

Kraków, 15 października 2010 r.

Recenzja

Przedmiotem recenzji jest rozprawa **doktorska mgr inż. Rafała Wąsika** zatytułowana „*Metoda automatycznej oceny żeliwa stosowanego na części maszyn, z wykorzystaniem sieci neuronowych*”. Opiniowana praca przygotowana została na Wydziale Transportu Politechniki Śląskiej pod kierunkiem naukowym **prof. dr hab. inż. Jan Piechy**. Recenzję przygotowano na zlecenie Dziekana Wydziału (pismo z dnia 20.09.2010 r.) działającego w oparciu o uchwałę Rady Wydziału Transportu Politechniki Śląskiej z dnia 16.09.2010.

Cel, zakres i charakter rozprawy

Opiniowana praca jest kolejnym opracowaniem powstającym w szeroko znanym i doskonale pracującym zespole naukowym prof. Jana Piechy. Głównym wątkiem ocenianej pracy są badania jakości materiałów przeznaczonych na części maszyn, a narzędziem użytym do tego celu są sieci neuronowe. Doceniam znaczenie techniczne badań jakości materiałów i pozytywnie oceniam wybór sieci neuronowych jako narzędzia wykorzystywanego do tego celu, dlatego uważam, że cel i zakres pracy zostały wybrane prawidłowo. Stwierdzam także, iż przedstawiona do recenzji praca mieści się w dziedzinie nauk technicznych, w obszarze kompetencji i działalności naukowej Wydziału Transportu Politechniki Śląskiej, gdzie praca została przedłożona w dyscyplinie Budowa i Eksploatacja Maszyn.

Mimo, że moje kompetencje jako recenzenta dotyczą głównie tych aspektów pracy, które związane są z wykorzystywanymi w niej narzędziami informatycznymi, to znaczy technikami komputerowego przetwarzania i analizy obrazów oraz metodami wykorzystania sieci neuronowej, to jednak czuję się na siłach zaopiniować tę rozprawę z następujących powodów. Otóż główne oryginalne dokonania Doktoranta są wprawdzie dedykowane do budowy i eksploatacji maszyn, ale oryginalna myśl naukowa zawarta w rozprawie głównie polega na znalezieniu metody automatycznego klasyfikowania metalograficznych obrazów preparatów żeliwa, co osiągnęte jest dzięki zastosowaniu wymienionych wyżej technik komputerowej analizy i przetwarzania obrazów z użyciem elementu decyzyjnego w postaci sieci neuronowej. Jak się wydaje, po pominięciu rozdziałów wstępnych, mających przeglądowy

i literaturowy charakter - jest to główny obszar problemowy, w którym Doktorant porusza się z godną uznania zręcznością. Z tego powodu czuję się upoważniony do wydania opinii o wartości naukowej rozprawy Pana mgra Wąsika.

Omówienie zawartości rozprawy

Ze względu na obecność obszernego wprowadzenia literaturowego, lokującego zadanie realizowane przez Doktoranta w kontekście problematyki budowy i eksploatacji maszyn, a w szczególności w kontekście oceny żeliwa stosowanego na części maszyn, właściwy cel pracy podany jest *explicite* dopiero w rozdziale 4. Tam także ulokowana jest teza rozprawy, będąca przy recenzowaniu doktoratów zawsze najważniejszą wskazówką dla recenzenta, na co powinien zwracać uwagę przy studiowaniu dysertacji, a co stanowi wątek uboczny. Jednak jest to późne wyartykułowanie formalne celu i tezy pracy uzasadnione jest faktem, że Doktorant był w trudnej sytuacji, rozwiązując problemy informatyczne ukierunkowane jednak na potrzeby budowy i eksploatacji maszyn. Z tego powodu w strukturze pracy trzeba było najpierw pokazać, w jaki sposób problematyka analizy i przetwarzania obrazów oraz sieci neuronowych jest osadzona w kontekście potrzeb bazowej dyscypliny naukowej. Dopiero po zarysowaniu tego tła można było tworzyć i oceniać odpowiednie narzędzia komputerowe, obiektywizujące procesy oceny rozważanych preparatów mikroskopowych struktury metalograficznej żeliwa.

Zaletą pracy jest dość przejrzyste zredagowany Wstęp, który pozwala poprawnie ukierunkować uwagę czytelnika, a rozdziały poprzedzające rozdział w którym zawarto tezę rozprawy stanowią dobrą podstawę dla wskazania, że teza ta jest ważna i z całą pewnością naukowo aktualna.

Charakteryzując zawartość tych właśnie rozdziałów wprowadzających stwierdzam, że wprawdzie nie zawierają one bezpośrednio wyników, które mógłbym wskazać jako osobiste dokonania naukowe Doktoranta, ale są to bardzo kompetentne opracowania literaturowe, mające duże znaczenie jako dowód, że Doktorant posiada obszerną i aktualną wiedzę naukową w obszarze uprawianej dyscypliny naukowej – co także jest jednym z wymagań przy nadawaniu stopnia naukowego doktora.

Przechodząc do szczegółów muszę przyznać, że rozdział 2, przedstawiający zagadnienia technologii wytwarzania żeliw oraz opis ich cech charakterystycznych całkowicie pozostaje poza zasięgiem moich możliwości oceny. Wyrażam więc nadzieję, że ten element pracy

oceniony będzie przez drugiego recenzenta, dla którego zagadnienia materiałowe nie mają zapewne żadnych tajemnic. Ja natomiast mogę się kompetentnie wypowiedzieć na temat rozdziału 3, w którym omówiono sieci neuronowych oraz zasady ich wykorzystania. Potwierdzam, że zrobiono to prawidłowo i z dużą znajomością rzeczy.

W rozdziale 4, jak już wyżej wspomniano, przedstawiono zakres pracy oraz jej tezę. Zawartość tego rozdziału omówiono w poprzednim punkcie recenzji.

W rozdziale 5 zamieszczono opis najbardziej znanych metod przetwarzania obrazów, w dużej mierze oparty (jak mi się wydaje) na moich pracach, które jednak chyba mylnie cytowano przedstawiając miejscami pozycję oznaczoną w spisie literatury jako [49] z pozycją wykazaną pod numerem [50]. W rozdziale tym obok danych zaczerpniętych z literatury dostrzegam także elementy oryginalne, związane bezpośrednio z planowanym celem i zakresem opiniowanej rozprawy. Takim oryginalnym elementem jest niewątpliwie podrozdział 5.2 a także końcowa część podrozdziału 5.3 – szkoda tylko, że nie oddzielona w bardziej zauważalny sposób od reszty tego podrozdziału. Dostyżnym pomysłem, zastosowanym w tym rozdziale, jest użycie zabiegu sztucznego generowania rekordów danych, zgodnych z przyjętymi elementami charakterystycznymi produktu trafionego. Cel takiego zabiegu jest oczywisty i w pełni akceptowalny – chodzi o uzyskanie bogatszego i bardziej reprezentatywnego zbioru uczącego, co znakomicie usprawnia proces uczenia sieci neuronowej. Jednak tego rodzaju rekordy wirtualne wprowadzają do obiektywnego procesu badawczego pewien element arbitralności, który budzi mój niepokój.

W rozdziale 6. przedstawiono metodykę badań zmierzających do określania rodzaju wydzielen grafitu na podstawie rzeczywistych obrazów zdjęć mikroskopowych struktury zeliwa. Jest to w mojej ocenie jeden z dwóch głównych rozdziałów pracy, w których przytoczone są oryginalne koncepcje Doktoranta i oryginalne wyniki jego własnych prac badawczych. W rozdziale tym zostały postawione i prawidłowo rozwiązane między innymi następujące zadania badawcze: segmentacja obrazu, wyodrębnianie badanych obiektów oraz ich identyfikacja. Służy to do udostępnienia następnemu etapowi badań (opisanemu w rozdziale 7) obiektów, które mogą być przedmiotem oceny ich liczby, wielkości i kształtu, pełniących rolę graficznych deskryptorów wydzielen grafitu. Dyskusja tego następnego etapu badań zostanie przeprowadzona za chwilę, natomiast tutaj chciałbym zwrócić uwagę na kilka wartościowych i oryginalnych elementów, będących własnym wkładem mgra Wąsika do komputerowej analizy cyfrowego obrazu mikroskopowego struktury metalograficznej próbek

żeliwa. I tak kolejno Doktorant sformułował i rozwiązał następujące zagadnienia: ekstrakcji mapy krawędzi obiektów na obrazie, algorytmu wypełniania obszarów i algorytmu identyfikacji obiektów. Dzięki rozwiązaniu wyżej wymienionych problemów obraz próbki metalograficznej zamieniony zostaje na kolekcję obiektów, których właściwości badane są w kolejnym rozdziale pod kątem możliwości stworzenia automatycznych metod służących do określenia ilościowych i jakościowych cech wtrąceń grafitu w strukturze żeliwa.

Automatyczne określanie cech charakterystycznych rozważanych obiektów (ziaren grafitu w strukturze żeliwa) jest przedmiotem rozdziału 7 rozprawy. W rozdziale tym określono kolejno parametry lokalne obiektu, współczynniki kształtu obiektów oraz cechy charakteryzujące kształt wydzielenia grafitu, co stanowi zasadniczy wynik nawiązujący do głównego celu rozprawy. Po dyskusji teoretycznej zaproponowanych miar cech charakterystycznych rozważanych obiektów przedstawiono (w podrozdziale 7.2) implementację zaproponowanych algorytmów. Implementacja ta w pracy potraktowana jest bardzo skrótowo, a wynik przedstawiony na rysunku 7.3. traktować trzeba jako jedynie ilustrację działania tej implementacji (zresztą zgodnie z jego tytułem), ale pracę uzupełnia załącznik i pewne dodatkowe uwagi przedstawię podczas dyskusji zawartości tego załącznika.

Pozostając jeszcze przez chwilę przy ocenie zawartości głównego woluminu opracowanej intrologatorsko rozprawy skomentować trzeba jeszcze rozdział 8 zawierający (zgodnie z tytułem) podsumowanie i wnioski oraz bibliografię pracy. O bibliografii mogą tylko powiedzieć, że jest dobrze dobrana do celu i zakresu pracy i wskazuje na to, że Autor dobrze zna aktualną literaturę przedmiotu i umie z niej właściwie korzystać. Kilka słów komentarza należą się natomiast rozdziałowi podsumowującemu pracę. W podsumowaniu tym zebrano i syntetycznie omówiono wszystkie najważniejsze elementy rozprawy, zmierzając do stwierdzenia, że cel dysertacji został osiągnięty, a postawiona teza rozprawy została wykazana. Istotna nowość wnoszona przez pracę mgr inż. Rafała Wąsika polega na koncentracji uwagi badacza głównie na dobrze mi znanym i sympatycznym narzędziu, jakim są sztuczne sieci neuronowe. Uważam, że w tym właśnie zastosowaniu, o jakim jest mowa w ocenianej rozprawie, narzędzie to może szczególnie dobrze wykazać swoje zalety i możliwości. Dlatego zawartość pracy już w tym miejscu oceniam pozytywnie.

Mam jednak do ocenianego tu podsumowania uwagę dyskusyjną. Jest dla mnie oczywiste, że rzeczywiście Doktorant przeprowadził odpowiednie badania i że tezę rozprawy udało się wykazać. Jednak na podstawie informacji, które zawarte są w ocenianej rozprawie w jej

głównym tekście – w niektórych składnikach pracy byłoby trudno to wykazać. Na przykład gdy czytam w podsumowaniu, że *sieci neuronowe mogą być wykorzystane do zdefiniowania modelu automatu iteracyjnego, w którym zapisano cechy obiektu, trudne lub niemożliwe do bezpośredniego opisu matematycznego*, to oczywiście zgadzam się z tą tezą, ale zapytuję: gdzie w tekście rozprawy konkretnie to wykazano na podstawie rozważanych problemów klasyfikacji żeliwa stosowanego na części maszyn?

Na szczęście treść pracy uzupełnia załącznik zatytułowany „*Elementy aplikacji metod wnioskowania i przygotowania danych wejściowych*”. Tytuł ten wydaje się niezbyt fortunny, bo z logicznego punktu widzenia przygotowanie danych wejściowych powinno być przed wnioskowaniem, a nie odwrotnie, ale nie czepiając się do trzeciorzędnych szczegółów mogę z satysfakcją stwierdzić, że całość **zawartości** tego załącznika opiniuję zdecydowanie pozytywnie, odnajdując w nim te wszystkie konkrety, których w głównym tekście pracy trochę mi brakowało. Wprawdzie do egzemplarza załącznika powinna być dodawana lupa, bo bez niej przestudiowanie na przykład informacji o użytych przez Autora zestawach cech obiektów (rys. VII na stronie 10 załącznika) jest prawie niemożliwe – ale dzięki zawartości części VI załącznika dowiedziałem się wreszcie konkretnie, jakich sieci neuronowych Doktorant używał (nie ma o tym mowy w głównym tekście pracy) i jakie osiągał wyniki – chociaż w tej ostatniej sprawie znowu kilka słów dodatkowego komentarza by się przydało (w szczególności nie wiem jak interpretować zawartość tabeli na rysunku XIII).

Podsumowując tę część opinii stwierdzam, że praca dotyczy automatyzacji procesu oceny obrazu mikroskopowego żeliwa z uwzględnieniem wydzieleni różnych typów. Osiągnięciem Autora pracy jest zdefiniowanie cech obrazu, które mogą się stać podstawą automatycznego klasyfikowania próbek żeliwa (czy dowolnego materiału o podobnym zapisie graficznym badanej mikrostruktury). Klasyfikatorem jest sieć neuronowa uczona metodą *backpropagation*, która wytrenowana na bazie obrazów sklasyfikowanych przez eksperta może potem wykonywać klasyfikację kolejnych próbek szybko i bez błędów. Wynik ten oceniam pozytywnie i stwierdzam, że osiągnięcia mgr Wąsika zawarte w opiniowanej pracy w pełni odpowiadają wymogom stosownej Ustawy (Ustawa o stopniach naukowych i o tytule naukowym oraz o stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 roku, Dziennik Ustaw Nr 65, poz. 595) i mogą stanowić podstawę do nadania Autorowi stopnia naukowego doktora.

Opinia na temat strony redakcyjnej rozprawy

Opiniowana praca została przygotowana starannie, o czym świadczą między innymi dobrze zredagowane rozdziały, a także ładnie dopracowane rysunki. Autor pisząc swą rozprawę wykazał, że umie stawiać zadania naukowe, umie je rozwiązywać a także poprawnie wyciągać wnioski oraz skutecznie je uzasadniać. Zawartość załącznika do pracy pokazuje ponad wszelką wątpliwość, że mgr Wąsik dysponuje także dużą biegłością w używaniu nowoczesnych narzędzi informatycznych (Matlab). Według mojej oceny strona redakcyjna ocenianej rozprawy jest zdecydowanie jej mocną stroną.

Dyskusja drobnych usterek rozprawy

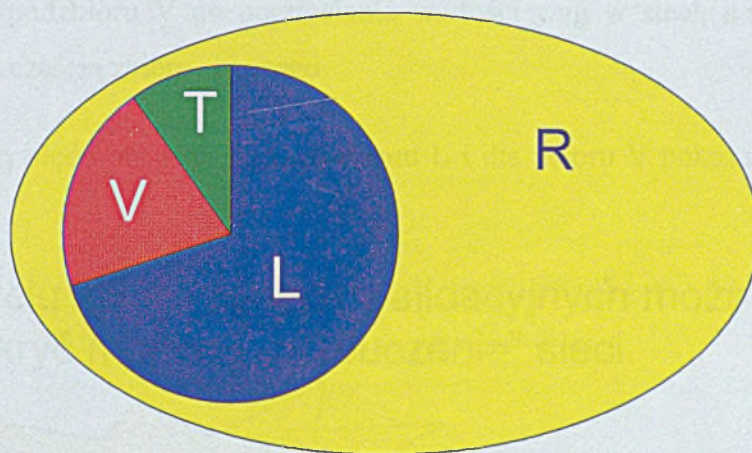
Opiniowana praca jest bardzo dobra merytorycznie a ponadto starannie przygotowana pod względem edytorskim, mam więc niewiele uwag dyskusyjnych, które chciałbym do niej zgłosić, spełniając podstawowy obowiązek recenzenta, którym jest wskazanie Doktorantowi ewentualnych usterek w celu skłonienia Go do doskonalenia kolejnych dalszych prac.

Moja jedyna poważniejsza wątpliwość merytoryczna wiąże się ze stosowaną przez mgr Wąsika regułą „stopu” algorytmu uczenia (rys. 3.5 na str.50). Nie wdając się w szczegóły mogę stwierdzić, że w opiniowanej pracy proces uczenia sieci neuronowych przerywa się, gdy błąd osiągnie wystarczająco niski poziom, **co nie jest rozwiązaniem właściwym.**

Postaram się wyjaśnić dlaczego:

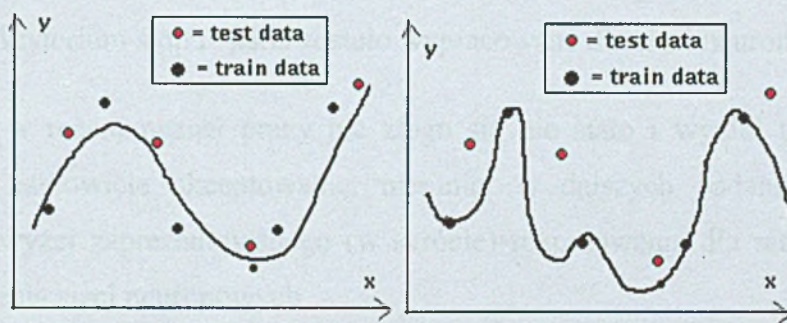
Otóż celem użycia sieci neuronowej jest zawsze odtworzenie pewnej nieznannej zależności, która ma funkcjonować dla **wszystkich** danych należących do pewnego zbioru R . Zależność tę staramy się zamodelować w sieci neuronowej, mając do dyspozycji wyłącznie pewien zasób danych Z , w których znane są zarówno dane wejściowe do budowanego modelu neuronowego, jak i dane wyjściowe, będące wzorcem, do którego dopasowywać będziemy parametry neuronowego modelu. Niebezpieczeństwo, które nam zagraża, polega na tym, że podczas uczenia sieci może dojść do nadmiernego dopasowania modelu do właściwości danych ze zbioru Z , nie koniecznie dobrze reprezentujących wszystkie dane w zbiorze R . Żeby tego uniknąć dokonujemy podziału zasobu Z na trzy części według następującego schematu: Wydzielamy (arbitralnie, ale losowo!) trzy podzbiory: L , V oraz T (patrz rysunek na następnej stronie).

Część danych L są te dane, które używamy do aktywnego uczenia sieci. To znaczny, że po każdym pokazaniu danych ze zbioru L następuje korekta wag neuronów w sieci zgodnie z wybranym algorytmem uczenia minimalizującym błąd.



Doświadczenie uczy, że do pewnego momentu polepszaniu działania sieci dla zbioru L towarzyszy polepszanie działania także dla zbioru R (czyli coraz lepsze rozwiązywanie całego postawionego zadania), ale w pewnym momencie sieć zaczyna się nadmiernie specjalizować w rozwiązywaniu problemów ze zbioru L, co pogarsza jej zdolność do generalizacji. Oznacza to, że wyniku uczenia, uzyskanego na zbiorze L, nie będziemy mogli potem użyć dla „obsługi” wszystkich przypadków ze zbioru R. Zjawisko to nazywa się **przeuczeniem** i jest zilustrowane na rysunku poniżej.

Na tych samych danych uczących sieć może zbudować model dobrze albo źle generalizujący



Model zachowujący dobrą zdolność generalizacji

Model który utracił zdolność generalizacji

Żeby wykryć, kiedy następuje przeuczenie, i żeby oczywiście przerwać w tym momencie uczenie, używa się danych ze zbioru V. Dane te przedstawia się sieci po każdej „epoce” procesu uczenia i określa się błąd, jaki sieć popełnia na tym zbiorze danych, ale absolutnie nie używa się danych z podzbioru V do poprawiania wartości wag w sieci, a więc ten zbiór (walidacyjny) nie jest częścią zbioru uczącego.

Przykładowy przebieg błędu otrzymanego dla zbioru L i dla zbioru V pokazano na rysunku poniżej.

Dzięki użyciu zbioru danych walidacyjnych można wykryć moment „przeuczenia” sieci.



— Błąd dla zbioru uczącego
— Błąd dla zbioru walidacyjnego

Gdy podczas walidacji pojawia się „efekt nożyc” (wzrostu błędu walidacji przy nadal malejącym błędzie na zbiorze uczącym) – uczenie należy przerwać. Jest to najbardziej racjonalne „kryterium stopu” jakie zostało wypracowane dla sieci neuronowych.

Oczywiście w recenzowanej pracy nic złego się nie stało i wyniki uzyskane przez mgra Wąsika są całkowicie akceptowalne, niemniej w dalszych badaniach gorąco zalecam stosowanie wyżej zaprezentowanego (w skrócie) rozumowania dla racjonalnego końca procesu uczenia sieci neuronowych.

Niejasne jest dla mnie, jakie zastosowanie w badaniach Doktoranta miały metody PCA (*Principal Component Analysis*), które opisano jako realizowane neuronowo metody do wstępnego przetworzenia danych (strony 52 i 53 rozprawy). Oczywiście zastosowanie PCA

w podobnych zadaniach zwykle jest celowe i zapewnia pozytywne rezultaty wydobywania najsilniej determinujących i w dodatku zdekorelowanych cech, jednak studiując rozprawę i załącznik nie zauważyłem miejsca, w którym by konkretnie była mowa o tym, jak tę technikę zastosowano do zadania oceny żeliwa i jakie uzyskano wyniki? Będę wdzięczny za uwzględnienie tego elementu w autoreferacie Doktoranta podczas obrony, bo jestem naprawdę ciekaw, co osiągnięto przy użyciu PCA w tym właśnie rozważanym tu zadaniu. Przy okazji chciałbym też prosić o wyjaśnienia, co oznaczają symbole (cyfry i jedna litera) wewnątrz sztucznych neuronów na rysunku 3.6 (str. 53) – bo absolutnie nie potrafię ich skojarzyć z niczym znanym.

Z drobnych usterek zauważonych w opiniowanej pracy warto może odnotować fakt, że we wzorach (3.6) i (3.7) powinno się zaznaczyć, że występujące tam wielkości mają charakter wektorowy. Opisano to wprawdzie przy dyskusji oznaczeń, ale zwyczajowo wektory i macierze wyróżnia się we wzorach wytłuszczonym krojem czcionek (lub strzałkami ponad odpowiednimi symbolami, co jednak jest bardziej kłopotliwe typograficznie). Szczególnie źle pod tym względem wyglądają strony 52 i 53 rozprawy, gdzie raz stosuje się wytłuszczenie dla zaznaczenia wektorowego charakteru pewnych oznaczeń (wzór (3.11)), a zaraz potem stosuje się wzory w który występują zarówno skalary jak i wektory oznaczane tym samym symbolem W – bez wyróżnienia tych, które są wektorami! (dotyczy wzorów (3.12) i (3.14)).

Zauważyłem też błędnie zapisany wzór (7.2) na stronie 101. Wzór ten w pracy wygląda następująco:

$$L = \frac{\pi}{4} \cdot \left[\cdot a \cdot \left(N_0 + N_{90} + \frac{a}{\sqrt{2}} \cdot (N_0 + N_{90}) \right) \right]$$

tymczasem poprawnie powinien mieć formę:

$$L = \frac{\pi}{4} \cdot \left[\cdot a \cdot (N_0 + N_{90}) + \frac{a}{\sqrt{2}} \cdot (N_0 + N_{90}) \right]$$

Niejasny jest też zapis wzoru (7.3) na stronie 102 – w szczególności powinien w tym wzorze być wyjaśniony związek indeksów i oraz j z argumentem funkcji k . Podobnie niejasna jest rola parametru k we wzorach (7.6) i (7.7).

Jak się wydaje niektóre określenia używane przez Doktoranta nie są do końca precyzyjne. Na przykład w podpisie rysunku 6.3. na stronie 90 napisano *Mapa krawędzi obrazu*, podczas gdy w moim przekonaniu rysunek ten nie przedstawia krawędzi tylko zbinaryzowane wnętrza prezentowanych obiektów.

Wniosek końcowy

Uważam, że opiniowana praca zatytułowana „*Metoda automatycznej oceny żeliwa stosowanego na części maszyn, z wykorzystaniem sieci neuronowych*” **zawiera** wartościowy i oryginalny dorobek naukowy Kandydata, a zakres i poziom uzyskanych wyników badawczych odpowiada ustawowym i zwyczajowym wymaganiom, stawianym rozprawom na stopień doktora nauk technicznych. Szczegółowe rozwinięcie tego stwierdzenia znajduje się w całej treści recenzji. Wnioskuje zatem do Wysokiej Rady Wydziału Transportu Politechniki Śląskiej o **przyjęcie** rozprawy i dopuszczenie jej Autora, **mgr inż. Rafała Wąsika** do publicznej obrony.