Powodowane jest to małym zróżnicowaniem otrzymywanych wartości, co obserwowane jest na poziomie statystyk opisowych dla każdego ze zbiorów. Komitet R_2 cechuje się również pomijalnym rozmiarem w efekcie w porównaniu z komitetem R_1 , co sugeruje na poziomie tej miary możliwość wykorzystania każdego z tych podejść. Odpowiednio średni efekt jest obserwowany w porównaniu z algorytmem AC ($0,4 \leq |r| < 0,7$) i mały dla metod GVF i RG ($0,1 \leq |r| < 0,4$). Wskazuje to na niewielkie, ale istotne różnice w przypadku dwóch ostatnich metod, które przemawiają za korzyścią wprowadzenia rozwiązania opartego o komitet klasyfikujący (Tab. A.14).

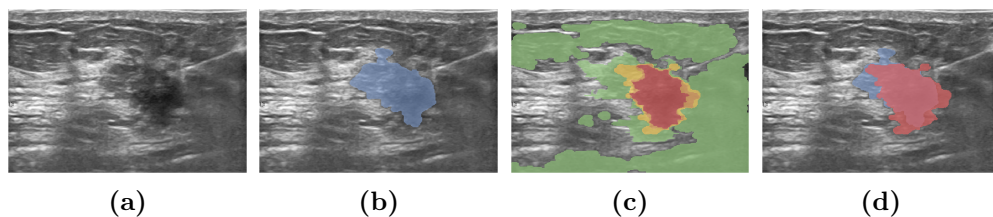
Test Wilcozona dla par obserwacji wykazał istotne statystycznie różnice dla każdej z porównywanych par w przypadku miary TNR ($p < 0,05$), co zostało

Tab. 5.16: Podsumowanie statystyk dla wyników segmentacji regionu zmiany ogniskowej z wykorzystaniem komitetu klasyfikującego na podstawie klasycznych metod segmentacji.

stat.	<i>ACC</i>	<i>TPR</i>	<i>TNR</i>	<i>IoU</i>	F_1
jedna metoda – R_1					
M	0,79	0,86	0,78	0,41	0,53
Me	0,84	0,94	0,85	0,36	0,53
SD	0,19	0,18	0,22	0,26	0,27
min	0,04	0,00	0,00	0,00	0,00
max	1,00	1,00	1,00	0,96	0,98
\geq dwie metody – R_2					
M	0,93	0,66	0,97	0,54	0,67
Me	0,96	0,71	0,99	0,57	0,73
SD	0,10	0,26	0,06	0,24	0,24
min	0,42	0,00	0,43	0,00	0,00
max	1,00	1,00	1,00	0,96	0,98
\geq trzy metody – R_3					
M	0,93	0,51	0,99	0,47	0,59
Me	0,97	0,56	1,00	0,51	0,67
SD	0,10	0,29	0,01	0,27	0,28
min	0,32	0,00	0,79	0,00	0,00
max	1,00	1,00	1,00	0,94	0,97
cztery metody – R_4					
M	0,92	0,35	0,99	0,34	0,45
Me	0,96	0,34	1,00	0,33	0,50
SD	0,11	0,27	0,01	0,26	0,30
min	0,31	0,00	0,90	0,00	0,00
max	1,00	0,98	1,00	0,96	0,98

podsumowane w tabeli A.15. Obserwowana jest również średnia wielkość efektu dla znaczącej większości par ($0,4 \leq |r| < 0,7$). Mała wielkość efektu cechuje porównanie metody rozrostu obszaru z algorytmami GVF, WS i komitetem R_2 oraz parę aktywny kontur i komitet R_4 ($0,1 \leq |r| < 0,4$). Pozostałe przypadki wskazują na istotność wyboru rozwiązania cechującego się najwyższymi wartościami statystyk, celem wskazania podejścia cechującego się jak najmniejszą liczbą fałszywie pozytywnych wskazań pikseli, celem uniknięcia włączenia do dalszej analizy próbek odpowiadających zdrowej tkance gruczołu sutkowego.

Analiza statystyczna wartości miary *IoU* wykazała brak istotnych statystycznie różnic dla pary GVF i R_3 ($p > 0,05$), co może sugerować, iż komitet klasyfikujący definiowany kryterium przyłączenia na poziomie wskazania



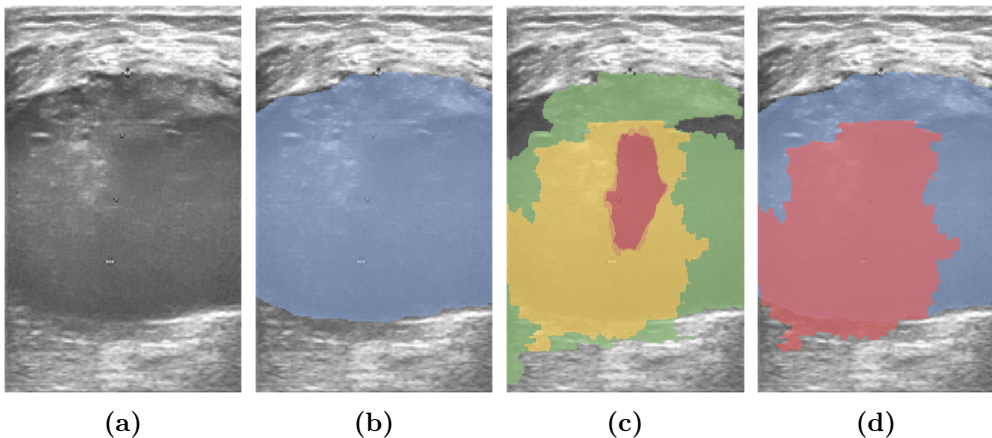
Rys. 5.24: Wynik segmentacji zmiany ogniskowej echoujemnej autorską metodą dla średniej wartości miary F_1 . Opisy są takie same, jak na rysunku 5.22.

piksela przez trzy i więcej metod, może cechować się niedosegmentowaniem zmiany ogniskowej w obrazach (Tab. A.16). W pozostałych porównywanych parach obserwacji występują istotne statystycznie różnice ($p < 0,05$) cechujące się w większości przypadków małą wielkością efektu ($0,1 \leq |r| < 0,4$). Wielkość efektu na poziomie średniego obserwowana jest w porównaniu komitetu R_4 z pozostałymi analizowanymi metodami wyznaczania regionu zmiany ogniskowej w obrazach ultrasonograficznych ($0,1 \leq |r| < 0,4$). Komitet R_4 jest najbardziej restrykcyjnym z analizowanych podejść i cechuje się najniższymi średnimi wartościami metryk. Na poziomie oceny miarą F_1 , komitet R_2 wykazuje pomijalną wielkość efektu w porównaniu z metodą rozrostu obszaru ($|r| < 0,1$). Metoda ta jest najprostszą z omawianych w niniejszej pracy, i jest oparta o podobieństwo analizowanego obszaru, dlatego sprawdza się w przypadku jednorodnych struktur, niemniej jednak jest obciążona wysokim ryzykiem znaczącego niedosegmentowania heterogenicznych zmian ogniskowych. Brak efektu może sugerować, iż restrykcyjne kryterium oceny pikseli przy zastosowaniu komitetów klasyfikujących będzie skutkowało znaczącym ograniczeniem segmentowanego obszaru, co może powodować błędy w procesie diagnostycznym.

Ocena statystyczna miary F_1 wskazuje brak istotnych różnic w porównaniu algorytmu wododziałowego z komitetem R_3 ($p > 0,05$), oraz istotne statystycznie różnice dla każdej z pozostałych par ($p < 0,05$), co zostało przedstawione w tabeli A.17. Większość porównywanych par cechuje mała wielkość efektu ($0,1 \leq |r| < 0,4$) z wyłączeniem par, które cechowały się odpowiednio najwyższymi i najniższymi średnimi wynikami. W tych przypadkach współczynnik korelacji dwuseryjnej dla par dopasowanych (r) wskazuje średnią wielkość efektu ($0,4 \leq |r| < 0,7$), co obserwowane jest w przypadku par: komitet R_4 porównywany z AC, GVF, R_2 i R_3 , jak również dla pary R_2 i RG. Wartości r mogą sugerować małe różnice odpowiednio pomiędzy średnimi w kontekście otrzymanych wyników metodami średnimi a metodami skrajnymi. Prowadzi to do wniosku o małym zróżnicowaniu otrzymywanych wyników oraz wskazuje, iż przy dużym, heterogenicznym zbiorze (w porównaniu do tych stosowanych dotychczas) danych proponowane rozwiązania będą oscylować w porównywalnym zakresie wartości.

Wyniki segmentacji zmiany ogniskowej z uwzględnieniem jej typu

Typ zmiany ogniskowej wpływa na jakość jej segmentacji. Wyróżniane są zmiany echododatnie i echoujemne. Uwzględnienie charakteru morfologicznego zmiany ogniskowej w kontekście metody segmentacji znajduje odzworowanie w wynikach. Wartości metryk oceny segmentacji z podziałem na typ zmiany ogniskowej zebrano w tabelach A.18 i A.19 odpowiednio dla klasycznych metod segmentacji oraz dla komitetów klasyfikujących. W przypadku zmiany ogniskowej echoujemnej wartości metryk są znacząco wyższe zarówno na poziomie oceny klasycznych metod segmentacji, jak również komitetów. Znaczące różnice pomiędzy poszczególnymi metrykami obserwowane są również na poziomie porównania odpowiednich dla typu zmiany wyników. Ocena komitetu R_2 wskazuje wzrost miary F_1 z $0,665 \pm 0,235$ do $0,740 \pm 0,206$ w przypadku zmiany echoujemnej i spadek do wartości średniej $0,576 \pm 0,236$ dla zmiany echododatniej. Podobna zależność obserwowana jest w przypadku pozostałych metryk. Pozwala to na wnioskowanie o istotności rozważenia charakteru morfologicznego zmiany ogniskowej (poziomu intensywności i kształtu) w kontekście doboru metod, szczególnie może znaleźć to przełożenie w przypadku metod uczenia głębokiego.



Rys. 5.25: Wynik segmentacji zmiany ogniskowej echododatniej autorską metodą dla średniej wartości miary F_1 . Opisy są takie same, jak na rysunku 5.22.

5.4 Podsumowanie

Niniejszy rozdział zawiera opis części eksperymentalnej pracy. Pierwsza część rozdziału poświęcona została zaproponowanemu systemowi rejestracji danych. Omówione zostały dokładności śledzenia obiektów za pomocą systemów OTS i EMTS z uwzględnieniem dokładności lokalizacji znaczników w sce-

nie oraz wytypowano najkorzystniejszy obszar detekcji ich położenia. Następnie przedstawiono wyniki dokładności kalibracji elementów systemu rejestracji: wskaźnika, pistoletu biopsyjnego, którego kalibracja uwzględniała autorską modyfikację oraz obrazu ultrasonograficznego (przestrzenną i czasową).

W drugiej części rozdziału przedstawione zostały dane obrazowe zebrane za pomocą systemu rejestracji, jak również publiczne źródła, które zostały wykorzystane w części eksperymentalnej. W przypadku obrazów ultrasonograficznych z zapisem biopsji gruboigłowej zaproponowano unikalny zbiór danych klinicznych wraz z obrysami eksperckimi igły biopsyjnej, który to zbiór został przygotowany w ramach niniejszej rozprawy. Zbiór ten został skonstruowany w oparciu o dane zebrane za pomocą systemu rejestracji w Oddziale w Gliwicach Narodowego Instytutu Onkologii im. Marii Skłodowskiej-Curie oraz rozszerzony o upublicznione na stronie <https://www.ultrasoundcases.info/> obrazy USG, dla których to obrazów przygotowano złoty standard. Przedstawiono również opis zbiorów obrazów zmian ogniskowych, dla których częściowo również zostały przygotowane obrysy eksperckie.

Trzecią część niniejszej rozprawy poświęcono autorskim metodom klasyfikacji i segmentacji struktur w dwuwymiarowym obrazie ultrasonograficznym. Przedstawione zostały wyniki efektywności działania autorskiej architektury opartej o konwolucyjne sieci neuronowe, która realizuje dwa zadania w uzależnieniu od zastosowanych w procesie treningu danych: (i) klasyfikacja obecności igły oraz (ii) klasyfikacja obecności i typu zmiany ogniskowej. Ewaluacja proponowanej architektury uwzględnia proces doboru hiperparametrów, rozmiaru reprezentacji początkowej obrazu, testowanie mechanizmu wcześniejszego zatrzymania, jak również obejmowała wpływ zastosowania techniki dostrajania poprzedzonej augmentacją przez filtrację zbioru obrazów na rezultaty. Autorska architektura w obu przypadkach klasyfikacji została porównana z publicznymi, znanymi architektuрами klasyfikacji: AlexNet, VGG16, VGG19, SqueezeNet, GoogLeNet, ResNet101 oraz InceptionResNetV2.

Następnie przedstawiona została autorska metoda segmentacji semantycznej połączonej z transformatą Radona celem wyznaczenia trajektorii igły biopsyjnej. Wyniki eksperymentów zostały przedstawione za pomocą miar oceny segmentacji igły biopsyjnej, jak również oszacowano miary dokładności wyznaczone na wyniku segmentacji trajektorii. Wyniki autorskiej metody zostały zestawione z podejściami przedstawionymi w literaturze.

W ostatnich podrozdziałach zawarto wyniki eksperymentów za pomocą miar oceny segmentacji dla autorskiej metody segmentacji zmian ogniskowych, w której zastosowano komitet klasyfikujący, który jako dane początkowe przyjmuje wyniki segmentacji zmiany ogniskowej metodami klasycznymi: aktywnym konturem, aktywnym konturem wykorzystującym obraz gradientowy, rozrost obszaru oraz algorytm wododziałowy. Wyniki komitetu klasyfikującego zestawiono z zastosowanymi metodami klasycznymi.

6. Podsumowanie

W niniejszej rozprawie przedstawiono metodologię śledzenia i przetwarzania obrazów ultrasonograficznych gruczołu sutkowego. W części poświęconej śledzeniu omówiono sposób kalibracji systemu rejestracji danych, natomiast w ramach przetwarzania obrazów przedstawiono nowe podejście do klasyfikacji i segmentacji struktur w obrazach USG, oparte zarówno o metody klasyczne jak i o metody głębokiego przetwarzania danych obrazowych. Poprawność kolejnych etapów metodologii zweryfikowano eksperymentalnie.

Autorskie elementy multimodalnego systemu śledzenia i rejestracji w zabiegach biopsji gruczołu sutkowego obejmują:

1. opracowanie systemu rejestracji danych,
2. modyfikację metody kalibracji pistoletu biopsyjnego,
3. opracowanie metody klasyfikacji obrazu USG opartej o nową architekturę wykorzystującą konwolucyjne sieci neuronowe do:
 - klasyfikacji obecności igły biopsyjnej,
 - klasyfikacji obecności i typu zmiany ogniskowej,
4. Opracowanie metody wyznaczania trajektorii igły biopsyjnej w obrazie USG opartej o segmentację semantyczną wykorzystującą autorską architekturę konwolucyjnej sieci neuronowej,
5. Opracowanie metody segmentacji zmiany ogniskowej za pomocą komitetu klasyfikującego opartego o klasyczne metody segmentacji,
6. Przygotowanie unikalnego zbioru klinicznych obrazów ultrasonograficznych wraz z obrysami eksperckimi igły biopsyjnej,
7. Przygotowanie zbioru obrazów klinicznych zmian ogniskowych o typie echodatnim i echoujemnym,
8. Opracowanie systemu wizualizacji położenia narzędzi i obrazu USG podczas zabiegu biopsji gruczołu sutkowego.

Przestawiony powyżej zakres zrealizowanych prac oraz uzyskane wyniki eksperymentów potwierdzają postawioną na początku pracy tezę badawczą:

Wykorzystanie obrazów ultrasonograficznych zebranych przy użyciu systemu śledzenia i rejestracji pozwala na określenie trajektorii igły biopsyjnej oraz na detekcję zmiany ogniskowej w zabiegu biopsji gruboigłowej gruczołu sutkowego.

6.1 Praktyczne zastosowanie systemu

Opisywany system rejestracji danych został sprawdzony podczas 33 zabiegów biopsji gruboigłowej pacjentek onkologicznych w Zakładzie Radiologii w Narodowym Instytucie Onkologii im. Marii Skłodowskiej-Curie – Oddziale w Gliwicach. Zabiegi te przeprowadzane były w ramach projektu STRATEG-MED Narodowego Centrum Badań i Rozwoju za zgodą komisji bioetycznej Instytutu, a pacjentki każdorazowo były informowane o działającym dodatkowym systemie śledzenia i wyrażały zgodę na takie badanie.

Podczas zabiegu śledzono i rejestrowano położenie narzędzi oraz obrazu ultrasonograficznego. Klasyfikacja i segmentacja realizowane były już poza salą zabiegową, w laboratorium. System rejestracji danych wykazał przydatność jako narzędzie do archiwizacji zabiegów. Może być także wykorzystywany do obliczania parametrów zmiany ogniskowej oraz wyznaczania trajektorii igły biopsyjnej względem zmiany ogniskowej.

6.2 Plany dalszych prac

System rejestracji danych opisany w niniejszej rozprawie może zostać rozszerzony o dodatkowe elementy zarówno sprzętowe jak i programowe.

Przestawiona metodologia przetwarzania obrazów ultrasonograficznych ma charakter modułowy, co pozwala na dalsze wykorzystanie jedynie jej wybranych fragmentów. Autorskie metody klasyfikacji i segmentacji zmian ogniskowych mogą zostać użyte nie tylko do przetwarzania na żywo obrazów rejestrowanych podczas zabiegów biopsji, ale również jako wsparcie w przesiewowych badaniach USG piersi w wykrywaniu obecności zmiany ogniskowej lub do obliczania powierzchni przekroju zmiany ogniskowej na podstawie wyniku segmentacji.

Innym obszarem dalszych prac jest rozbudowa zbioru obrazów ultrasonograficznych z igłą biopsyjną o obrazy pochodzące z różnych aparatów USG zarejestrowanych przez różnych lekarzy. Zwiększenie różnorodności bazy obrazów poprawi skuteczność algorytmów detekcji. Wskazane będzie również przeprowadzenie analizy metryk wykorzystywanych do ewaluacji wyników klasyfikacji i segmentacji igły aby umożliwić obiektywne porównanie rezultatów działania różnych algorytmów.

Nowym kierunkiem rozszerzenia przedstawionego systemu jest wykorzystanie kamery *time-of-flight* oraz kamery termicznej, co pozwoliłoby na przygotowanie trójwymiarowego modelu pacjentki z rozkładem temperatury na powierzchni gruczołu. Istnieją doniesienia, że rozkład temperatury koreluje z obecnością nowotworu złośliwego piersi [95]. Rozszerzenie systemu o trójwymiarowy model pacjentki pozwoliłoby na umieszczenie wewnątrz gruczołu sutkowego zrekonstruowanego modelu guza, a także możliwe by było wskazanie trajektorii igły biopsyjnej w trójwymiarowej przestrzeni.

A. Dodatek – tabele z wynikami

Tab. A.1: Wartości p wyznaczone za pomocą testu Shapiro Wilka dla miar oceniających segmentację realizowaną za pomocą autorskiej metody (PM) oraz metod referencyjnych: Lee i wsp. oraz Hatt i wsp. Zielonym kolorem zaznaczono wyniki potwierdzające hipotezę o rozkładzie normalnym badanej cechy ($p > 0,05$), natomiast czerwonym kolorem zaznaczono wartości ($p < 0,05$), wskazujące na brak normalności badanego wektora cech .

Method	ACC_{glob}	ACC_{mean}	IoU	IoU_w	F_1
PM	0,012	0,001	0,796	<0,001	<0,001
Lee i wsp.	<0,001	0,007	0,001	<0,001	<0,001
Hatt i wsp.	<0,001	<0,001	<0,001	<0,001	<0,001

Tab. A.2: Wyniki testu Wilcoxon dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary ACC_{glob} obliczonej dla autorskiej metody segmentacji oraz metod porównawczych: Lee i wsp. oraz Hatt i wsp. Górny trójkąt tabeli prezentuje wartość p , która w przypadku $> 0,05$ zaznaczona jest na zielony kolor, w przeciwnym na czerwony kolor. W dolnym trójkącie przedstawione zostały wartości rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r). Tło komórek zostało oznaczone w zależności od wielkości efektu: mały efekt zaznaczony jest zielonym kolorem, średni żółtym oraz duży czerwonym .

ACC_{glob}	PM	Lee i wsp.	Hatt i wsp.
PM		<0,001	<0,001
Lee i wsp.	0,867		<0,001
Hatt i wsp.	0,585	-0,867	

Tab. A.3: Wyniki testu Wilcozona dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary ACC_{mean} obliczonej dla autorskiej metody segmentacji oraz metod porównawczych: Lee i wsp. oraz Hatt i wsp. Oznaczenia tabeli zgodne z tabelą A.2.

ACC_{mean}	PM	Lee i wsp.	Hatt i wsp.
PM		<0,001	<0,001
Lee i wsp.	0,298		<0,001
Hatt i wsp.	0,867	0,867	

Tab. A.4: Wyniki testu Wilcozona dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary IoU obliczonej dla autorskiej architektury, metod Lee i wsp. oraz Hatt i wsp. Oznaczenia tabeli zgodne z tabelą A.2.

IoU	PM	Lee i wsp.	Hatt i wsp.
PM		<0,001	<0,001
Lee i wsp.	0,867		<0,001
Hatt i wsp.	0,867	-0,867	

Tab. A.5: Wyniki testu Wilcozona dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary IoU_w obliczonej dla autorskiej architektury, metod Lee i wsp. oraz Hatt i wsp. Oznaczenia tabeli zgodne z tabelą A.2.

IoU_w	PM	Lee i wsp.	Hatt i wsp.
PM		<0,001	<0,001
Lee i wsp.	0,867		<0,001
Hatt i wsp.	0,812	-0,867	

Tab. A.6: Wyniki testu Wilcozona dla par obserwacji oraz rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) wyznaczonych dla wartości miary F_1 wyznaczonej dla autorskiej architektury, metod Lee i wsp. oraz Hatt i wsp. Oznaczenia tabeli zgodne z tabelą A.2.

F_1	PM	Lee i wsp.	Hatt i wsp.
PM		<0,001	<0,001
Lee i wsp.	0,867		<0,001
Hatt i wsp.	0,850	-0,867	

Tab. A.7: Wyniki testu Shapiro Wilka (p -wartości) parametrów λ oraz θ oraz błędów trajektorii igły $\Delta\lambda$ i $\Delta\theta$ wyznaczone dla złotego standardu (GT), autorskiej metody detekcji igły (PM) oraz metod porównawczych: Lee i wsp. oraz Hatt i wsp. Czerwonym kolorem zaznaczono p -wartości, które $< 0,05$.

Metoda	λ	θ	$\Delta\lambda$	$\Delta\theta$
GT	0,011	0,029	–	–
PM	0,032	0,104	<0,001	<0,001
Lee i wsp.	0,010	<0,001	<0,001	<0,001
Hatt i wsp.	0,011	<0,001	<0,001	<0,001

Tab. A.8: Wyniki testu Wilcoxon dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla parametru λ określonego dla trajektorii igły wyznaczonej za pomocą złotego standardu (GT) oraz autorskiej metody detekcji igły oraz metod porównawczych: Lee i wsp. oraz Hatt i wsp. Parametr $\lambda_R = \lambda_{GT}$ dla złotego standardu. Oznaczenia tabeli zgodne z tabelą A.2.

λ_R	GT	PM	Lee i wsp.	Hatt i wsp.
GT		0,210	0,001	<0,001
PM	–		<0,001	<0,001
Lee i wsp.	-0,236	-0,262		<0,001
Hatt i wsp.	-0,531	-0,501	-0,268	

Tab. A.9: Wyniki testu Wilcoxon dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla parametru θ określonego dla trajektorii igły wyznaczonej za pomocą złotego standardu (GT) oraz autorskiej metody detekcji igły oraz metod porównawczych: Lee i wsp. oraz Hatt i wsp. Parametr $\theta_R = \theta_{GT}$ dla złotego standardu. Oznaczenia tabeli zgodne z tabelą A.2.

θ_R	GT	PM	Lee i wsp.	Hatt i wsp.
GT		0,780	0,001	<0,001
PM	–		<0,001	<0,001
Lee i wsp.	0,230	0,283		0,007
Hatt i wsp.	0,398	0,402	0,182	

Tab. A.10: Wyniki testu Wilcoxon dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla błędu dystansu detekcji trajektorii $\Delta\lambda$ określonego dla trajektorii igły wyznaczonej za pomocą autorskiej metody detekcji igły oraz metod porównawczych: Lee i wsp. oraz Hatt i wsp. Oznaczenia tabeli zgodne z tabelą A.2.

$\Delta\lambda_R$	PM	Lee i wsp.	Hatt i wsp.
PM		<0,001	<0,001
Lee i wsp.	-0,664		<0,001
Hatt i wsp.	-0,773	-0,399	

Tab. A.11: Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla błędu kąta przy detekcji trajektorii $\Delta\theta_R$ określonego dla trajektorii igły wyznaczonej za pomocą autorskiej metody detekcji igły oraz metod porównawczych: Lee i wsp. oraz Hatt i wsp. Oznaczenia tabeli zgodne z tabelą A.2.

$\Delta\theta$	PM	Lee i wsp.	Hatt i wsp.
PM		<0,001	<0,001
Lee i wsp.	-0,642		<0,001
Hatt i wsp.	-0,797	-0,372	

Tab. A.12: Wartości p wyznaczone za pomocą testu Shapiro Wilka dla miar segmentacji zmian ogniskowych niezależnie od jej typu. Zielonym kolorem zaznaczono wyniki potwierdzające hipotezę o rozkładzie normalnym badanej cechy ($p > 0,05$), natomiast czerwonym kolorem zaznaczono wartości ($p < 0,05$), wskazujące na brak normalności badanego wektora cech.

Metoda	ACC	TPR	TNR	IoU	F_1
AC	<0,001	<0,001	<0,001	<0,001	<0,001
GVF	<0,001	<0,001	<0,001	<0,001	<0,001
RG	<0,001	<0,001	<0,001	<0,001	<0,001
WS	<0,001	<0,001	<0,001	<0,001	<0,001
R ₁	<0,001	<0,001	<0,001	<0,001	<0,001
R ₂	<0,001	<0,001	<0,001	<0,001	<0,001
R ₃	<0,001	<0,001	<0,001	<0,001	<0,001
R ₄	<0,001	<0,001	<0,001	<0,001	<0,001

Tab. A.13: Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla dokładności segmentacji zmian ogniskowych niezależnie od jej typu. Oznaczenia tabeli zgodne z tabelą A.2.

ACC	AC	GVF	RG	WS	R ₁	R ₂	R ₃	R ₄
AC		<0,001	<0,001	<0,001	<0,001	<0,001	<0,001	<0,001
GVF	-0,237		<0,001	<0,001	<0,001	<0,001	<0,001	<0,001
RG	0,433	0,434		<0,001	<0,001	<0,001	<0,001	<0,001
WS	0,274	0,291	-0,207		<0,001	<0,001	<0,001	<0,001
R ₁	0,453	0,457	0,250	0,351		<0,001	<0,001	<0,001
R ₂	-0,162	-0,120	-0,492	-0,409	0,409		<0,001	<0,001
R ₃	-0,152	0,089	-0,442	-0,281	0,499	-0,112		<0,001
R ₄	0,377	0,366	-0,301	-0,141	0,112	0,448	-0,448	

Tab. A.14: Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla współczynnika wyników prawdziwie pozytywnych segmentacji zmian ogniskowych niezależnie od jej typu. Oznaczenia tabeli zgodne z tabelą A.2.

<i>TPR</i>	AC	GVF	RG	WS	R ₁	R ₂	R ₃	R ₄
AC		<0,001	<0,001	<0,001	<0,001	<0,001	<0,001	<0,001
GVF	-0,541		0,078	<0,001	<0,001	<0,001	<0,001	<0,001
RG	-0,153	–		<0,001	<0,001	<0,001	<0,001	<0,001
WS	-0,363	-0,226	-0,133		<0,001	0,004	<0,001	<0,001
R ₁	-0,612	-0,604	-0,603	-0,601		0,004	<0,001	<0,001
R ₂	-0,605	-0,297	-0,182	0,056	-0,056		<0,001	<0,001
R ₃	0,101	0,549	0,191	0,417	-0,610	-0,611		<0,001
R ₄	0,610	0,611	0,583	0,611	0,611	0,610	-0,610	

Tab. A.15: Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla współczynnika wyników prawdziwie negatywnych segmentacji zmian ogniskowych niezależnie od jej typu. Oznaczenia tabeli zgodne z tabelą A.2.

<i>TNR</i>	AC	GVF	RG	WS	R ₁	R ₂	R ₃	R ₄
AC		<0,001	<0,001	<0,001	<0,001	<0,001	<0,001	<0,001
GVF	0,537		<0,001	<0,001	<0,001	<0,001	<0,001	<0,001
RG	0,437	0,387		<0,001	<0,001	<0,001	<0,001	<0,001
WS	0,523	0,463	-0,155		<0,001	<0,001	<0,001	<0,001
R ₁	0,610	0,599	0,603	0,595		<0,001	<0,001	<0,001
R ₂	0,544	0,146	-0,380	-0,479	0,479		<0,001	<0,001
R ₃	-0,240	-0,547	-0,469	-0,564	0,610	0,579		<0,001
R ₄	-0,526	-0,562	-0,527	-0,597	-0,579	-0,517	0,517	

Tab. A.16: Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla miary *IoU* segmentacji zmian ogniskowych niezależnie od jej typu. Oznaczenia tabeli zgodne z tabelą A.2.

<i>IoU</i>	AC	GVF	RG	WS	R ₁	R ₂	R ₃	R ₄
AC		<0,001	<0,001	0,003	<0,001	<0,001	0,089	<0,001
GVF	-0,300		<0,001	<0,001	<0,001	<0,001	<0,001	<0,001
RG	0,296	0,321		<0,001	<0,001	<0,001	<0,001	0,003
WS	0,057	0,112	-0,221		<0,001	<0,001	0,024	<0,001
R ₁	0,151	0,197	-0,188	0,095		<0,001	<0,001	<0,001
R ₂	-0,371	-0,255	-0,478	-0,300	0,300		<0,001	<0,001
R ₃	–	0,231	-0,294	-0,043	0,335	-0,355		<0,001
R ₄	0,516	0,513	0,098	0,285	0,355	0,551	-0,551	

Tab. A.17: Wyniki testu Wilcoxona dla par obserwacji i rangowego współczynnika korelacji dwuseryjnej dla par dopasowanych (r) przeprowadzonych dla miary F_1 segmentacji zmian ogniskowych niezależnie od jej typu. Oznaczenia tabeli zgodne z tabelą A.2.

F_1	AC	GVF	RG	WS	R_1	R_2	R_3	R_4
AC		<0,001	<0,001	0,043	<0,001	<0,001	0,043	<0,001
GVF	-0,335		<0,001	<0,001	<0,001	<0,001	<0,001	<0,001
RG	0,287	0,316		<0,001	<0,001	<0,001	<0,001	0,043
WS	0,039	0,097	-0,238		<0,001	<0,001	0,211	<0,001
R_1	0,142	0,189	-0,200	0,103		<0,001	<0,001	<0,001
R_2	-0,373	-0,250	-0,479	-0,297	0,297		<0,001	<0,001
R_3	-0,039	0,255	-0,283	–	0,334	-0,362		<0,001
R_4	0,523	0,527	0,101	0,305	0,362	0,559	-0,559	

Tab. A.18: Podsumowanie statystyk dla wyników segmentacji regionu odpowiednio zmiany ogniskowej echoujemnej i echododatniej klasycznymi metodami segmentacji.

stat.	zmiana echoujemna					zmiana echododatnia				
	<i>ACC</i>	<i>TPR</i>	<i>TNR</i>	<i>IoU</i>	<i>F₁</i>	<i>ACC</i>	<i>TPR</i>	<i>TNR</i>	<i>IoU</i>	<i>F₁</i>
aktywny kontur – AC										
M	0,95	0,63	0,99	0,57	0,69	0,90	0,41	0,99	0,36	0,48
Me	0,98	0,68	1,00	0,61	0,76	0,94	0,38	1,00	0,33	0,50
SD	0,06	0,26	0,01	0,24	0,23	0,12	0,29	0,02	0,25	0,27
min	0,57	0,00	0,92	0,00	0,00	0,35	0,00	0,79	0,00	0,00
max	1,00	1,00	1,00	0,95	0,97	1,00	1,00	1,00	0,94	0,97
aktywny kontur oparty o obraz gradientowy – GVF										
M	0,95	0,69	0,99	0,59	0,71	0,90	0,46	0,99	0,38	0,50
Me	0,98	0,76	1,00	0,65	0,79	0,94	0,43	1,00	0,36	0,53
SD	0,06	0,27	0,01	0,23	0,22	0,12	0,30	0,02	0,24	0,27
min	0,58	0,00	0,89	0,00	0,00	0,36	0,00	0,78	0,00	0,00
max	1,00	1,00	1,00	0,95	0,97	1,00	1,00	1,00	0,88	0,94
rozrost obszaru – RG										
M	0,86	0,63	0,88	0,43	0,54	0,77	0,60	0,80	0,28	0,40
Me	0,95	0,67	1,00	0,44	0,61	0,81	0,64	0,86	0,22	0,36
SD	0,17	0,26	0,18	0,28	0,29	0,19	0,30	0,22	0,23	0,26
min	0,24	0,00	0,18	0,00	0,00	0,09	0,00	0,02	0,00	0,00
max	1,00	1,00	1,00	0,97	0,98	1,00	1,00	1,00	0,93	0,96
algorytm wododziałowy – WS										
M	0,92	0,73	0,93	0,51	0,64	0,85	0,60	0,91	0,37	0,51
Me	0,95	0,80	0,98	0,52	0,68	0,90	0,60	0,96	0,36	0,53
SD	0,11	0,24	0,12	0,24	0,24	0,14	0,26	0,14	0,20	0,22
min	0,25	0,00	0,23	0,00	0,00	0,04	0,00	0,00	0,00	0,00
max	1,00	1,00	1,00	0,97	0,98	1,00	1,00	1,00	0,93	0,97

Tab. A.19: Podsumowanie statystyk dla wyników segmentacji regionu zmiany ogni-skowej echoujemnej i echodatniej metodą komitetu klasyfikującego na podstawie wyników klasycznych metod segmentacji.

stat.	zmiana echoujemna					zmiana echodatnia				
	<i>ACC</i>	<i>TPR</i>	<i>TNR</i>	<i>IoU</i>	<i>F₁</i>	<i>ACC</i>	<i>TPR</i>	<i>TNR</i>	<i>IoU</i>	<i>F₁</i>
jedna metoda – R ₁										
M	0,84	0,89	0,83	0,47	0,59	0,74	0,83	0,73	0,33	0,46
Me	0,93	0,95	0,93	0,48	0,64	0,76	0,91	0,77	0,28	0,44
SD	0,19	0,15	0,21	0,28	0,28	0,19	0,20	0,22	0,22	0,23
min	0,18	0,00	0,18	0,00	0,00	0,04	0,00	0,00	0,00	0,00
max	1,00	1,00	1,00	0,96	0,98	1,00	1,00	1,00	0,87	0,93
≥ dwie metody – R ₂										
M	0,95	0,73	0,98	0,62	0,74	0,90	0,58	0,97	0,44	0,58
Me	0,98	0,79	1,00	0,68	0,81	0,94	0,60	0,99	0,44	0,61
SD	0,07	0,22	0,05	0,22	0,21	0,11	0,28	0,06	0,23	0,24
min	0,52	0,00	0,54	0,00	0,00	0,41	0,00	0,43	0,00	0,00
max	1,00	1,00	1,00	0,96	0,98	1,00	1,00	1,00	0,94	0,97
≥ trzy metody – R ₃										
M	0,95	0,61	0,99	0,56	0,69	0,90	0,40	0,99	0,36	0,47
Me	0,98	0,66	1,00	0,61	0,76	0,94	0,37	1,00	0,33	0,50
SD	0,06	0,26	0,01	0,24	0,24	0,12	0,29	0,01	0,25	0,28
min	0,57	0,00	0,91	0,00	0,00	0,32	0,00	0,79	0,00	0,00
max	1,00	1,00	1,00	0,94	0,97	1,00	1,00	1,00	0,92	0,96
cztery metody – R ₄										
M	0,94	0,44	1,00	0,43	0,55	0,89	0,25	1,00	0,24	0,33
Me	0,97	0,45	1,00	0,44	0,61	0,94	0,17	1,00	0,16	0,28
SD	0,07	0,27	0,01	0,26	0,28	0,13	0,24	0,01	0,23	0,28
min	0,56	0,00	0,93	0,00	0,00	0,31	0,00	0,90	0,00	0,00
max	1,00	0,97	1,00	0,96	0,98	1,00	0,96	1,00	0,87	0,93

Bibliografia

- [1] N. Agarwal, A. K. Yadav, A. Gupta i M. F. Orlando. „Real-time needle tip localization in 2D ultrasound images using Kalman filter”. W: *2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. 2019, s. 1008–1012. DOI: 10.1109/AIM.2019.8868799 (cytowane na stronach 25, 26, 27, 30 i 83).
- [2] D. Alsbeih, M. I. Douad, A.-K. Al-Tamimi i M. A. Al-Jarrah. „A dynamic system for tracking biopsy needle in two dimensional ultrasound images”. W: *2020 IEEE 5th Middle East and Africa Conference on Biomedical Engineering (MECBME)*. 2020, s. 1–4. DOI: 10.1109/MECBME47393.2020.9265166 (cytowane na stronach 25, 26 i 28).
- [3] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami i M. K. Khan. „Medical image analysis using convolutional neural networks: a review”. W: *Journal of Medical Systems* 42.11 (paź. 2018). DOI: 10.1007/s10916-018-1088-1 (cytowane na stronie 59).
- [4] M. Arif, A. Moelker i T. van Walsum. „Automatic needle detection and real-time Bi-planar needle visualization during 3D ultrasound scanning of the liver”. W: *Medical Image Analysis* 53 (2019), s. 104–110. ISSN: 1361-8415. DOI: 10.1016/j.media.2019.02.002 (cytowane na stronach 23, 26, 28, 29 i 31).
- [5] G. Ayana, K. Dese i S.-w. Choe. „Transfer learning in breast cancer diagnoses via ultrasound imaging”. W: *Cancers* 13.4 (2021). ISSN: 2072-6694. DOI: 10.3390/cancers13040738 (cytowane na stronach 21 i 60).
- [6] N. Azamjah, Y. Soltan-Zadeh i F. Zayeri. „Global trend of breast cancer mortality rate: a 25-year study”. W: *Asian Pacific Journal of Cancer Prevention* 20.7 (2019), s. 2015–2020. DOI: 10.31557/APJCP.2019.20.7.2015 (cytowane na stronie 2).
- [7] V. Badrinarayanan, A. Kendall i R. Cipolla. „SegNet: a deep convolutional encoder-decoder architecture for image segmentation”. W: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.12 (2017), s. 2481–2495. DOI: 10.1109/TPAMI.2016.2644615 (cytowane na stronie 83).

- [8] R. S. Bandaru, A. R. Sornes, J. Hermans, E. Samset i J. D'hooge. „Delay and standard deviation beamforming to enhance specular reflections in ultrasound imaging”. W: *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control* 63.12 (2016), s. 2057–2068. DOI: 10.1109/TUFFC.2016.2613963 (cytowane na stronach 5 i 22).
- [9] F. Banterle, M. Corsini, P. Cignoni i R. Scopigno. „A low-memory, straightforward and fast bilateral filter through subsampling in spatial domain”. W: *Computer Graphics Forum* 31.1 (2012), s. 19–32. DOI: <https://doi.org/10.1111/j.1467-8659.2011.02078.x> (cytowane na stronach 90 i 91).
- [10] P. Beigi, R. Rohling i G. Salcudean S.and Ng. „CASPER: computer-aided segmentation of imperceptible motion-a learning-based tracking of an invisible needle in ultrasound”. W: *International journal of computer assisted radiology and surgery* 12.11 (list. 2017), s. 1857–1866. ISSN: 1861-6410. DOI: 10.1007/s11548-017-1631-4 (cytowane na stronach 26, 27 i 30).
- [11] P. Beigi, R. Rohling, T. Salcudean, V. A. Lessoway i G. C. Ng. „Detection of an invisible needle in ultrasound using a probabilistic SVM and time-domain features”. W: *Ultrasonics* 78 (2017), s. 18–22. ISSN: 0041-624X. DOI: <https://doi.org/10.1016/j.ultras.2017.02.010> (cytowane na stronach 28 i 31).
- [12] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006. ISBN: 978-0387-31073-2 (cytowane na stronach 59, 62, 68, 73, 78, 80, 82 i 86).
- [13] L. Brattain, B. Telfer, M. Dhyani, J. Grajo i A. Samir. „Machine learning for medical ultrasound: status, methods, and future opportunities”. W: *Abdominal Radiology* 43.4 (2018), s. 786–799. DOI: doi : 10.1007/s00261-018-1517-0 (cytowane na stronie 5).
- [14] Breast Cancer Facts & Figures 2019-2020. *American Cancer Society*. Atlanta, USA: American Cancer Society, Inc., 2019. ISBN: 9781417642595 (cytowane na stronach 1 i 12).
- [15] R. F. Brem, M. J. Lenihan, J. Lieberman i J. Torrente. „Screening breast ultrasound: past, present, and future”. W: *American Journal of Roentgenology* 204.2 (2015). PMID: 25615743, s. 234–240. DOI: 10.2214/AJR.13.12072 (cytowane na stronach 12 i 13).
- [16] A. Buades, B. Coll i J.-M. Morel. „A non-local algorithm for image denoising”. W: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. T. 2. 2005, 60–65 vol. 2. DOI: 10.1109/CVPR.2005.38 (cytowane na stronach 76 i 90).

- [17] M. Byra, P. Jarosik, A. Szubert, M. Galperin, H. Ojeda-Fournier, L. Olson, M. O'Boyle, C. Comstock i M. Andre. „Breast mass segmentation in ultrasound with selective kernel U-Net convolutional neural network”. W: *Biomedical Signal Processing and Control* 61 (2020), s. 102027. ISSN: 1746-8094. DOI: <https://doi.org/10.1016/j.bspc.2020.102027> (cytowane na stronach 14, 15, 18, 19, 20 i 124).
- [18] M. Byra. „Discriminant analysis of neural style representations for breast lesion classification in ultrasound”. W: *Biocybernetics and Biomedical Engineering* 38.3 (2018), s. 684–690. ISSN: 0208-5216. DOI: <https://doi.org/10.1016/j.bbe.2018.05.003> (cytowane na stronach 21, 23, 24 i 60).
- [19] G. Carbajal, A. Lasso, Á. Gómez i G. Fichtinger. „Improving N-wire phantom-based freehand ultrasound calibration”. W: *International Journal of Computer Assisted Radiology and Surgery* 8 (2013), s. 1063–1072. DOI: 10.1007/s11548-013-0904-9 (cytowane na stronach 51, 52, 54, 116 i 118).
- [20] R.-F. Chang, W.-J. Wu, W. K. Moon, W.-M. Chen, W. Lee i D.-R. Chen. „Segmentation of breast tumor in three-dimensional ultrasound images using three-dimensional discrete active contour model”. W: *Ultrasound in Medicine and Biology* 29.11 (2003), s. 1571–1581. DOI: 10.1016/S0301-5629(03)00992-X (cytowane na stronach 17, 18 i 19).
- [21] J. Cho, T. Jensen, K. Reiersen, B. Mathews, A. Bhagra, R. Franco-Sadud, L. Grikis, M. Mader, R. Dancel, B. Lucas, N. Soni, S. Abdel-Ghani, R. Arntfield, J. Bates, M. Blaiwas, D. Brotman, C. Candotti, R. Hoppmann, S. Hunt, V. Kalidindi, K. Kobaidze, J. Lenchus, P. Mayo, S. Nichani, V. Noble, M. Perez, N. Puri, A. Pustavoitau, S. Rodgers, G. Salame, D. Schnobrich, K. Spencer, V. Tayal i D. Tierney. „Recommendations on the use of ultrasound guidance for adult abdominal paracentesis: A position statement of the society of hospital medicine”. W: *Journal of hospital medicine (Online)* 14 (2019), E7–E15. ISSN: 1553-5606. DOI: 10.12788/jhm.3095 (cytowane na stronie 22).
- [22] V. Corsetti, N. Houssami, M. Ghirardi, A. Ferrari, M. Speziani, S. Bellarosa, G. Remida, C. Gasparotti, E. Galligioni i S. Ciatto. „Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: Interval breast cancers at 1year follow-up”. W: *European Journal of Cancer* 47.7 (2011), s. 1021–1026. ISSN: 0959-8049. DOI: <https://doi.org/10.1016/j.ejca.2010.12.002> (cytowane na stronach 5, 12 i 14).

- [23] C. P. Crum, S. C. Lester i R. S. Cotran. „Żeński układ płciowy i pierś”. W: V. Kumar, R. Cotran i S. L. Robbins. *Robbins Patologia*. Wrocław: Wydawnictwo Medyczne Urban & Partner, 2005. Rozd. 19, s. 779–822. ISBN: 0-7216-9274-5 (cytowane na stronie 2).
- [24] J. Czajkowska, B. Pyciński i E. Pietka. „HoG feature based detection of tissue deformations in ultrasound data”. W: *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2015, s. 6326–6329. DOI: 10.1109/EMBC.2015.7319839 (cytowane na stronach 27, 28 i 30).
- [25] J. Czajkowska, B. Pyciński, J. Juszczak i E. Pietka. „Biopsy needle tracking technique in US images”. W: *Computerized Medical Imaging and Graphics* 65 (2018), s. 93–101. ISSN: 0895-6111. DOI: 10.1016/j.compmedimag.2017.07.001 (cytowane na stronach 25, 26, 27, 30, 50 i 75).
- [26] R. Dancel, D. Schnobrich, N. Puri, R. Franco-Sadud, J. Cho, L. Grikis, B. P. Lucas, M. El-Barbary i N. J. Soni. „Recommendations on the use of ultrasound guidance for adult thoracentesis: a position statement of the society of hospital medicine”. W: *Journal of Hospital Medicine* 13.2 (lut. 2018), s. 126–135. DOI: 10.12788/jhm.2940 (cytowane na stronach 13 i 22).
- [27] M. I. Daoud, S. Abdel-Rahman i R. Alazrai. „Breast ultrasound image classification using a pre-trained convolutional neural network”. W: *2019 15th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*. 2019, s. 167–171. DOI: 10.1109/SITIS.2019.00037 (cytowane na stronie 21).
- [28] M. I. Daoud, S. Abdel-Rahman, T. M. Bdair, M. S. Al-Najar, F. H. Al-Hawari i R. Alazrai. „Breast tumor classification in ultrasound images using combined deep and handcrafted features”. W: *Sensors* 20.23 (2020). ISSN: 1424-8220. DOI: 10.3390/s20236838 (cytowane na stronie 21).
- [29] M. I. Daoud, A. F. Abu-Hani i R. Alazrai. „Reliable and accurate needle localization in curvilinear ultrasound images using signature-based analysis of ultrasound beamformed radio frequency signals”. W: *Medical Physics* 47.6 (2020), s. 2356–2379. DOI: <https://doi.org/10.1002/mp.14126> (cytowane na stronach 26, 28 i 87).
- [30] M. I. Daoud, T. M. Bdair, M. Al-Najar i R. Alazrai. „A fusion-based approach for breast ultrasound image classification using multiple-ROI texture and morphological analyses”. W: *Computational and Mathematical Methods in Medicine* 2016 (2016). DOI: 10.1155/2016/6740956 (cytowane na stronach 20, 21, 23 i 24).

- [31] M. I. Daoud, S. Khraiweh, A. Zayadeen i R. Alazrai. „Accurate needle localization in two-dimensional ultrasound images”. W: *2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*. 2017, s. 578–582 (cytowane na stronach 22, 25 i 28).
- [32] M. I. Daoud, A. Shtaiyat i R. Alazrai. „Enhanced needle detection in ultrasound images using acoustic excitation and ultrasound image analyses”. W: *2018 11th Biomedical Engineering International Conference (BMEiCON)*. 2018, s. 1–5. DOI: 10.1109/BMEiCON.2018.8609920 (cytowane na stronach 25, 26 i 28).
- [33] C. Desantis, J. Ma, M. M. Gaudet, L. A. Newman, K. D. Miller, A. Goding Sauer, A. Jemal i R. L. Siegel. „Breast cancer statistics, 2019”. W: *CA: A Cancer Journal for Clinicians* 69 (paź. 2019). DOI: 10.3322/caac.21583 (cytowane na stronie 2).
- [34] W. Al-Dhabyani, M. Gomaa, H. Khaled i A. Fahmy. „Dataset of breast ultrasound images”. W: *Data in Brief* 28 (2020), s. 104863. ISSN: 2352-3409. DOI: <https://doi.org/10.1016/j.dib.2019.104863> (cytowane na stronach 20, 21, 123 i 124).
- [35] D. DiGiacinto, J. Bagley i A. M. Goldsbury. „The value of sonography in the assessment of skin cancers and their metastases”. W: *Journal of Diagnostic Medical Sonography* 32.3 (2016), s. 140–146. DOI: 10.1177/8756479316643959 (cytowane na stronach 4 i 120).
- [36] C. A. Duarte-Salazar, A. E. Castro-Ospina, M. A. Becerra i E. Delgado-Trejos. „Speckle noise reduction in ultrasound images for improving the metrological evaluation of biomedical applications: an overview”. W: *IEEE Access* 8 (2020), s. 15983–15999. DOI: 10.1109/ACCESS.2020.2967178 (cytowane na stronach 5, 13, 18, 88 i 99).
- [37] R. Elfring, M. de la Fuente i K. Radermacher. „Assessment of optical localizer accuracy for computer aided surgery systems”. W: *Computer Aided Surgery* 15.1-3 (2010). PMID: 20233129, s. 1–12. DOI: 10.3109/10929081003647239 (cytowane na stronie 113).
- [38] Euklides. *Elements of Geometry. The Greek text of J. L. Heiberg (1883–1885) from „Euclidis Elementa, edidit et Latine interpretatus est I. L. Heiberg, in aedibus B. G. Teubneri, 1883-1885”, edited, and provided with a modern English translation, by Richard Fitzpatrick*. Księga 1, Postulat 5. ISBN: 978-0-615-17984-1 (cytowane na stronie 122).
- [39] A. Fattahi, A. Tavassoli, M. Kalantari, S. Noorshafiee i M. Rahmani. „Evaluation of the value of core needle biopsy in the diagnosis of a breast mass”. W: *Archives of Breast Cancer* 3.2 (2016), s. 56–61. DOI: 10.19187/abc.20163256-61 (cytowane na stronach 7, 22 i 25).

- [40] T. Fawcett. „An introduction to ROC analysis”. W: *Pattern Recognition Letters* 27.8 (2006). ROC analysis in pattern recognition, s. 861–874. ISSN: 0167-8655. DOI: <https://doi.org/10.1016/j.patrec.2005.10.010> (cytowane na stronie 104).
- [41] J. Ferlay, M. Colombet, I. Soerjomataram, C. Mathers, D. Parkin, M. Piñeros, A. Znaor i F. Bray. „Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods”. W: *International Journal of Cancer* 144.8 (2019), s. 1941–1953. DOI: <https://doi.org/10.1002/ijc.31937> (cytowane na stronie 11).
- [42] A. M. Franz, T. Haidegger, W. Birkfellner, K. Cleary, T. M. Peters i L. Maier-Hein. „Electromagnetic tracking in medicine—A review of technology, validation, and applications”. W: *IEEE Transactions on Medical Imaging* 33.8 (2014), s. 1702–1725. DOI: [10.1109/TMI.2014.2321777](https://doi.org/10.1109/TMI.2014.2321777) (cytowane na stronach 7 i 25).
- [43] M. Galińska, W. Ogiegło, A. Wijata, J. Juszczak i J. Czajkowska. „Breast cancer segmentation method in ultrasound images”. W: *Innovations in Biomedical Engineering*. Red. M. Gzik, E. Tkacz, Z. Paszenda i E. Piętka. Cham: Springer International Publishing, 2018, s. 23–31. ISBN: 978-3-319-70063-2. DOI: [10.1007/978-3-319-70063-2_3](https://doi.org/10.1007/978-3-319-70063-2_3) (cytowane na stronach 14, 16, 17, 18, 20, 88 i 156).
- [44] A. Garg i V. Khandelwal. „Speckle noise reduction in medical ultrasound images using coefficient of dispersion”. W: *2016 International Conference on Signal Processing and Communication (ICSC)*. 2016, s. 208–212. DOI: [10.1109/ICSPCom.2016.7980577](https://doi.org/10.1109/ICSPCom.2016.7980577) (cytowane na stronach 13 i 19).
- [45] X. Glorot i Y. Bengio. „Understanding the difficulty of training deep feedforward neural networks”. W: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Red. Y. W. Teh i M. Titterton. T. 9. Proceedings of Machine Learning Research. Chia Laguna Resort, Sardinia, Italy: PMLR, maj 2010, s. 249–256 (cytowane na stronach 59, 62, 64, 67, 72, 78, 82, 84, 85 i 86).
- [46] Główny Urząd Statystyczny. *Rocznik Demograficzny*. Warszawa: Zakład Wydawnictw Statystycznych, 2020 (cytowane na stronie 1).
- [47] W. Gomez-Flores i W. Coelho de Albuquerque Pereira. „A comparative study of pre-trained convolutional neural networks for semantic segmentation of breast tumors in ultrasound”. W: *Computers in Biology and Medicine* 126 (2020), s. 104036. ISSN: 0010-4825. DOI: [10.1016/j.compbiomed.2020.104036](https://doi.org/10.1016/j.compbiomed.2020.104036) (cytowane na stronach 13, 14, 15, 18, 19, 20 i 124).

- [48] I. Goodfellow, Y. Bengio i A. Courville. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016 (cytowane na stronach 59, 62, 63, 64, 68, 70 i 71).
- [49] K. Gremban, C. Thorpe i T. Kanade. „Geometric camera calibration using systems of linear equations”. W: *Proceedings. 1988 IEEE International Conference on Robotics and Automation*. T. 1. 1988, s. 562–567. DOI: 10.1109/ROBOT.1988.12111 (cytowane na stronie 33).
- [50] L. A. Groves, B. VanBerlo, T. M. Peters i E. C. S. Chen. „Deep learning approach for automatic out-of-plane needle localisation for semi-automatic ultrasound probe calibration”. W: *Healthcare Technology Letters* 6.6 (2019), s. 204–209. DOI: 10.1049/htl.2019.0075 (cytowane na stronach 25, 26, 27, 29, 30 i 108).
- [51] P. Gu, W.-M. Lee, M. A. Roubidoux, J. Yuan, X. Wang i P. L. Carson. „Automated 3D ultrasound image segmentation to aid breast cancer image interpretation”. W: *Ultrasonics* 65 (2016), s. 51–58. ISSN: 0041-624X. DOI: <https://doi.org/10.1016/j.ultras.2015.10.023> (cytowane na stronach 17, 18 i 19).
- [52] A. Gulland. „Shortage of health workers is set to double, says WHO”. W: *BMJ* 347 (2013). DOI: 10.1136/bmj.f6804 (cytowane na stronie 3).
- [53] I. Hacihaliloglu, P. Beigi, G. Ng, R. N. Rohling, S. Salcudean i P. Abolmaesumi. „Projection-based phase features for localization of a needle tip in 2D curvilinear ultrasound”. W: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Red. N. Navab, J. Hornegger, W. M. Wells i A. Frangi. Cham: Springer International Publishing, 2015, s. 347–354. ISBN: 978-3-319-24553-9 (cytowane na stronie 28).
- [54] J. A. Hanley. „Receiver operating characteristic (ROC) methodology: the state of the art”. W: *Critical reviews in diagnostic imaging* 29.3 (1989), s. 307–35 (cytowane na stronie 104).
- [55] C. Hatt, G. Ng i V. Parthasarathy. „Enhanced needle localization in ultrasound using beam steering and learning-based segmentation”. W: *Computerized Medical Imaging and Graphics* 41 (2015), s. 46–54. ISSN: 0895-6111. DOI: 10.1016/j.compmedimag.2014.06.016 (cytowane na stronach 23, 25, 26, 27, 28, 29, 30, 87, 107, 108, 146, 149 i 150).
- [56] K. He, X. Zhang, S. Ren i J. Sun. „Deep residual learning for image recognition”. W: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, s. 770–778. DOI: 10.1109/CVPR.2016.90 (cytowane na stronach 60, 61, 62, 139 i 144).

- [57] K. He, X. Zhang, S. Ren i J. Sun. *Delving deep into rectifiers: surpassing human-level performance on ImageNet classification*. 2015. DOI: 10.1109/ICCV.2015.123. arXiv: 1502.01852 [cs.CV] (cytowane na stronach 59, 62, 67, 69, 78, 83 i 86).
- [58] W. R. Hendee i E. R. Ritenour. *Medical Imaging Physics, 4th Edition*. Wiley-Liss, 2003. ISBN: 978-0-471-46113-5 (cytowane na stronach 4, 5, 6 i 52).
- [59] P. S. Hiremath, P. T. Akkasaligar i S. Badiger. „Speckle noise reduction in medical ultrasound images”. W: 2013. DOI: 10.5772/56519 (cytowane na stronach 13 i 88).
- [60] Y. Ho i S. Wookey. „The real-world-weight cross-entropy loss function: modeling the costs of mislabeling”. W: *IEEE Access* 8 (2020), s. 4806–4813. ISSN: 2169-3536. DOI: 10.1109/access.2019.2962617 (cytowane na stronach 68 i 156).
- [61] H. H. Holm i B. Skjoldbye. „Interventional ultrasound”. W: *Ultrasound in Medicine & Biology* 22.7 (1996), s. 773–789. ISSN: 0301-5629. DOI: [https://doi.org/10.1016/0301-5629\(96\)00086-5](https://doi.org/10.1016/0301-5629(96)00086-5) (cytowane na stronie 22).
- [62] B. K. P. Horn. „Closed-form solution of absolute orientation using unit quaternions”. W: *J. Opt. Soc. Am. A* 4.4 (kw. 1987), s. 629–642. DOI: 10.1364/JOSAA.4.000629 (cytowane na stronie 55).
- [63] P. W. Hsu, R. W. Prager, A. H. Gee i G. M. Treece. „Freehand 3D ultrasound calibration: a review”. W: *Advanced Imaging in Biology and Medicine: Technology, Software Environments, Applications*. Red. C. W. Sensen i B. Hallgrímsson. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, s. 47–84. ISBN: 978-3-540-68993-5. DOI: 10.1007/978-3-540-68993-5_3 (cytowane na stronie 51).
- [64] P. W. Hsu, R. W. Prager, A. H. Gee i G. M. Treece. „Real-time freehand 3D ultrasound calibration”. W: *Ultrasound in Medicine and Biology* 34.2 (2008), s. 239–251. DOI: doi:10.1016/j.ultrasmedbio.2007.07.020 (cytowane na stronach 51 i 52).
- [65] G. Huang, Z. Liu, L. Van Der Maaten i K. Q. Weinberger. „Densely connected convolutional networks”. W: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017, s. 2261–2269. DOI: 10.1109/CVPR.2017.243 (cytowane na stronie 61).
- [66] J. Huang, J. Triedman, N. Vasilyev, Y. Suematsu, R. Cleveland i P. Dupont. „Imaging artifacts of medical instruments in ultrasound-guided interventions”. W: *Journal of Ultrasound in Medicine* 26.10 (2007),

- s. 1303–1322. DOI: 10.7863/jum.2007.26.10.1303 (cytowane na stronach 7, 23 i 25).
- [67] Y.-L. Huang i D.-R. Chen. „Watershed segmentation for breast tumor in 2-D sonography”. W: *Ultrasound in Medicine and Biology* 30.5 (2004), s. 625–632. DOI: 10.1016/j.ultrasmedbio.2003.12.001 (cytowane na stronach 15, 16, 17, 18, 97 i 156).
- [68] Q. Huang, F. Yang, L. Liu i X. Li. „Automatic segmentation of breast lesions for interaction in ultrasonic computer-aided diagnosis”. W: *Information Sciences* 314 (2015), s. 293–310. ISSN: 0020-0255. DOI: <https://doi.org/10.1016/j.ins.2014.08.021> (cytowane na stronach 15, 16 i 17).
- [69] S.-F. Huang, R.-F. Chang, D.-R. Chen i W. K. Moon. „Characterization of spiculation on ultrasound lesions”. W: *IEEE Transactions on Medical Imaging* 23.1 (2004), s. 111–121. DOI: 10.1109/TMI.2003.819918 (cytowane na stronie 13).
- [70] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally i K. Keutzer. *SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size*. 2016. arXiv: 1602.07360 [cs.CV] (cytowane na stronach 61, 139 i 144).
- [71] J. J. In Chan, J. Ma, Y. Leng, K. K. Tan, C. W. Tan, R. Sultana, A. T. H. Sia i B. L. Sng. „Machine learning approach to needle insertion site identification for spinal anesthesia in obese patients”. W: *BMC Anesthesiology* 21.246 (2021). DOI: 10.1186/s12871-021-01466-8 (cytowane na stronie 22).
- [72] S. Ioffe i C. Szegedy. *Batch normalization: accelerating deep network training by reducing internal covariate shift*. 2015. arXiv: 1502.03167 [cs.LG] (cytowane na stronach 62 i 70).
- [73] S. Jadon. „A survey of loss functions for semantic segmentation”. W: *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* (paź. 2020). DOI: 10.1109/cibcb48159.2020.9277638 (cytowane na stronach 68, 69 i 85).
- [74] P. L. Jansen, A. Ghallab, N. Vartak, R. Reif, F. G. Schaap, J. Hampe i J. G. Hengstler. „The ascending pathophysiology of cholestatic liver disease”. W: *Hepatology* 65.2 (2017), s. 722–738. DOI: <https://doi.org/10.1002/hep.28965> (cytowane na stronach 4 i 120).
- [75] V. L. Jaya i R. Gopikakumari. „IEM: a new image enhancement metric for contrast and sharpness measurements”. W: *International Journal of Computer Applications* 79.9 (2013), s. 1–9 (cytowane na stronie 93).

- [76] P. Jiang, J. Peng, G. Zhang, E. Cheng, V. Megalooikonomou i H. Ling. „Learning-based automatic breast tumor detection and segmentation in ultrasound images”. W: *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*. 2012, s. 1587–1590. DOI: 10.1109/ISBI.2012.6235878 (cytowane na stronach 15, 16 i 17).
- [77] J. Kang, J. Y. Lee i Y. Yoo. „A new feature-enhanced speckle reduction method based on multiscale analysis for ultrasound B-mode imaging”. W: *IEEE Transactions on Biomedical Engineering* 63.6 (2016), s. 1178–1191. DOI: 10.1109/TBME.2015.2486042 (cytowane na stronie 75).
- [78] M. Kass, A. Witkin i D. Terzopoulos. „Snakes: Active contour models”. W: *INTERNATIONAL JOURNAL OF COMPUTER VISION* 1.4 (1988), s. 321–331. DOI: 10.1007/BF00133570 (cytowane na stronach 17 i 98).
- [79] M. Kaya, E. Senel i O. Bebek. „Gabor filter-based localization of straight and curved needles in 2D ultrasound images”. W: *Turkish Journal of Electrical Engineering and Computer Sciences* 26.5 (2020), s. 2940–2955. DOI: 10.3906/elk-1912-181 (cytowane na stronach 25, 26, 27 i 30).
- [80] Y. R. Kim, H. S. Kim i H.-W. Kim. „Are irregular hypoechoic breast masses on ultrasound always malignancies?: A pictorial essay”. W: *Korean Journal of Radiology* 16.6 (2015), s. 1266–1275. DOI: 10.3348/kjr.2015.16.6.1266 (cytowane na stronie 20).
- [81] D. P. Kingma i J. Ba. „Adam: A Method for Stochastic Optimization”. W: *Proc. 3rd International Conference on Learning Representations, ICLR 2015*. Red. Y. Bengio i Y. LeCun. 2015 (cytowane na stronach 77, 82 i 86).
- [82] Kriti, J. Virmani i R. Agarwal. „Characterization of breast tumors using selected laws’ mask texture features”. W: *2017 Fourth International Conference on Image Information Processing (ICIIP)*. 2017, s. 1–6. DOI: 10.1109/ICIIP.2017.8313743 (cytowane na stronach 119 i 125).
- [83] A. Krizhevsky, I. Sutskever i G. E. Hinton. „ImageNet classification with deep convolutional neural networks”. W: *Commun. ACM* 60.6 (maj 2017), s. 84–90. ISSN: 0001-0782. DOI: 10.1145/3065386 (cytowane na stronach 60, 61, 71, 139 i 144).
- [84] V. Kumar, R. Cotran i S. L. Robbins. „Nowotwory”. W: *Robbins Patologia*. Wrocław: Wydawnictwo Medyczne Urban & Partner, 2005. Rozd. 6, s. 185–238. ISBN: 0-7216-9274-5 (cytowane na stronie 1).

- [85] G. Lapouge, P. Poignet i J. Troccaz. „Towards 3D ultrasound guided needle steering robust to uncertainties, noise, and tissue heterogeneity”. W: *IEEE Transactions on Biomedical Engineering* 68.4 (2021), s. 1166–1177. DOI: 10.1109/TBME.2020.3022619 (cytowane na stronach 28 i 31).
- [86] A. Lasso, T. Heffter, A. Rankin, C. Pinter, T. Ungi i G. Fichtinger. „PLUS: Open-source toolkit for ultrasound-guided intervention systems”. W: *IEEE Transactions on Biomedical Engineering* 61.10 (2014), s. 2527–2537. DOI: 10.1109/TBME.2014.2322864 (cytowane na stronach 6, 25, 32, 42, 50, 115, 118 i 120).
- [87] J. F. Lazo, S. Moccia, E. Frontoni i E. D. Momi. *Comparison of different CNNs for breast tumor classification from ultrasound images*. 2020. arXiv: 2012.14517 [eess.IV] (cytowane na stronach 21 i 124).
- [88] C.-Y. Lee, T.-F. Chang, Y.-H. Chou i K.-C. Yang. „Fully automated lesion segmentation and visualization in automated whole breast ultrasound (ABUS) images”. W: *Quantitative Imaging in Medicine and Surgery* 10.3 (2020), s. 568–584. DOI: <https://doi.org/10.21037/qims.2020.01.12> (cytowane na stronach 17, 18 i 19).
- [89] J. Lee, M. Islam, J. Woh, T. Washeem, L. Ngoh, W. Wong i H. Ren. „Ultrasound needle segmentation and trajectory prediction using excitation network”. W: *International Journal of Computer Assisted Radiology and Surgery* 15 (2020), s. 437–443. DOI: 10.1007/s11548-019-02113-x (cytowane na stronach 26, 27, 29, 30, 106, 108, 146, 149 i 150).
- [90] L. K. Lee i S. C. Liew. „Breast ultrasound automated ROI segmentation with region growing”. W: *2015 4th International Conference on Software Engineering and Computer Systems (ICSECS)*. 2015, s. 177–182. DOI: 10.1109/ICSECS.2015.7333106 (cytowane na stronach 15, 16, 18, 88, 98 i 156).
- [91] Y. Lei, X. He, J. Yao, T. Wang, L. Wang, W. Li, W. Curran, T. Liu, D. Xu i X. Yang. „Breast tumor segmentation in 3D automatic breast ultrasound using Mask scoring R-CNN”. W: *Medical Physics* 48.1 (2021), s. 204–214. DOI: 10.1002/mp.14569 (cytowane na stronach 17, 18 i 19).
- [92] X. Li, Q. Long, X. Chen, D. He i H. He. „Assessment of the SonixGPS system for its application in real-time ultrasonography navigation-guided percutaneous nephrolithotomy for the treatment of complex kidney stones”. W: *Urolithiasis* 45 (2017), s. 221–227. DOI: 10.1007/s00240-016-0897-2 (cytowane na stronie 25).
- [93] Y. Lin, Z. Wang, X. Guo, H. Wang i F. Wang. *Closed-form freehand 3D ultrasound spatial calibration based on N-wedge phantom*. EasyChair Preprint no. 426. 2018. DOI: 10.29007/zrzd (cytowane na stronie 52).

- [94] B. Liu, H. Cheng, J. Huang, J. Tian, X. Tang i J. Liu. „Fully automatic and segmentation-robust classification of breast tumors based on local texture analysis of ultrasound images”. W: *Pattern Recognition* 43.1 (2010), s. 280–298. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2009.06.002 (cytowane na stronach 21, 23, 24 i 75).
- [95] A. Lozano, J. C. Hayes, L. M. Compton, J. Azarnoosh i F. Hassanipour. „Determining the thermal characteristics of breast cancer based on high-resolution infrared imaging, 3D breast scans, and magnetic resonance imaging”. W: *Scientific Reports* 10.1 (2020). DOI: 10.1038/s41598-020-66926-6 (cytowane na stronie 169).
- [96] J. Ma, D. Wang, X.-P. Wang i X. Yang. *A characteristic function-based algorithm for geodesic active contours*. 2021. arXiv: 2007.00525 [cs.CV] (cytowane na stronie 124).
- [97] W. Malina i M. Smiatacz. *Cyfrowe przetwarzanie obrazów*. Informatyka - Akademicka Oficyna Wydawnicza EXIT. Akademicka Oficyna Wydawnicza EXIT, 2008. ISBN: 9788360434543 (cytowane na stronach 88 i 89).
- [98] F. R. Margolin, L. Kaufman, S. R. Denny, R. P. Jacobs i J. D. Schrupf. „Metallic marker placement after stereotactic core biopsy of breast calcifications: comparison of two clips and deployment techniques”. W: *American Journal of Roentgenology* 181.6 (2003), s. 1685–1690 (cytowane na stronie 3).
- [99] R. V. Menon, P. Raha, S. Kothari, S. Chakraborty, I. Chakrabarti i R. Karim. „Automated detection and classification of mass from breast ultrasound images”. W: *2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)*. 2015, s. 1–4. DOI: 10.1109/NCVPRIPG.2015.7490070 (cytowane na stronach 13, 17, 18, 20, 21, 23, 24, 88, 99 i 125).
- [100] E. Michel-González, M. H. Cho i S. Y. Lee. „Geometric nonlinear diffusion filter and its application to X-ray imaging”. W: *BioMedical Engineering OnLine* 10.1 (2011), s. 47 (cytowane na stronach 91 i 95).
- [101] A. Mikołajczyk i M. Grochowski. „Data augmentation for improving deep learning in image classification problem”. W: *2018 International Interdisciplinary PhD Workshop (IIPhDW)*. 2018, s. 117–122. DOI: 10.1109/IIPhDW.2018.8388338 (cytowane na stronie 74).
- [102] Minavathi, S. Murali i M. S. Dinesh. „Classification of mass in breast ultrasound images using image processing techniques”. W: *International Journal of Computer Applications* 42 (2012), s. 29–36. DOI: 10.5120/5730-7801 (cytowane na stronach 5, 13, 16, 17, 18, 21, 23, 24 i 88).

- [103] F. Mohamed i C. V. Siang. „A survey on 3D ultrasound reconstruction techniques”. W: *Applications in Medicine and Biology* (2019). DOI: 10.5772/intechopen.81628 (cytowane na stronie 44).
- [104] W. K. Moon, Y.-W. Lee, H.-H. Ke, S. H. Lee, C.-S. Huang i R.-F. Chang. „Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks”. W: *Computer Methods and Programs in Biomedicine* 190 (2020), s. 105361. ISSN: 0169-2607. DOI: 10.1016/j.cmpb.2020.105361 (cytowane na stronie 124).
- [105] W. K. Moon, Y.-W. Shen, M. S. Bae, C.-S. Huang, J.-H. Chen i R.-F. Chang. „Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images”. W: *IEEE Transactions on Medical Imaging* 32.7 (2013), s. 1191–1200. DOI: 10.1109/TMI.2012.2230403 (cytowane na stronach 13, 17, 18 i 19).
- [106] O. Moussa, H. Khachnaoui, R. Guetari i N. Khelifa. „Thyroid nodules classification and diagnosis in ultrasound images using fine-tuning deep convolutional neural network”. W: *International Journal of Imaging Systems and Technology* 30.1 (2020), s. 185–195. DOI: <https://doi.org/10.1002/ima.22363> (cytowane na stronach 4 i 120).
- [107] M. H. Mozaffari i W.-S. Lee. „Freehand 3-D ultrasound imaging: a systematic review”. W: *Ultrasound in Medicine & Biology* 43.10 (2017), s. 2099–2124. ISSN: 0301-5629. DOI: 10.1016/j.ultrasmedbio.2017.06.009 (cytowane na stronie 44).
- [108] K. P. Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012 (cytowane na stronach 77, 78, 82, 86, 102, 103, 104, 105 i 106).
- [109] C. Mwikirize, J. Noshier i I. Hacihaliloglu. „Learning needle tip localization from digital subtraction in 2D ultrasound”. W: *International Journal of Computer Assisted Radiology and Surgery* 14 (list. 2019), s. 1017–1026. DOI: 10.1007/s11548-019-01951-z (cytowane na stronach 27 i 28).
- [110] C. Mwikirize, J. Noshier i I. Hacihaliloglu. „Single shot needle tip localization in 2D ultrasound”. W: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*. Cham: Springer International Publishing, 2019, s. 637–645. ISBN: 978-3-030-32254-0. DOI: https://doi.org/10.1007/978-3-030-32254-0_71 (cytowane na stronach 26, 27 i 30).
- [111] C. Mwikirize, J. Noshier i I. Hacihaliloglu. „Convolution neural networks for real-time needle detection and localization in 2D ultrasound.” W: *International Journal of Computer Assisted Radiology and Surgery* 13.5

- (2018), s. 647–657. DOI: 10.1007/s11548-018-1721-y (cytowane na stronach 26, 27, 28, 30, 108 i 123).
- [112] C. Mwikirize, A. B. Kimbowa, S. Imanirakiza, A. Katumba, J. L. Noshier i I. Hacihaliloglu. „Time-aware deep neural networks for needle tip localization in 2D ultrasound.” W: *International Journal of Computer Assisted Radiology and Surgery* 16 (2021), s. 819–827. DOI: 10.1007/s11548-021-02361-w (cytowane na stronach 26, 27, 28 i 77).
- [113] J. Nagi, F. Ducatelle, G. A. Di Caro, D. Cireşan, U. Meier, A. Giusti, F. Nagi, J. Schmidhuber i L. M. Gambardella. „Max-pooling convolutional neural networks for vision-based hand gesture recognition”. W: *2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. 2011, s. 342–347. DOI: 10.1109/ICSIPA.2011.6144164 (cytowane na stronach 64, 65 i 69).
- [114] V. Nair i G. E. Hinton. „Rectified linear units improve restricted boltzmann machines”. W: *Proceedings of the 27th International Conference on International Conference on Machine Learning*. ICML’10. Haifa, Israel: Omnipress, 2010, s. 807–814. ISBN: 9781605589077. DOI: 10.5555/3104322.3104425 (cytowane na stronie 69).
- [115] K. S. Ni i T. Q. Nguyen. „An adaptable k -nearest neighbors algorithm for MMSE image interpolation”. W: *IEEE Transactions on Image Processing* 18.9 (2009), s. 1976–1987. DOI: 10.1109/TIP.2009.2023706 (cytowane na stronie 58).
- [116] Northern Digital Inc. *Aurora V3 User Guide*. Revision 1. Waterloo, Ontario, 2014 (cytowane na stronach 32, 34, 35, 36, 39, 50 i 113).
- [117] Northern Digital Inc. *Polaris Vicra User Guide*. Revision 8. Waterloo, Ontario, 2019 (cytowane na stronach 32, 33, 39, 47, 50 i 113).
- [118] Northern Digital Inc. *The Polaris Family of Products*. Revision 3. Waterloo, Ontario, 2008 (cytowane na stronie 34).
- [119] P. M. Novotny, J. A. Stoll, N. V. Vasilyev, P. J. del Nido, P. E. Dupont, T. E. Zickler i R. D. Howe. „GPU based real-time instrument tracking with three-dimensional ultrasound”. W: *Medical Image Analysis* 11.5 (2007). Special Issue on the Ninth International Conference on Medical Image Computing and Computer-Assisted Interventions - MICCAI 2006, s. 458–464. ISSN: 1361-8415. DOI: 10.1016/j.media.2007.06.009 (cytowane na stronie 87).
- [120] A. Nugroho, R. Hidayat, H. A. Nugroho i J. Debayle. „Combinatorial active contour bilateral filter for ultrasound image segmentation”. W: *Journal of Medical Imaging* 7.5 (2020), s. 057003. DOI: 10.1117/1.JMI.7.5.057003 (cytowane na stronach 18, 88 i 99).

- [121] C. O'Reilly i T. Nielsen. „Automatic sleep spindle detection: benchmarking with fine temporal resolution using open science tools”. W: *Frontiers in Human Neuroscience* 9 (2015). ISSN: 1662-5161. DOI: 10.3389/fnhum.2015.00353 (cytowane na stronach 102, 103, 104, 105 i 106).
- [122] S. H. C. Ortiz, J. Macione, T. Chiu i M. D. Fox. „Super-resolution of ultrasound images by displacement, averaging, and interlacing”. W: *Medical Imaging 2009: Ultrasonic Imaging and Signal Processing*. Red. S. A. McAleavey i J. D'hooge. T. 7265. International Society for Optics i Photonics. SPIE, 2009, s. 726519. DOI: 10.1117/12.812883 (cytowane na stronie 58).
- [123] S. Osowski. „Głębokie sieci neuronowe i ich zastosowania w eksploracji danych”. W: *Przegląd Telekomunikacyjny - Wiadomości Telekomunikacyjne* 5 (2018), s. 112–121. DOI: 10.15199/59.2018.5.2 (cytowane na stronach 64, 69, 70 i 85).
- [124] G. Palubinskas. „Image similarity/distance measures: what is really behind MSE and SSIM?” W: *International Journal of Image and Data Fusion* 8.1 (2017), s. 32–53. DOI: 10.1080/19479832.2016.1273259 (cytowane na stronie 91).
- [125] E. Peli. „Contrast in complex images”. W: *Journal of the Optical Society of America* 7.10 (1990), s. 2032–2040 (cytowane na stronie 93).
- [126] P. Perona i J. Malik. „Scale-space and edge detection using anisotropic diffusion”. W: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12.7 (1990), s. 629–639. DOI: 10.1109/34.56205 (cytowane na stronie 89).
- [127] R. Pflanzner, M. Hofmann, A. Shelke, A. Habib, W. Derwich, T. Schmitz-Rixen, A. Bernd, R. Kaufmann i J. Bereiter-Hahn. „Advanced 3D-sonographic imaging as a precise technique to evaluate tumor volume”. W: *Translational Oncology* 7.6 (2014), s. 681–686 (cytowane na stronach 3 i 12).
- [128] A. Pourtaherian, F. Ghazvinian Zanjani, S. Zinger, N. Mihajlovic, G. Ng, H. Korsten i P. de With. „Robust and semantic needle detection in 3D ultrasound using orthogonal-plane convolutional neural networks”. W: *International Journal of Computer Assisted Radiology and Surgery* 13.9 (2018), s. 1321–1333. DOI: 10.1007/s11548-018-1798-3 (cytowane na stronach 26, 28 i 31).
- [129] A. Pourtaherian, S. Zinger, P. H. N. de With, H. H. M. Korsten i N. Mihajlovic. „Gabor-based needle detection and tracking in three-dimensional ultrasound data volumes”. W: *Proc. 2014 IEEE International Conference on Image Processing (ICIP)*. 2014, s. 3602–3606. DOI: 10.1109/ICIP.2014.7025731 (cytowane na stronach 25, 26, 28, 31 i 83).

- [130] D. M. W. Powers. *Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation*. 2020. arXiv: 2010.16061 [cs.LG] (cytowane na stronach 102, 103, 104, 105 i 106).
- [131] B. Pyciński. „Estimation of pointer calibration error in optical tracking system”. W: *Innovations in Biomedical Engineering*. Red. M. Gzik, E. Tkacz, Z. Paszcenda i E. Piętka. Cham: Springer International Publishing, 2017, s. 228–239. ISBN: 978-3-319-47154-9. DOI: 10.1007/978-3-319-47154-9_27 (cytowane na stronie 46).
- [132] B. Pyciński. *Zintegrowany system śledzenia i rejestracji obrazów ultrasonograficznych i tomografii komputerowej jamy brzusznej*. Gliwice: Politechnika Śląska, Wydział Automatyki, Elektroniki i Informatyki, 2015. ISBN: 978-83-934357-8-4 (cytowane na stronach 29 i 118).
- [133] B. Pyciński, J. Juszczak, A. Wijata, M. Galinska, J. Czajkowska i E. Pietka. „Image guided core needle biopsy of the breast”. W: *Information Technology in Biomedicine*. T. 762. Advances in Intelligent Systems and Computing. 2019, s. 160–171. DOI: 10.1007/978-3-319-91211-0_14 (cytowane na stronie 120).
- [134] A. Ramachandran, K. Shiva Balan, S. Kiran i M. Azharudeen. „Neural network pattern recognition of ultrasound image gray scale intensity histogram of breast lesions to differentiate between benign and malignant lesions”. W: *medRxiv* (2020). DOI: 10.1101/2020.05.01.20088245 (cytowane na stronach 21, 23, 24, 60 i 124).
- [135] R. Rocha, R. Pinto, D. Tavares i C. Goncalves. „Step-by-step of ultrasound-guided core-needle biopsy of the breast: review and technique”. W: *Radiologia Brasileira* 46 (sierp. 2013), s. 234–241. ISSN: 0100-3984. DOI: 10.1590/S0100-39842013000400010 (cytowane na stronach 7, 13 i 22).
- [136] P. Rodrigues. „Breast ultrasound image”. W: Mendeley Data, 2017. DOI: 10.17632/wmy84gzngw.1 (cytowane na stronach 22 i 124).
- [137] A. Rodtook i S. S. Makhanov. „Multi-feature gradient vector flow snakes for adaptive segmentation of the ultrasound images of breast cancer”. W: *Journal of Visual Communication and Image Representation* 24.8 (2013), s. 1414–1430. ISSN: 1047-3203. DOI: <https://doi.org/10.1016/j.jvcir.2013.09.009> (cytowane na stronach 15, 16, 17, 19, 20, 99 i 156).
- [138] O. Ronneberger, P. Fischer i T. Brox. „U-Net: convolutional networks for biomedical image segmentation”. W: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Red. N. Navab, J. Hornegger, W. M. Wells i A. F. Frangi. Cham: Springer International Publishing, 2015, s. 234–241. DOI: 10.1007/978-3-319-24574-4_28 (cytowane na stronach 19 i 83).

- [139] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg i L. Fei-Fei. „ImageNet large scale visual recognition challenge”. W: *International Journal of Computer Vision (IJCV)* 115.3 (2015), s. 211–252. DOI: 10.1007/s11263-015-0816-y (cytowane na stronach 60, 61, 62, 63, 141 i 145).
- [140] T. Sadad, A. Hussain, A. Munir, M. Habib, S. Ali Khan, S. Hussain, S. Yang i M. Alawairdhi. „Identification of breast malignancy by marker-controlled watershed transformation and hybrid feature set for health-care”. W: *Applied Sciences* 10.6 (2020). ISSN: 2076-3417. DOI: 10.3390/app10061900 (cytowane na stronach 23, 24 i 124).
- [141] H. Sadjadi, K. Hashtrudi-Zaad i G. Fichtinger. „Simultaneous electromagnetic tracking and calibration for dynamic field distortion compensation”. W: *IEEE Transactions on Biomedical Engineering* 63.8 (2016), s. 1771–1781. DOI: 10.1109/TBME.2015.2502138 (cytowane na stronie 34).
- [142] T. A. Sadoon i M. H. Ali. „Deep learning model for glioma, meningioma and pituitary classification”. W: *International Journal of Advances in Applied Sciences* 10.1 (2021), s. 88–98. DOI: 10.11591/ijaas.v10.i1.pp88-98 (cytowane na stronie 67).
- [143] L. Santiago, B. E. Adrada, M. L. Huang, W. Wei i R. P. Candelaria. „Breast cancer neoplastic seeding in the setting of image-guided needle biopsies of the breast”. W: *Breast Cancer Research and Treatment* 166.1 (lip. 2017), s. 29–39. DOI: 10.1007/s10549-017-4401-7 (cytowane na stronie 155).
- [144] U. Sara, M. Akter i M. S. Uddin. „Image quality assessment through FSIM, SSIM, MSE and PSNR A comparative study”. W: *Journal of Computer and Communications* 7.3 (2019). DOI: 10.4236/jcc.2019.73002 (cytowane na stronach 91 i 93).
- [145] F. Sattar, L. Floreby, G. Salomonsson i B. Löfvström. „Image enhancement based on a nonlinear multiscale method”. W: *IEEE Trans. Image Processing* 6.6 (1997), s. 888–895 (cytowane na stronach 91, 94 i 95).
- [146] S. L. Savaridas, Y. T. Sim, S. J. Vinnicombe, C. A. Purdie, A. M. Thompson i A. Evans. „Are baseline ultrasound and mammographic features associated with rates of pathological complete response in patients receiving neoadjuvant chemotherapy for breast cancer?” W: *Cancer Imaging* 19 (2019). DOI: 10.1186/s40644-019-0251-3 (cytowane na stronach 13 i 20).

- [147] R. Shanmugamani. *Deep Learning for Computer Vision. Expert techniques to train advanced neural networks using TensorFlow and Keras*. 2018. ISBN: 978-17-882-9335-8 (cytowane na stronach 21, 60, 66, 70, 71, 73, 74, 78, 80, 81, 82, 85, 133, 137, 139 i 144).
- [148] E. Shelhamer, J. Long i T. Darrell. „Fully convolutional networks for semantic segmentation”. W: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.4 (2017), s. 640–651. DOI: 10.1109/TPAMI.2016.2572683 (cytowane na stronach 19, 66 i 106).
- [149] B. Shinkins, M. Thompson, S. Mallett i R. Perera. „Diagnostic accuracy studies: how to report and analyse inconclusive test results”. W: *BMJ* 346 (2013). DOI: 10.1136/bmj.f2778 (cytowane na stronie 103).
- [150] F. J. Sipel, B. Maris, M. K. Welleweerd, V. Groenhuis, P. Fiorini i S. Stramigioli. „Needle and biopsy robots: a review”. W: *Current Robotics Reports* 2 (2021), s. 73–84. DOI: 10.1007/s43154-020-00042-1 (cytowane na stronie 22).
- [151] K. Simonyan i A. Zisserman. *Very deep convolutional networks for large-scale image recognition*. 2015. arXiv: 1409.1556 [cs.CV] (cytowane na stronach 60, 61, 139 i 144).
- [152] B. K. Singh, K. Verma i A. Thoke. „Fuzzy cluster based neural network classifier for classifying breast tumors in ultrasound images”. W: *Expert Systems with Applications* 66 (2016), s. 114–123. ISSN: 0957-4174. DOI: 10.1016/j.eswa.2016.09.006 (cytowane na stronach 21, 23 i 24).
- [153] R. J. van Sloun, L. Demi, A. W. Postema, J. JMCH De La Rosette, H. Wijkstra i M. Mischi. „Entropy of ultrasound-contrast-agent velocity fields for angiogenesis imaging in prostate cancer”. W: *IEEE Transactions on Medical Imaging* 36.3 (2017), s. 826–837. DOI: 10.1109/TMI.2016.2629851 (cytowane na stronie 75).
- [154] SonoSkills BV. *Free ultrasound library offered to you by SonoSkills and Hitachi Medical Systems Europe*. <https://www.ultrasoundcases.info/cases/breast-and-axilla/>. [accessed: June 2020]. 2020 (cytowane na stronach 119, 120 i 124).
- [155] A. Sorriento, M. B. Porfido, S. Mazzoleni, G. Calvosa, M. Tenucci, G. Ciuti i P. Dario. „Optical and electromagnetic tracking systems for biomedical applications: A critical review on potentialities and limitations”. W: *IEEE Reviews in Biomedical Engineering* 13 (2020), s. 212–232. DOI: 10.1109/RBME.2019.2939091 (cytowane na stronie 32).

- [156] D. Spinczyk i Ł. Kowalewska. „Przegląd metod segmentacji i interpretacji stosowanych w analizie obrazów mikroskopowych”. W: *Acta Bio-Optica et Informatica Medica. Inżynieria Biomedyczna* 23.4 (2017), s. 261–269. ISSN: 1234-5563 (cytowane na stronie 100).
- [157] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever i R. Salakhutdinov. „Dropout: a simple way to prevent neural networks from overfitting”. W: *J. Mach. Learn. Res.* 15.1 (sty. 2014), s. 1929–1958. ISSN: 1532-4435. DOI: 10.5555/2627435.2670313 (cytowane na stronie 71).
- [158] A. Stecco, C. Pirri, R. D. Caro i P. Raghavan. „Stiffness and echogenicity: Development of a stiffness-echogenicity matrix for clinical problem solving”. W: *European journal of translational myology* 29.3 (sierp. 2019), s. 8476. ISSN: 2037-7452. DOI: 10.4081/ejtm.2019.8476 (cytowane na stronie 13).
- [159] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin i M. Jorge Cardoso. „Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations”. W: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. T. 10553. Springer International Publishing, 2017, s. 240–248. DOI: 10.1007/978-3-319-67558-9_28 (cytowane na stronach 69 i 85).
- [160] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal i F. Bray. „Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries”. W: *CA: A Cancer Journal for Clinicians* 71.3 (2021), s. 209–249. DOI: <https://doi.org/10.3322/caac.21660> (cytowane na stronie 2).
- [161] C. Szegedy, S. Ioffe, V. Vanhoucke i A. Alemi. *Inception-v4, Inception-ResNet and the impact of residual connections on learning*. 2016. arXiv: 1602.07261 [cs.CV] (cytowane na stronach 60, 61, 62, 139 i 144).
- [162] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke i A. Rabinovich. „Going deeper with convolutions”. W: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015, s. 1–9. DOI: 10.1109/CVPR.2015.7298594 (cytowane na stronach 60, 61, 62, 139 i 144).
- [163] K. Szopiński i M. Szopińska. „Current opinion on imaging and biopsy of breast masses”. W: *Acta Clinica* 2.3 (2002), s. 273–294 (cytowane na stronie 3).
- [164] T. Ślubowski, M. Ślubowska i A. Wojciechowski. „Techniki diagnostyczne w raku piersi. Część II: Metody fizyczne i biopsyjne”. W: *Ginekologia Polska* 78 (2007), s. 479–483 (cytowane na stronach 3, 7, 13 i 22).

- [165] S. Tauscher, J. Tokuda, G. Schreiber, T. Neff, N. Hata i T. Ortmaier. „OpenIGTLink interface for state control and visualisation of a robot for image-guided therapy systems”. W: *International Journal of Computer Assisted Radiology and Surgery* 10 (2014), s. 285–292. DOI: I:10.1007/s11548-014-1081-1 (cytowane na stronie 42).
- [166] P. A. Toft. *The Radon transform – theory and implementation*. Technical University of Denmark, 1996 (cytowane na stronach 87 i 122).
- [167] C. Tomasi i R. Manduchi. „Bilateral filtering for gray and color images”. W: *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*. 1998, s. 839–846. DOI: 10.1109/ICCV.1998.710815 (cytowane na stronach 90 i 91).
- [168] G. M. Treece, A. H. Gee, R. W. Prager, C. J. Cash i L. H. Berman. „High-definition freehand 3-D ultrasound”. W: *Ultrasound in Medicine and Biology* 29.4 (2003), s. 529–546. DOI: 10.1016/S0301-5629(02)00735-4 (cytowane na stronie 51).
- [169] T. Ungi, A. Lasso i G. Fichtinger. „Tracked ultrasound in navigated spine interventions”. W: *Lecture Notes in Computational Vision and Biomechanics* 18 (2015), s. 469–494. DOI: 10.1007/978-3-319-12508-4_15 (cytowane na stronach 32, 39, 42, 47, 51, 53, 54, 113, 115 i 117).
- [170] T. Ungi, A. Lasso i G. Fichtinger. „Open-source platforms for navigated image-guided interventions”. W: *Medical Image Analysis* 33 (2016). 20th anniversary of the Medical Image Analysis journal (MedIA), s. 181–186. ISSN: 1361-8415. DOI: <https://doi.org/10.1016/j.media.2016.06.011> (cytowane na stronach 6, 44 i 55).
- [171] A. Vacavant, A.-R. Ali, M. Grand-Brochier, A. Albouy-Kissi, A. Alfidja, P. Chabrot i J.-Y. Boire. „Smoothed shock filtered defuzzification with Zernike moments for liver tumor extraction in MR images”. W: *2015 International Conference on Image Processing Theory, Tools and Applications (IPTA)*. 2015, s. 17–22. DOI: 10.1109/IPTA.2015.7367088 (cytowane na stronie 75).
- [172] K. W. Wan, C. H. Wong, F. Ip, D. Fan, P. L. Yuen, H. Y. Fong i M. Ying. „Evaluation of the performance of traditional machine learning algorithms, convolutional neural network and AutoML Vision in ultrasound breast lesions classification: a comparative study”. W: *Quantitative imaging in medicine and surgery* 11.4 (2021), s. 1381–1393. DOI: 10.21037/qims-20-922 (cytowane na stronach 13, 20, 21, 23, 24, 60 i 124).

- [173] K. Wang, S. Liang i Y. Zhang. „Residual feedback network for breast lesion segmentation in ultrasound image”. W: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*. Red. M. de Bruijne, P. C. Cattin, S. Cotin, N. Padoy, S. Speidel, Y. Zheng i C. Essert. Cham: Springer International Publishing, 2021, s. 471–481. ISBN: 978-3-030-87193-2. DOI: https://doi.org/10.1007/978-3-030-87193-2_45 (cytowane na stronach 119 i 125).
- [174] Z. Wang i A. C. Bovik. „A universal image quality index”. W: *IEEE Signal Processing Letters* 9.3 (2002), s. 81–84 (cytowane na stronie 94).
- [175] Z. Wang, A. C. Bovik, H. R. Sheikh i E. P. Simoncelli. „Image quality assessment: from error visibility to structural similarity”. W: *IEEE Transactions on Image Processing* 13.4 (2004), s. 600–612 (cytowane na stronie 93).
- [176] M. Wei, Y. Du, X. Wu, Q. Su, J. Zhu, L. Zheng, G. Lv i J. Zhuang. „A benign and malignant breast tumor classification method via efficiently combining texture and morphological features on ultrasound images”. W: *Computational and Mathematical Methods in Medicine* (2020). DOI: 10.1155/2020/5894010 (cytowane na stronach 6 i 21).
- [177] Z. Wei, L. Gardi, D. B. Downey i A. Fenster. „Oblique needle segmentation and tracking for 3D TRUS guided prostate brachytherapy”. W: *Medical Physics* 32.9 (2005), s. 2928–2941. DOI: <https://doi.org/10.1118/1.2011108> (cytowane na stronie 25).
- [178] T. Wen, C. Wang, Y. Zhang i S. Zhou. „A novel ultrasound probe spatial calibration method using a combined phantom and stylus”. W: *Ultrasound in Medicine & Biology* 46.8 (2020), s. 2079–2089. ISSN: 0301-5629. DOI: <https://doi.org/10.1016/j.ultrasmedbio.2020.03.018> (cytowane na stronie 51).
- [179] A. Wiacek, E. Falomo, K. Myers, O. M. Hoel Rindal, K. Fabrega-Foster, S. Harvey i M. A. Lediju Bell. „Clinical feasibility of coherence-based to distinguish solid from fluid hypoechoic breast masses”. W: *2018 IEEE International Ultrasonics Symposium (IUS)*. 2018, s. 1–4. DOI: 10.1109/ULTSYM.2018.8579846 (cytowane na stronach 13, 20 i 125).
- [180] W. Wieclawek, M. Rudzki, A. Wijata i M. Galinska. „Preliminary development of an automatic breast tumour segmentation algorithm from ultrasound volumetric images”. W: *Information Technology in Biomedicine*. Red. E. Pietka, P. Badura, J. Kawa i W. Wieclawek. Springer International Publishing, 2019, s. 77–88. ISBN: 978-3-319-91211-0 (cytowane na stronach 14, 17, 18, 88 i 97).

- [181] A. Wijata, Ż. Ranoż, M. Galińska, J. Juszczuk i J. Czajkowska. „Detection and tracking of the biopsy needle using ultrasound images”. W: *Innovations in Biomedical Engineering*. Springer International Publishing, 2018, s. 33–41. ISBN: 978-3-319-70063-2. DOI: 10.1007/978-3-319-70063-2_4 (cytowane na stronach 25, 26, 27, 28, 30 i 75).
- [182] A. Wijata, J. Andrzejewski i B. Pyciński. „An automatic biopsy needle detection and segmentation on ultrasound images using a convolutional neural network”. W: *Ultrasonic Imaging* 43.5 (czer. 2021), s. 262–272. ISSN: 0161-7346. DOI: 10.1177/01617346211025267 (cytowane na stronach 26, 29 i 85).
- [183] A. Wijata, B. Pyciński, M. Galińska i D. Spinczyk. „Breast lesion segmentation method using ultrasound images”. W: *Innovations in Biomedical Engineering*. Cham: Springer International Publishing, 2019, s. 20–27. ISBN: 978-3-030-15472-1. DOI: 10.1007/978-3-030-15472-1_3 (cytowane na stronach 17, 18, 88 i 99).
- [184] A. M. Wijata i J. Nalepa. „Unbiased validation of the algorithms for automatic needle localization in ultrasound-guided breast biopsies”. W: *2022 IEEE International Conference on Image Processing (ICIP)*. 2022, s. 3571–3575. DOI: 10.1109/ICIP46576.2022.9897449 (cytowane na stronie 151).
- [185] A. D. Wiles, D. G. Thompson i D. D. Frantz. „Accuracy assessment and interpretation for optical tracking systems”. W: *Medical Imaging 2004: Visualization, Image-Guided Procedures, and Display*. Red. R. L. G. Jr. T. 5367. International Society for Optics i Photonics. SPIE, 2004, s. 421–432 (cytowane na stronie 33).
- [186] U. Wojciechowska i J. Didkowska. *Zachorowania i zgony na nowotwory złośliwe w Polsce*. Dostępne na stronie <http://onkologia.org.pl/raporty/>, dostęp z dnia 02.02.2021. Krajowy Rejestr Nowotworów, Narodowy Instytut Onkologii im. Marii Skłodowskiej-Curie – Państwowy Instytut Badawczy, 2021 (cytowane na stronie 1).
- [187] World Health Organization. *Cancer*. <https://www.who.int/news-room/fact-sheets/detail/cancer>, dostęp z dnia 25.03.2022. 2022 (cytowane na stronie 2).
- [188] J. Wysocka. „New WHO classification of breast tumours – as published in 2019”. W: *Nowotwory. Journal of Oncology* 70.6 (2020), s. 250–252. ISSN: 2300-2115. DOI: 10.5603/NJO.a2020.0038 (cytowane na stronie 2).

- [189] W. Xia, S. J. West, M. C. Finlay, J.-M. Mari, S. Ourselin, A. L. David i A. E. Desjardins. „Looking beyond the imaging plane: 3D needle tracking with a linear array ultrasound probe”. W: *Scientific Reports* 7.1 (czer. 2017). DOI: 10.1038/s41598-017-03886-4 (cytowane na stronie 25).
- [190] M. Xian, Y. Zhang i H. Cheng. „Fully automatic segmentation of breast ultrasound images based on breast characteristics in space and frequency domains”. W: *Pattern Recognition* 48.2 (2015), s. 485–497. ISSN: 0031-3203. DOI: <https://doi.org/10.1016/j.patcog.2014.07.026> (cytowane na stronach 15, 16, 17, 18, 20 i 88).
- [191] C. Xu i J. Prince. „Gradient vector flow: a new external force for snakes”. W: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1997, s. 66–71. DOI: 10.1109/CVPR.1997.609299 (cytowane na stronie 99).
- [192] C. Xu i J. Prince. „Snakes, shapes, and gradient vector flow”. W: *IEEE Transactions on Image Processing* 7.3 (1998), s. 359–369. DOI: 10.1109/83.661186 (cytowane na stronie 99).
- [193] C. Xue, L. Zhu, H. Fu, X. Hu, X. Li, H. Zhang i P.-A. Heng. „Global guidance network for breast lesion segmentation in ultrasound images”. W: *Medical Image Analysis* 70 (2021), s. 101989. ISSN: 1361-8415. DOI: <https://doi.org/10.1016/j.media.2021.101989> (cytowane na stronach 14, 15 i 18).
- [194] H. Yang, C. Shan, A. Bouwman, A. F. Kolen i P. H. de With. „Efficient and robust instrument segmentation in 3d ultrasound using patch-of-interest-FuseNet with hybrid loss”. W: *Medical Image Analysis* 67 (2021), s. 101842. ISSN: 1361-8415. DOI: 10.1016/j.media.2020.101842 (cytowane na stronach 26, 28, 29, 31, 60 i 106).
- [195] Z. Yaniv, E. Wilson, D. Lindisch i K. Cleary. „Electromagnetic tracking in the clinical environment”. W: *Medical Physics* 36.3 (2009), s. 876–892. DOI: 10.1118/1.3075829 (cytowane na stronach 32 i 34).
- [196] M. Yap, M. Goyal, F. Osman, R. Marti, E. Denton, A. Juetten i R. Zwiggelaar. „Breast ultrasound lesions recognition: end-to-end deep learning approaches”. W: *Journal of Medical Imaging* 6.1 (2019), s. 786–799. DOI: [doi:10.1117/1.JMI.6.1.011007](https://doi.org/10.1117/1.JMI.6.1.011007) (cytowane na stronach 14, 16, 18, 19, 20, 22 i 124).
- [197] M. H. Yap, G. Pons, J. Martí, S. Ganau, M. Sentís, R. Zwiggelaar, A. K. Davison i R. Martí. „Automated breast ultrasound lesions detection using convolutional neural networks”. W: *IEEE Journal of Biomedical and Health Informatics* 22.4 (2018), s. 1218–1226. DOI: 10.1109/JBHI.2017.2731873 (cytowane na stronach 13, 14, 16, 18, 19, 20 i 124).

- [198] C. Yeong-Jun. *Weighted intersection over union (wIoU): a new evaluation metric for image segmentation*. 2021. arXiv: 2107.09858 (cytowane na stronie 106).
- [199] Q. Zhang, Y. Xiao, W. Dai, J. Suo, C. Wang, J. Shi i H. Zheng. „Deep learning based classification of breast tumors with shear-wave elastography”. W: *Ultrasonics* 72 (2016), s. 150–157. ISSN: 0041-624X. DOI: <https://doi.org/10.1016/j.ultras.2016.08.004> (cytowane na stronach 21, 23 i 24).
- [200] T. Zhang i B. Yu. „Boosting with early stopping: Convergence and consistency”. W: *The Annals of Statistics* 33.4 (2005), s. 1538–1579. DOI: 10.1214/009053605000000255 (cytowane na stronach 78, 82, 86, 135 i 144).
- [201] Y. Zhao, A. Bernard, C. Cachard i H. Liebgott. „Biopsy needle localization and tracking using ROI-RK method”. W: *Abstract and Applied Analysis* 2014 (2014), s. 1–7. DOI: 10.1155/2014/973147 (cytowane na stronach 25, 26, 28, 29 i 31).
- [202] Y. Zhao, Y. Shen, A. Bernard, C. Cachard i H. Liebgott. „Evaluation and comparison of current biopsy needle localization and tracking methods using 3D ultrasound”. W: *Ultrasonics* 73 (2017), s. 206–220. ISSN: 0041-624X. DOI: 10.1016/j.ultras.2016.09.006 (cytowane na stronach 25, 26, 28, 29 i 31).
- [203] Y. Zhou, H. Chen, Y. Li, Q. Liu, X. Xu, S. Wang, P.-T. Yap i D. Shen. „Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images”. W: *Medical Image Analysis* 70 (2021), s. 101918. ISSN: 1361-8415. DOI: <https://doi.org/10.1016/j.media.2020.101918> (cytowane na stronach 17, 18 i 19).
- [204] Z. Zhuang, Y. Kang, A. N. Joseph Raj, Y. Yuan, W. Ding i S. Qiu. „Breast ultrasound lesion classification based on image decomposition and transfer learning”. W: *Medical Physics* 47.12 (2020), s. 6257–6269. DOI: <https://doi.org/10.1002/mp.14510> (cytowane na stronach 13, 21 i 120).
- [205] Z. Zhuang, N. Li, A. N. Joseph Raj, V. G. V. Mahesh i S. Qiu. „An RDAU-NET model for lesion segmentation in breast ultrasound images”. W: *PLOS ONE* 14.8 (2019), s. 1–23. DOI: 10.1371/journal.pone.0221535 (cytowane na stronach 18 i 119).