DOI:10.34918/85110

Piotr WILCZEK[1,*]

Chapter 10. AN APPLICATION OF THE LOCAL BINARY PATTERN ALGORITHM AND ITS UNIFORM VARIANT TO IMPROVE THE RECURRENCE AND CROSS--RECURRENCE QUANTIFICATION ANALYSES OF THE PHARMACOLOGICALLY AND PHYSIOLOGICALLY IMPORTANT TIME SERIES

10.1. Introduction

Since its introduction in 1987 by J.-P. Eckmann et al. [2] and the development of several quantification approaches, the so-called *recurrence plots* (RPs) were widely used for the study of complex systems in a variety of scientific disciplines, as physiology, ecology, finance or earth sciences [10]. The RP and related methodologies are aimed to visualize and quantitatively assess the recurrences of nonlinear dynamical systems. For instance, if we consider a trajectory $\{\vec{x}_i\}_{i=1}^{n}$ of some dynamical system in its phase space, then it is straightforward to realize that the evolution of this system can be described by a time series whose components are the above vectors. The corresponding RP efficiently visualizes recurrences of the analyzed system and can be formally obtained from the following *recurrence matrix* (R) [10]:

$$R_{i,j}(\varepsilon) \coloneqq \Theta\left(\varepsilon - \left\|\vec{x}_i - \vec{y}_j\right\|_2\right) \tag{1}$$

for $i,j = 1,\dots,n$ where n is the number of *measured time points*, ε is a *threshold distance*, $\Theta(\cdot)$ is the *Heaviside function* (i.e., $\Theta(\cdot) = 0$ if $x < 0$ and $\Theta(x) = 0$ otherwise) and $\|\cdot\|_2$ is the *Euclidean norm*. For ε-recurrent states, i.e., for states which are in an ε-neighbourhood, it is possible to introduce the subsequent condition [10]:

$$\vec{x}_i \approx \vec{x}_j \leftrightarrow R_{i,j} = 1. \tag{2}$$

[1] Computer Laboratory, Poznań, Poland.

[*] Corresponding author: piotr.wilczek.net@onet.pl.

Then, the RP structure is obtained by plotting a black dot at the coordinates (i,j) if $R_{i,j}=1$ and a white dot if $R_{i,j}=0$. Both axes of the RP are identified with time axes. Since $R_{i,i}=1$ by convention, the RP has always a black main diagonal line, i.e., the line of identity (LOI). Moreover, the RP is symmetric by definition with respect to the main diagonal, i.e., $R_{i,j}=R_{j,i}$. Roughly speaking, the RP object juxtaposes the states of a nonlinear system at times i and j, i.e., if states are similar, then $R_{i,j}=1$ and otherwise $R_{i,j}=0$.

In turn, the *cross-recurrence plot* (CRP) can be understood as a bivariate extension of the RP and was proposed in order to analyze the dependencies between two different nonlinear dynamical systems by comparing their states [10]. Namely, if we consider two systems, each one represented by the trajectories $\vec{x}_i$ and $\vec{y}_j$ in the same phase space, then the *cross-recurrence matrix* (CR) is defined by the condition [10]:

$$CR_{i,j}^{\vec{x},\vec{y}}(\varepsilon) \coloneqq \Theta\left(\varepsilon - \left\|\vec{x}_i - \vec{y}_j\right\|_2\right) \tag{3}$$

for $i=1,\dots,n$ and $j=1,\dots,m$ where the length of the trajectories of both dynamical systems is not required to be identical and, consequently, the CR matrix is not necessarily square. Then, for ε-recurrent states, i.e., for states which are in an ε-neighbourhood, it is possible to introduce the subsequent condition [10]:

$$\vec{x}_i \approx \vec{y}_j \leftrightarrow CR_{i,j}=1. \tag{4}$$

Finally, the CRP is obtained by plotting a black dot at the coordinates (i,j) if $CR_{i,j}=1$ and a white dot if $CR_{i,j}=0$. Thus, it can be easily observed that the CRP structure compares the states of two nonlinear systems at times i and j, i.e., if states are similar, then $CR_{i,j}=1$ and otherwise $CR_{i,j}=0$. In other words, the CRP object detects those time points when a state of the first system recurs to one of the second system. There is usually not a black main diagonal line on the CRP structure because the values of the main diagonal $CR_{i,i}$ are not necessarily equal to one.

In the present paper, we will show the usefulness of the *recurrence quantification analysis* (RQA) in investigating the pharmacological and physiological nonlinear time series. Moreover, we will propose to apply the *local binary pattern* (LBP) algorithm and its uniform variant (ULBP) to improve the performance of the RP and CRP methodologies. The paper is structured as follows: Section 3 details the RQA complexity descriptors, Section 4 presents the novel (U)LBP-based RQA complexity indices, Section 5 describes the datasets and computational procedure whereas Section 6 includes the results of the analysis of the exemplary time series. Finally, Section 7 concludes the paper.

10.2. The RQA complexity descriptors

A crucial parameter of the RP algorithm is the threshold distance ε [10]. According to J.P. Zbilut and coworkers, it is possible to choose ε with respect to the *recurrence rate* ($RR(\varepsilon)$) of the R matrix [21]. For the time series of the length n, this matrix--theoretical index is expressed by the formula [10]:

$$RR(\varepsilon) := \frac{1}{n^2}\sum_{i,j=1}^{n} R_{i,j}(\varepsilon). \tag{5}$$

It was suggested to select ε such that the $RR(\varepsilon)$ coefficient is approximately equal to 1% [21]. Moreover, the RQA developed several complexity descriptors in order to analyze nonlinear time series data and to extract meaningful statistics and other characteristics of the data [10]. All listed below RQA indices refer to the RP object obtained from a time series of the length n.

10.2.1. The complexity measures based on diagonal lines

The *histogram of diagonal lines of length* l ($P(\varepsilon, l)$) has the form [10]:

$$P(\varepsilon, l) \coloneqq \sum_{i,j=1}^{n} \Big(1 - R_{i-1,j-1}(\varepsilon)\Big)\Big(1 - R_{i+l,j+l}(\varepsilon)\Big)\prod_{k=0}^{l-1} R_{i+k,j+k}(\varepsilon). \tag{6}$$

In the rest of the present article, the symbol ε will be omitted from the RQA measures for the sake of simplicity (i.e., $P(l) = P(\varepsilon, l)$). Then, the *determinism* (DET) of the RP structure is identified with the ratio of recurrence points that constitute diagonal structures (of at least length l_{min}) to all recurrence points [10], i.e.,

$$DET \coloneqq \frac{\sum_{l=l_{min}}^{n} lP(l)}{\sum_{l=1}^{n} lP(l)}. \tag{7}$$

This descriptor is aimed to quantify the predictability of the dynamical system. In turn, the *average diagonal line length* (L_{mean}) of the RP structure, the *length of the longest diagonal line* (L_{max}) found in the RP and its *inverse* (i.e., the so-called *divergence – DIV*) are expressed by the subsequent conditions [10]:

$$L_{mean} \coloneqq \frac{\sum_{l=l_{min}}^{n} lP(l)}{\sum_{l=l_{min}}^{n} P(l)}, \tag{8}$$

$$L_{max} \coloneqq \max\left(\{l_i\}_{i=1}^{n_l}\right) \tag{9}$$

and by

$$DIV \coloneqq \frac{1}{L_{max}} \tag{10}$$

where $n_l \coloneqq \sum_{l \geq l_{min}} P(l)$ refers to the total number of diagonal lines, respectively. If in the computation of the L_{mean} index, we exclude the main diagonal line, then the resultant descriptor is denoted by the symbol L_{mean}^{-}. If we identify the *probability* of finding a diagonal line of the length l in the RP structure with the formula $p(l) \coloneqq \frac{P(l)}{n_l}$, then the *Shannon entropy of this distribution* ($ENTR$) is defined as [10]:

$$ENTR \coloneqq -\sum_{l=l_{min}}^{n} p(l) \ln p(l). \tag{11}$$

A further descriptor used to quantitatively assess the RP object is the *ratio* ($RATIO$) defined as the ratio between DET and RR. Formally, this index has the form [10]:

$$RATIO \coloneqq n^2 \frac{\sum_{l=l_{min}}^{n} lP(l)}{(\sum_{l=1}^{n} lP(l))^2}. \tag{12}$$

The RR measure computed for a certain line parallel to the LOI with the distance τ from the LOI is called the *τ-recurrence rate* ($\tau - RR$). The $\tau - RR$ for the diagonal lines with the distance τ from the LOI is given by the formula [10]:

$$RR_\tau \coloneqq \frac{1}{n-\tau} \sum_{i=1}^{n-\tau} R_{i,i+\tau} = \frac{1}{n-\tau} \sum_{l=1}^{n-\tau} lP_\tau(l) \tag{13}$$

where the symbol $P_\tau(l)$ refers to the number of diagonal lines of the length l on each diagonal $R_{i,i+\tau}$ parallel to the LOI. Namely, $\tau = 0$ corresponds to the main diagonal, $\tau > 0$ to diagonals above and $\tau < 0$ to diagonals below the LOI (i.e., $R_{i+|\tau|,i}$). Then, a linear regression coefficient over the $\tau - RR$ of the diagonals parallel to the LOI as a function of the distance between these diagonals and the LOI is called the *trend* ($TREND$) and is defined as [10]:

$$TREND \coloneqq \frac{\sum_{\tau=1}^{\bar{n}} \left(\tau - \frac{\bar{n}}{2}\right) (RR_\tau - \langle RR_\tau \rangle)}{\sum_{\tau=1}^{\bar{n}} \left(\tau - \frac{\bar{n}}{2}\right)^2} \tag{14}$$

where $\bar{n} < n$ and the symbol $\langle \cdot \rangle$ denotes the average value.

10.2.2. The complexity measures based on vertical lines

The total number of vertical lines of the length v in the RP is calculated from the following histogram [10]:

$$P(v) \coloneqq \sum_{i,j=1}^{n} \left(1 - R_{i,j}\right)\left(1 - R_{i,j+v}\right) \prod_{k=0}^{v-1} R_{i,j+k}. \tag{15}$$

Then, the ratio between the recurrence points constituting vertical structures and the entire set of recurrence points is known as the *laminarity* (LAM) of the RP object. Formally, this descriptor is defined as [10]:

$$LAM \coloneqq \frac{\sum_{v=v_{min}}^{n} vP(v)}{\sum_{v=1}^{n} vP(v)}. \tag{16}$$

Moreover, the *maximal length of vertical lines* (V_{max}) in the RP structure as well as the average length of vertical lines (V_{mean}) in the RP object are given by the conditions [10]:

$$V_{max} \coloneqq max\left(\{v_l\}_{l=1}^{n_v}\right) \tag{17}$$

where n_v is the absolute number of vertical lines and

$$V_{mean} \coloneqq \frac{\sum_{v=v_{min}}^{n} vP(v)}{\sum_{v=v_{min}}^{n} P(v)}, \tag{18}$$

respectively. Equivalently, the V_{mean} complexity descriptor is called the *trapping time* [10].

10.3. The novel (U)LBP-based RQA complexity descriptors

The recurrence and cross-recurrence matrices whose definitions include the threshold distance ε can be regarded as the *thresholded* recurrence and cross-recurrence two--dimensional arrays. Besides these thresholded objects, it is possible to define the

unthresholded recurrence matrix (UR) as well as the *unthresholded cross-recurrence matrix* (UCR) [10]. Formally, for a nonlinear dynamical system, its UR matrix has the form [10]:

$$UR_{i,j} \coloneqq \left\| \vec{x}_i - \vec{y}_j \right\|_2 \tag{19}$$

for $i, j = 1, \ldots, n$ where n is the number of measured time points. In turn, if we study two dynamical systems, each one represented by the trajectories $\vec{x}_i$ and $\vec{y}_j$ in the same phase space, then their UCR matrix is given by the condition [10]:

$$UCR_{i,j}^{\vec{x},\vec{y}} \coloneqq \left\| \vec{x}_i - \vec{y}_j \right\|_2 \tag{20}$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, m$ where the length of both trajectories is not required to be identical and, therefore, the UCR array is not necessarily square. Moreover, the entries of both untresholded matrices can be transformed according to the following rules [3]:

$$\mathcal{UR}_{i,j} \coloneqq e^{-UR_{i,j}} \tag{21}$$

and

$$\mathcal{UCR}_{i,j}^{\vec{x},\vec{y}} \coloneqq e^{-UCR_{i,j}^{\vec{x},\vec{y}}}, \tag{22}$$

respectively. The resultant negatively exponentially transformed matrices are denoted by the calligraphic letters. Thus, for very similar states, the values of the entries of both matrices near one and for identical states they are exactly equal to one. In turn, for very dissimilar states, the values of the entries of both two-dimensional arrays approach zero [3]. Moreover, in the present paper, we propose to divide the values of the entries of both transformed matrices into ten equivalence classes. These classes are identified with the values 0.1, 0.2, …, 0.9, 1 and the resultant *coarse-grained* unthresholded recurrence and cross-recurrence matrices are denoted by the symbols $cg\mathcal{UR}$ and $cg\mathcal{UCR}$, respectively. In details, using the notation of the form $M \in \{\mathcal{UR}, \mathcal{UCR}\}$ and $cgM \in \{cg\mathcal{UR}, cg\mathcal{UCR}\}$, the equivalence classes, i.e., the entries of both coarse-grained matrices are defined by the conditions included in Table 1[2]. Next, we propose to run the *local binary pattern* (LBP) algorithm and/or its simplified counterpart on both unthresholded matrices and their coarse-grained counterparts in order to obtain the novel arrays whose entries are 8-*digit binary numbers* usually converted to *decimal numbers* for convenience. Recall that the LBP

[2] In Table 1, the symbol ∧ refers to the logical *conjunction*.

technique is a type of visual texture descriptor used for classification tasks in the realm of *computer vision* [1, 8, 11, 12]. Although the LBP methodology started in 1994 [12], the present contribution can be regarded as a first attempt to apply the LBP algorithm and its modified variant in the analysis of the nonlinear dynamical systems, for instance, in studying the dose-response time series data.

Table 1

The formal conditions defining the coarse-grained unthresholded matrices

ENTRY	CONDITION	ENTRY	CONDITION
$cgM_{i,j} = 0.1$	$M_{i,j} < 0.1$	$cgM_{i,j} = 0.6$	$M_{i,j} \geq 0.5 \wedge M_{i,j} < 0.6$
$cgM_{i,j} = 0.2$	$M_{i,j} \geq 0.1 \wedge M_{i,j} < 0.2$	$cgM_{i,j} = 0.7$	$M_{i,j} \geq 0.6 \wedge M_{i,j} < 0.7$
$cgM_{i,j} = 0.3$	$M_{i,j} \geq 0.2 \wedge M_{i,j} < 0.3$	$cgM_{i,j} = 0.8$	$M_{i,j} \geq 0.7 \wedge M_{i,j} < 0.8$
$cgM_{i,j} = 0.4$	$M_{i,j} \geq 0.3 \wedge M_{i,j} < 0.4$	$cgM_{i,j} = 0.9$	$M_{i,j} \geq 0.8 \wedge M_{i,j} < 0.9$
$cgM_{i,j} = 0.5$	$M_{i,j} \geq 0.4 \wedge M_{i,j} < 0.5$	$cgM_{i,j} = 1$	$M_{i,j} \geq 0.9$

More specifically, our paper proposes to scrutinize the recurrences of the studied dynamical system by extracting from its (coarse-grained) unthresholded recurrence and cross-recurrence matrices the (uniform) local binary patterns in the form of feature vectors. The LBP procedure can be run on *any* real matrices (for instance, on the $\mathcal{UR}, \mathcal{UCR}$ or on their coarse-grained analogues) whose entries are interpreted as *pixels*. A single instantiation of the LBP procedure is defined on 3×3 submatrix (of the input $n \times m$ two-dimensional real array) consisting the central pixel at the position (2,2) and eight neighboring pixels. Formally, this type of neighborhood is called the *Moore neighborhood*, i.e., all eight neighboring pixels are located at the *Chebyshev distance* equal to one from the central pixel. In the LBP algorithm, all neighboring pixels that have the values higher than or equal to the value of the central pixel are set to one otherwise they are set to zero. Then, starting from the pixel located at the cell (1,2) and moving clockwise to the pixel located at the cell (1,1), we obtain the 8-digit binary number. For convenience, the resultant binary number is converted to the corresponding decimal number. This decimal value is regarded as the LBP pattern of the central pixel. Consequently, if we repeat the same process for all 3×3 submatrices of the input $n \times m$ real matrix, then we obtain the histogram of all LBP values. Moreover, this histogram can be treated as the *feature vector*. From the fact that the number of 8-digit binary *words* is $2^8 = 256$, it follows that the dimensionality of the output feature vector is equal to 256. In turn, its length is equal to $nm - 2n - 2m + 4$. We will denote this numeric object by the symbol VB_{256}. Alternatively, the VB_{256} vector can be matricized into the $(n-2) \times (m-2)$ matrix denoted by the symbol B_{256}.

This matricization is performed column-wise, i.e., the first $(n-2)$ values of the VB_{256} object constitute the first column of the B_{256} matrix and the last $(m-2)$ values of the VB_{256} object form the last column of the B_{256} matrix. Some useful extension of the original LBP algorithm is the so-called *uniform local binary pattern* (ULBP) operator. This modified procedure can be used to reduce the length of the original feature vector. The conceptual idea of this modified procedure is motivated by the fact that some binary patterns are more frequent in texture images than other. The LBP is called *uniform* if it contains at most two 0-1 or 1-0 *transitions*. For instance, the word 0001000 includes two transitions and is treated as uniform, but the word 01010100 includes six transitions and, consequently, can not be regarded as uniform. In the computation of the ULBP histogram, the resultant histogram has a separate bin for every uniform pattern whereas all non-uniform patterns are assigned to one bin. It can be easily observed that among 256 8-digit binary words, it is possible to single out 58 uniform words. Consequently, by using the ULBP algorithm, the dimensionality of the resultant feature vector is reduced from 256 to 59. In this case, for the input $n \times m$ two-dimensional real array, the output 59-dimensional feature vector has also the length equal to $nm - 2n - 2m + 4$. In the ULBP approach, all uniform words are indexed by the decimal number from 0 to 57 whereas all non-uniform words are assigned to the decimal value equal to 58. Therefore, the obtained 59-dimensional vector is denoted by VB_{59} whereas its matricized form is denoted by B_{59}. In the remainder of the present article, the both types of feature vectors, i.e., VB_{59} and VB_{256} as well as their matricized forms, i.e., B_{59} and B_{256} derived from the arrays $\mathcal{UR}, \mathcal{UCR}$ are denoted by the symbols: $VB_{59}^{\mathcal{UR}}$, $VB_{59}^{\mathcal{UCR}}$, $VB_{256}^{\mathcal{UR}}$, $VB_{256}^{\mathcal{UCR}}$ and by $B_{59}^{\mathcal{UR}}$, $B_{59}^{\mathcal{UCR}}$, $B_{256}^{\mathcal{UR}}$, $B_{256}^{\mathcal{UCR}}$, respectively. In turn, the feature vectors and their matrix counterparts derived from the coarse-grained structures are denoted by the symbols: $VB_{59}^{cg\mathcal{UR}}$, $VB_{59}^{cg\mathcal{UCR}}$, $VB_{256}^{cg\mathcal{UR}}$, $VB_{256}^{cg\mathcal{UCR}}$ and by $B_{59}^{cg\mathcal{UR}}$, $B_{59}^{cg\mathcal{UCR}}$, $B_{256}^{cg\mathcal{UR}}$, $B_{256}^{cg\mathcal{UCR}}$, respectively.

Example 1. For the 6-element time series $ts = (2, 6, 4, 9, 0, 5)$, its UR matrix and its negatively exponentially transformed $\mathcal{UR}$ derivative have the forms

$$UR = \begin{vmatrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 0 & 4 & 2 & 7 & 2 & 3 \\ 2 & 4 & 0 & 2 & 3 & 6 & 1 \\ 3 & 2 & 2 & 0 & 5 & 4 & 1 \\ 4 & 7 & 3 & 5 & 0 & 9 & 4 \\ 5 & 2 & 6 & 4 & 9 & 0 & 5 \\ 6 & 3 & 1 & 1 & 4 & 5 & 0 \end{vmatrix},$$

$$\mathcal{UR} = \begin{vmatrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 1 & 0.018 & 0.135 & 0.001 & 0.135 & 0.05 \\ 2 & 0.018 & 1 & 0.135 & 0.05 & 0.002 & 0.368 \\ 3 & 0.135 & 0.135 & 1 & 0.007 & 0.018 & 0.368 \\ 4 & 0.001 & 0.05 & 0.007 & 1 & 0 & 0.018 \\ 5 & 0.135 & 0.002 & 0.018 & 0 & 1 & 0.007 \\ 6 & 0.05 & 0.368 & 0.368 & 0.018 & 0.007 & 1 \end{vmatrix},$$

respectively.

The entries of the coarse-grained counterpart of the $\mathcal{UR}$ array are given below

$$cg\mathcal{UR} = \begin{vmatrix} & 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 1 & 0.1 & 0.2 & 0.1 & 0.2 & 0.1 \\ 2 & 0.1 & 1 & 0.2 & 0.1 & 0.1 & 0.4 \\ 3 & 0.2 & 0.2 & 1 & 0.1 & 0.1 & 0.4 \\ 4 & 0.1 & 0.1 & 0.1 & 1 & 0.1 & 0.1 \\ 5 & 0.2 & 0.1 & 0.1 & 0.1 & 1 & 0.1 \\ 6 & 0.1 & 0.4 & 0.4 & 0.1 & 0.1 & 1 \end{vmatrix}.$$

The feature vectors obtained by the LBP and ULBP approaches have the forms

$$VB_{58}^{\mathcal{UR}} = \{58, 58, 58, 56, 58, 58, 58, 58, 58, 58, 58, 57, 56, 58, 57, 58\},$$

$$VB_{256}^{\mathcal{UR}} = \{17, 226, 197, 254, 142, 17, 235, 93, 71, 175, 17, 255, 254, 117, 255, 17\},$$

$$VB_{58}^{cg\mathcal{UR}} = \{58, 58, 57, 57, 58, 58, 57, 57, 57, 57, 58, 57, 57, 57, 57, 58\}$$

and

$$VB_{256}^{cg\mathcal{UR}} = \{17, 226, 255, 255, 142, 17, 255, 255, 255, 255, 17, 255, 255, 255, 255, 17\},$$

respectively. The corresponding matricized counterparts are given below

$$B_{58}^{\mathcal{UR}} = \begin{vmatrix} & 1 & 2 & 3 & 4 \\ 1 & 58 & 58 & 58 & 56 \\ 2 & 58 & 58 & 58 & 58 \\ 3 & 58 & 58 & 58 & 57 \\ 4 & 56 & 58 & 57 & 58 \end{vmatrix}, B_{256}^{\mathcal{UR}} = \begin{vmatrix} & 1 & 2 & 3 & 4 \\ 1 & 17 & 142 & 71 & 254 \\ 2 & 226 & 17 & 175 & 117 \\ 3 & 197 & 235 & 17 & 255 \\ 4 & 254 & 93 & 255 & 17 \end{vmatrix}$$

and

$$B_{58}^{cg\mathcal{UR}} = \begin{vmatrix} & 1 & 2 & 3 & 4 \\ 1 & 58 & 58 & 57 & 57 \\ 2 & 58 & 58 & 57 & 57 \\ 3 & 57 & 57 & 58 & 57 \\ 4 & 57 & 57 & 57 & 58 \end{vmatrix}, B_{256}^{cg\mathcal{UR}} = \begin{vmatrix} & 1 & 2 & 3 & 4 \\ 1 & 17 & 142 & 255 & 255 \\ 2 & 226 & 17 & 255 & 255 \\ 3 & 255 & 255 & 17 & 255 \\ 4 & 255 & 255 & 255 & 17 \end{vmatrix},$$

respectively.

For the unthresholded non-square cross-recurrence matrices and their coarse-grained derivatives the above algorithm is identical.

In the current paper, to improve the performance of the RQA methodology, we propose to introduce the several novel LBP-based and ULBP-based RQA descriptors to analyze nonlinear time series. These newly suggested complexity indices are classified as vector-based and matrix-based complexity measures. Namely, the vector-based metrics are extracted from the feature vectors obtained by the action of the (U)LBP algorithm on the arrays given by the formulae (21) and (22) as well as on their coarse-grained analogues. In turn, the matrix-based metrics are extracted from the matricized versions of these vectors. Accordingly, for any feature vector VB instantiated by the cases

$$VB \in \left\{VB_{58}^{\mathcal{UR}}, VB_{256}^{\mathcal{UR}}, VB_{58}^{cg\mathcal{UR}}, VB_{256}^{cg\mathcal{UR}}, VB_{58}^{\mathcal{UCR}}, VB_{256}^{\mathcal{UCR}}, VB_{58}^{cg\mathcal{UCR}}, VB_{256}^{cg\mathcal{UCR}}\right\} \tag{23}$$

and its matricized derivative B instantiated by the cases

$$B \in \left\{B_{58}^{\mathcal{UR}}, B_{256}^{\mathcal{UR}}, B_{58}^{cg\mathcal{UR}}, B_{256}^{cg\mathcal{UR}}, B_{58}^{\mathcal{UCR}}, B_{256}^{\mathcal{UCR}}, B_{58}^{cg\mathcal{UCR}}, B_{256}^{cg\mathcal{UCR}}\right\}, \tag{24}$$

we suggest to consider two types of complexity indices whose formulae are included in Subsections 4.1 and 4.2, respectively.

10.3.1. The vector-based complexity measures

For the feature vector whose elements are indexed by $i \in \{1, 2, \ldots, |VB|\}$, we will test the following complexity descriptors

10.3.1.1. The *arithmetical mean* ($\langle\cdot\rangle$), i.e.,

$$\langle VB\rangle := \frac{1}{|VB|}\sum_{i=1}^{|VB|} VB_i \tag{25}$$

where $\langle\cdot\rangle$ refers to the arithmetical mean and $|VB|$ is the length of the studied vector.

10.3.1.2. The *variance* (Var), i.e.,

$$Var(VB) := \frac{1}{|VB|}\sum_{i=1}^{|VB|}(VB_i - \langle VB\rangle)^2. \tag{26}$$

10.3.1.3. The *coefficient of variation* (CV) also known as the *relative standard deviation*, i.e.,

$$CV(VB) \coloneqq \frac{\sqrt{Var(VB)}}{\langle VB \rangle}. \tag{27}$$

10.3.1.4. The *Gini index* (Gi), i.e.,

$$Gi(VB) \coloneqq \frac{2\sum_{i=1}^{|VB|} iVB_i^*}{|VB|\sum_{i=1}^{|VB|} VB_i} - \frac{|VB|+1}{|VB|} \tag{28}$$

where VB_i^* denotes the vector of the VB_i values arranged in non-decreasing order, i.e., $VB_i \leq VB_{i+1}$.

10.3.1.5. The *structural information content* ($I(VB)$) also known as the *parametric entropy* [16, 19]. To present this index, the subsequent auxiliary notions must be introduced: First, we define the following quantity [16, 19]:

$$p(VB_i) \coloneqq \frac{VB_i}{\sum_{j=1}^{|VB|} VB_j}. \tag{29}$$

In *Chemical Graph Theory*, the quantity expressed by the condition (29) is known as the *information functional* and is identified with a *positive* and *monotonous* function that captures structural information included in a *chemical graph* by defining the finite probability value for each graph node. In this case, the input feature vector is identified with some sequence of the vertex invariants [16, 19]. Next, from the reason that the following equation $p(VB_1) + p(VB_2) + \cdots + p\left(VB_{|VB|}\right) = 1$ is *a priori* valid, it follows that the vector $P(VB) = \left(p(VB_1), p(VB_2), \ldots, p\left(VB_{|VB|}\right)\right)$ constitutes a *finite probability distribution*. Its structural information content is expressed by the following formula [16, 19]:

$$I(VB) \coloneqq -\sum_{i=1}^{|VB|} \frac{VB_i}{\sum_{j=1}^{|VB|} VB_j} \log_2\left(\frac{VB_i}{\sum_{j=1}^{|VB|} VB_j}\right). \tag{30}$$

10.3.1.6. The *mean information content* ($\bar{I}(VB)$) [16, 19]. First, we define the equivalence relation $\simeq$ on the analyzed vector given by the condition: $VB_i \simeq VB_j \leftrightarrow VB_i = VB_j$. Therefore, it is possible to obtain the partition of VB, where the resultant

partitions are symbolized by $VB^1, VB^2, \ldots, VB^h, \ldots, VB^k$ where $k \leq |VB|$. In this context, the quantities $p_h = \frac{|VB^h|}{|VB|}$ for $h \in \{1,2,\ldots,k\}$ can be regarded as probabilities for each singled out partition VB^h. Moreover, it is apparent that $0 \leq p_h \leq 1$ as well as $\sum_{h=1}^{k} p_h = 1$. Accordingly, the vector $P(VB) = (p_1, p_2, \ldots, p_h, \ldots, p_k)$ can be also treated as a *finite probability distribution* over the VB object. Its mean information content is given by the expression [16, 19]:

$$\bar{I}(VB) \coloneqq -\sum_{h=1}^{k} p_h \log_2 p_h. \tag{31}$$

Thus, for any feature vector VB which can be partitioned into k many disjoint subsets of the cardinalities equal to $|VB^h|$ for $h \in \{1,2,\ldots,k\}$, we obtain the following information-theoretical complexity measure [16, 19]:

$$\bar{I}(VB) \coloneqq -\sum_{h=1}^{k} \frac{|VB^h|}{|VB|} \log_2 \left(\frac{|VB^h|}{|VB|} \right). \tag{32}$$

Note that the structural and mean information content complexity indices defined on the graph centrality measures were successfully used to quantify the intricacy of the complex networks with respect to these vertex invariants [19].

10.3.2. The matrix-based complexity measures

For the matrix B whose rows and columns are indexed by $i \in \{1, 2, \ldots, n-2\}$ and by $j \in \{1, 2, \ldots, m-2\}$, respectively, we will test the following matrix-based complexity descriptors

10.3.2.1. The *Frobenius norm* ($\|B\|_F$), i.e.,

$$\|B\|_F \coloneqq \sqrt{tr(B^T B)} \tag{33}$$

where $tr(\cdot)$ refers to the trace function whereas the subscript ^{T} denotes the matrix transposition.

10.3.2.2. The *maximum norm* ($\|B\|_{max}$), i.e.,

$$\|B\|_{max} \coloneqq \max_{i,j} |B_{i.j}|. \tag{34}$$

10.3.2.3. The *maximum absolute column sum norm* ($\|B\|_1$), i.e.,

$$\|B\|_1 \coloneqq \max_j \left(\sum_{i=1}^{n-2} |B_{i,j}| \right). \tag{35}$$

10.3.2.4. The *maximum absolute row sum norm* ($\|B\|_\infty$), i.e.,

$$\|B\|_\infty \coloneqq \max_i \left(\sum_{j=1}^{m-2} |B_{i,j}| \right). \tag{36}$$

10.3.2.5. The *spectral norm* ($\|B\|_2$), i.e.,

$$\|B\|_2 \coloneqq \sigma_{max}(B) \tag{37}$$

where σ_{max} is the largest singular value of B.

10.3.2.6. The *rank* ($rk(B)$), i.e.,

$$rk(B) \coloneqq \text{the number of non} - \text{zero singular values.} \tag{38}$$

The performance of the above descriptors will be quantitatively assessed in the time series classification task experiments presented in Section 6.

10.4. The datasets and computational methodology

In the pharmacological time series analysis, the use of time-dependent measurements of the cell proliferation can relate drug-responses to cell phenotypes. The HTS007 dataset downloaded from https://demo.thunor.net includes a panel of eight breast cancer cell lines treated with 27 anticancer drugs at several different concentrations. The cell proliferation was quantified over 5 days [9]. All time series contained in the above dataset consist of 24 measured time points and they have two class labels: control and non-control. In our computational experiments, we selected eight dose-response time series corresponding to the following anticancer drugs: *abemaciclib*, *alpelisib*, *bleomycin*, *cediranib*, *dasatinib*, *etoposide*, *everolimus* and *ipatasertib* whose efficiency was tested on the BT20 cell line. In sum, our numerical experiments were performed on 393 time-dependent cell proliferation measurements. Moreover, for a comparative purpose, we tested our improved RQA methodology on the

ECG200 dataset composed of 200 time series. This dataset was downloaded from https://www.timeseriesclassification.com. Each time series in this dataset contains 96 measured time points. All time series contained in the ECG200 dataset have also two class labels: normal heartbeat and myocardial infarction. In the numerical classification experiments, we perform the 10-fold cross validation using *1-nearest neighbors classifier* (1-NN classifier). In the traditional RQA experiments, we followed the proposal of J.P. Zbilut and coworkers and we set the threshold distance ε such that the RR coefficient is approximately equal to 1% (cf. Table 2) [21]. Moreover, the RQA parameters [10], i.e., l_{min}, v_{min}, the *embedding dimension* and the *time lag* were set to two, two, one and to zero, respectively. All computations were performed in the R programming language [4, 6, 7, 13–15, 18].

10.5. Results and discussion

The LPB and ULPB algorithms were run on the matrices (21) and (22). The matrix (21) was obtained from the raw time series whereas the matrix (22) was obtained from the raw time series and its first differenced variant. In our numerical classification experiments, we try to predict the class labels of the time series contained in the HTS007 and ECG200 datasets. The results of the performance of the 1-NN classifier, i.e., the number of correctly classified time series are recorded in percentage (scaled to the range from zero to one). In the tables below, the names of time series are identified with the names of tested anticancer drugs. In all cases, the best results are in bold.

Table 2

The RQA results for five time series

INDEX	ABEMACICLIB	ALPELISIB	BLEOMYCIN	CEDIRANIB	DASATINIB
RR	**0.784**	0.674	0.608	0.625	0.562
$RATIO$	0.765	**0.696**	0.549	0.688	0.646
DET	0.647	0.435	0.471	0.604	0.562
DIV	0.689	0.674	**0.647**	0.688	0.667
L_{max}	0.745	0.674	**0.647**	0.708	0.667
L_{mean}	0.686	0.652	0.451	0.562	0.583
L^{-}_{mean}	0.6	0.522	0.471	0.667	0.604
$ENTR$	0.725	0.565	0.608	0.688	**0.771**
$TREND$	0.667	0.543	**0.647**	**0.729**	0.542
LAM	0.667	**0.696**	0.588	0.667	0.5
V_{max}	0.725	**0.696**	0.627	0.708	**0.771**
V_{mean}	0.706	0.587	0.549	0.688	0.604

Table 3

The RQA results for four time series

INDEX	ETOPOSIDE	EVEROLIMUS	IPATASERTIB	ECG
RR	0.64	0.66	**0.612**	0.575
$RATIO$	0.64	0.62	**0.612**	0.56
DET	0.48	0.6	0.51	0.595
DIV	**0.72**	**0.72**	0.592	0.64
L_{max}	0.7	0.7	0.551	0.615
L_{mean}	0.54	0.54	0.429	0.585
L_{mean}^{-}	0.64	0.66	0.469	**0.67**
$ENTR$	0.6	0.6	0.51	0.59
$TREND$	0.6	0.62	0.571	0.6
LAM	0.58	0.62	0.571	0.59
V_{max}	0.66	0.66	**0.612**	0.66
V_{mean}	**0.72**	**0.72**	0.551	0.63

Table 4

The (U)LBP-based RQA results for Abemaciclib using the vector-based descriptors

Index	$VB_{58}^{\mathcal{UR}}$	$VB_{256}^{\mathcal{UR}}$	$VB_{58}^{cg\mathcal{UR}}$	$VB_{256}^{cg\mathcal{UR}}$	$VB_{58}^{\mathcal{UCR}}$	$VB_{256}^{\mathcal{UCR}}$	$VB_{58}^{cg\mathcal{UCR}}$	$VB_{256}^{cg\mathcal{UCR}}$
$\langle VB \rangle$	0.667	0.608	0.667	0.745	0.765	0.608	0.608	0.667
$Var(VB)$	0.569	0.745	0.608	0.804	0.706	0.667	0.412	0.706
$CV(VB)$	0.471	0.647	0.627	**0.843**	0.667	0.706	0.569	0.784
$Gi(VB)$	0.608	0.608	0.686	0.745	0.647	0.706	0.627	0.667
$I(VB)$	0.451	0.627	0.706	0.745	0.706	0.608	0.471	0.725
$\bar{I}(VB)$	0.706	0.706	0.608	0.824	0.647	0.765	0.667	0.392

Table 5

The (U)LBP-based RQA results for Abemaciclib using the matrix-based descriptors

Index	$B_{58}^{\mathcal{UR}}$	$B_{256}^{\mathcal{UR}}$	$B_{58}^{cg\mathcal{UR}}$	$B_{256}^{cg\mathcal{UR}}$	$B_{58}^{\mathcal{UCR}}$	$B_{256}^{\mathcal{UCR}}$	$B_{58}^{cg\mathcal{UCR}}$	$B_{256}^{cg\mathcal{UCR}}$
$\|B\|_F$	0.647	0.706	0.706	0.784	0.49	0.686	0.608	0.706
$\|B\|_2$	0.588	0.569	0.647	0.784	0.824	0.549	0.49	0.686
$\|B\|_{max}$	0.608	0.824	0.608	0.608	0.608	0.765	0.608	0.608
$\|B\|_\infty$	0.627	0.549	0.608	0.686	0.725	0.667	0.608	0.647
$\|B\|_1$	0.569	0.745	0.471	0.549	0.667	0.686	0.608	0.647
$rk(B)$	0.608	0.608	0.608	0.608	0.725	0.745	0.49	0.49

Table 6

The (U)LBP-based RQA results for Alpelisib using the vector-based descriptors

Index	$VB_{58}^{\mathcal{UR}}$	$VB_{256}^{\mathcal{UR}}$	$VB_{58}^{cg\mathcal{UR}}$	$VB_{256}^{cg\mathcal{UR}}$	$VB_{58}^{\mathcal{UCR}}$	$VB_{256}^{\mathcal{UCR}}$	$VB_{58}^{cg\mathcal{UCR}}$	$VB_{256}^{cg\mathcal{UCR}}$
$\langle VB \rangle$	0.565	0.522	0.543	0.5	0.609	0.587	**0.739**	0.5
$Var(VB)$	0.478	0.565	0.587	0.565	0.457	0.652	0.543	0.674
$CV(VB)$	0.522	0.478	0.587	0.63	0.609	0.587	0.565	0.717
$Gi(VB)$	0.391	0.457	0.63	0.522	0.587	0.5	0.522	0.5
$I(VB)$	0.543	0.522	0.587	0.587	0.5	0.5	0.5	0.63
$\bar{I}(VB)$	0.630	0.652	0.478	0.565	0.565	0.587	0.543	0.565

Table 7

The (U)LBP-based RQA results for Alpelisib using the matrix-based descriptors

Index	$B_{58}^{\mathcal{UR}}$	$B_{256}^{\mathcal{UR}}$	$B_{58}^{cg\mathcal{UR}}$	$B_{256}^{cg\mathcal{UR}}$	$B_{58}^{\mathcal{UCR}}$	$B_{256}^{\mathcal{UCR}}$	$B_{58}^{cg\mathcal{UCR}}$	$B_{256}^{cg\mathcal{UCR}}$
$\|B\|_F$	0.587	0.478	0.717	0.587	0.522	0.587	0.63	0.543
$\|B\|_2$	0.63	0.565	0.5	0.609	0.543	**0.739**	0.609	0.543
$\|B\|_{max}$	0.674	0.717	0.674	0.674	0.674	0.522	0.674	0.674
$\|B\|_\infty$	0.609	0.609	0.522	0.652	0.63	0.587	0.674	0.674
$\|B\|_1$	0.652	0.63	0.543	0.609	0.543	0.37	0.609	0.609
$rk(B)$	0.674	0.674	0.674	0.674	0.609	0.63	0.609	0.609

Table 8

The (U)LBP-based RQA results for Bleomycin using the vector-based descriptors

Index	$VB_{58}^{\mathcal{UR}}$	$VB_{256}^{\mathcal{UR}}$	$VB_{58}^{cg\mathcal{UR}}$	$VB_{256}^{cg\mathcal{UR}}$	$VB_{58}^{\mathcal{UCR}}$	$VB_{256}^{\mathcal{UCR}}$	$VB_{58}^{cg\mathcal{UCR}}$	$VB_{256}^{cg\mathcal{UCR}}$
$\langle VB \rangle$	0.451	0.49	0.549	0.51	0.608	0.529	0.627	0.725
$Var(VB)$	0.627	0.588	0.392	**0.765**	0.588	0.471	0.627	0.667
$CV(VB)$	0.49	0.529	0.608	0.608	0.667	0.529	0.451	0.706
$Gi(VB)$	0.51	0.608	0.647	0.627	0.608	0.51	0.549	0.608
$I(VB)$	0.549	0.549	0.627	0.431	0.549	0.529	0.451	0.549
$\bar{I}(VB)$	0.529	0.529	0.529	0.549	0.333	0.471	0.431	0.471

Table 9

The (U)LBP-based RQA results for Bleomycin using the matrix-based descriptors

Index	$B_{58}^{\mathcal{UR}}$	$B_{256}^{\mathcal{UR}}$	$B_{58}^{cg\mathcal{UR}}$	$B_{256}^{cg\mathcal{UR}}$	$B_{58}^{\mathcal{UCR}}$	$B_{256}^{\mathcal{UCR}}$	$B_{58}^{cg\mathcal{UCR}}$	$B_{256}^{cg\mathcal{UCR}}$
$\|B\|_F$	0.549	0.51	0.569	0.647	0.529	0.51	0.647	0.608
$\|B\|_2$	0.627	0.647	0.627	0.667	0.627	0.451	0.627	0.529
$\|B\|_{max}$	0.608	0.647	0.608	0.608	0.608	0.608	0.608	0.608
$\|B\|_\infty$	0.51	0.588	0.588	0.569	0.588	0.667	0.608	0.647
$\|B\|_1$	0.608	0.471	0.49	0.588	0.529	0.451	0.431	0.588
$rk(B)$	0.608	0.608	0.608	0.608	0.725	0.686	0.51	0.451

Table 10

The (U)LBP-based RQA results for Cediranib using the vector-based descriptors

Index	$VB_{58}^{\mathcal{UR}}$	$VB_{256}^{\mathcal{UR}}$	$VB_{58}^{cg\mathcal{UR}}$	$VB_{256}^{cg\mathcal{UR}}$	$VB_{58}^{\mathcal{UCR}}$	$VB_{256}^{\mathcal{UCR}}$	$VB_{58}^{cg\mathcal{UCR}}$	$VB_{256}^{cg\mathcal{UCR}}$
$\langle VB \rangle$	0.562	0.604	0.604	0.521	0.646	0.583	0.646	0.604
$Var(VB)$	0.729	**0.792**	0.542	0.729	0.583	0.5	0.625	**0.792**
$CV(VB)$	0.688	0.646	0.688	0.646	0.729	0.583	0.604	0.729
$Gi(VB)$	0.771	0.604	0.792	0.562	0.562	0.542	0.479	0.625
$I(VB)$	0.646	0.646	0.75	0.542	0.729	0.542	0.583	0.562
$\bar{I}(VB)$	0.562	0.542	0.667	0.5	0.5	0.562	0.688	0.458

Table 11
The (U)LBP-based RQA results for Cediranib using the matrix-based descriptors

Index	$B_{58}^{\mathcal{UR}}$	$B_{256}^{\mathcal{UR}}$	$B_{58}^{cg\mathcal{UR}}$	$B_{256}^{cg\mathcal{UR}}$	$B_{58}^{\mathcal{UCR}}$	$B_{256}^{\mathcal{UCR}}$	$B_{58}^{cg\mathcal{UCR}}$	$B_{256}^{cg\mathcal{UCR}}$
$\|B\|_F$	0.604	0.625	0.625	0.646	0.5	0.458	0.646	0.688
$\|B\|_2$	0.646	0.708	0.667	0.667	0.625	0.5	0.438	0.583
$\|B\|_{max}$	0.646	0.75	0.646	0.646	0.646	0.688	0.646	0.646
$\|B\|_\infty$	0.708	0.604	0.583	0.75	0.458	0.625	0.688	0.708
$\|B\|_1$	0.604	0.562	0.438	0.688	0.625	0.542	0.542	0.562
$rk(B)$	0.646	0.646	0.646	0.646	0.667	0.604	0.75	0.729

Table 12
The (U)LBP-based RQA results for Dasatinib using the vector-based descriptors

Index	$VB_{58}^{\mathcal{UR}}$	$VB_{256}^{\mathcal{UR}}$	$VB_{58}^{cg\mathcal{UR}}$	$VB_{256}^{cg\mathcal{UR}}$	$VB_{58}^{\mathcal{UCR}}$	$VB_{256}^{\mathcal{UCR}}$	$VB_{58}^{cg\mathcal{UCR}}$	$VB_{256}^{cg\mathcal{UCR}}$
$\langle VB \rangle$	0.604	0.688	0.646	0.458	0.729	0.667	0.562	0.542
$Var(VB)$	0.729	0.708	0.5	0.625	0.562	0.604	0.417	0.667
$CV(VB)$	0.562	0.688	0.604	0.562	0.604	0.667	0.438	0.75
$Gi(VB)$	0.646	0.562	0.479	0.5	0.521	0.562	0.438	0.479
$I(VB)$	0.667	0.688	0.667	0.583	0.646	0.521	0.521	0.583
$\bar{I}(VB)$	0.562	0.625	0.604	0.646	0.521	0.625	0.625	0.562

Table 13
The (U)LBP-based RQA results for Dasatinib using the matrix-based descriptors

Index	$B_{58}^{\mathcal{UR}}$	$B_{256}^{\mathcal{UR}}$	$B_{58}^{cg\mathcal{UR}}$	$B_{256}^{cg\mathcal{UR}}$	$B_{58}^{\mathcal{UCR}}$	$B_{256}^{\mathcal{UCR}}$	$B_{58}^{cg\mathcal{UCR}}$	$B_{256}^{cg\mathcal{UCR}}$
$\|B\|_F$	0.625	0.562	0.583	0.583	0.5	0.625	0.562	0.5
$\|B\|_2$	0.562	0.5	0.542	0.542	0.562	0.542	0.396	0.583
$\|B\|_{max}$	0.646	0.771	0.646	0.646	0.646	0.729	0.646	0.646
$\|B\|_\infty$	0.604	0.583	0.625	0.604	**0.854**	0.562	0.604	0.729
$\|B\|_1$	0.604	0.542	0.625	0.625	0.417	0.667	0.562	0.542
$rk(B)$	0.646	0.646	0.646	0.646	0.625	0.625	0.625	0.625

Table 14
The (U)LBP-based RQA results for Etoposide using the vector-based descriptors

Index	$VB_{58}^{\mathcal{UR}}$	$VB_{256}^{\mathcal{UR}}$	$VB_{58}^{cg\mathcal{UR}}$	$VB_{256}^{cg\mathcal{UR}}$	$VB_{58}^{\mathcal{UCR}}$	$VB_{256}^{\mathcal{UCR}}$	$VB_{58}^{cg\mathcal{UCR}}$	$VB_{256}^{cg\mathcal{UCR}}$
$\langle VB \rangle$	0.64	0.5	0.44	0.56	0.6	0.56	0.64	0.54
$Var(VB)$	0.62	0.66	0.56	0.56	0.52	0.52	0.46	0.58
$CV(VB)$	0.62	0.54	0.6	0.56	0.5	0.68	0.52	0.56
$Gi(VB)$	0.78	0.7	0.54	0.5	0.6	0.58	0.5	0.48
$I(VB)$	0.62	0.52	0.68	0.56	0.5	0.54	0.38	0.66
$\bar{I}(VB)$	0.42	0.4	0.64	0.58	0.52	0.7	0.54	0.48

Table 15
The (U)LBP-based RQA results for Etoposide using the matrix-based descriptors

Index	$B_{58}^{\mathcal{UR}}$	$B_{256}^{\mathcal{UR}}$	$B_{58}^{cg\mathcal{UR}}$	$B_{256}^{cg\mathcal{UR}}$	$B_{58}^{\mathcal{UCR}}$	$B_{256}^{\mathcal{UCR}}$	$B_{58}^{cg\mathcal{UCR}}$	$B_{256}^{cg\mathcal{UCR}}$
$\|B\|_F$	0.64	0.5	0.52	0.56	0.4	0.6	0.5	0.62
$\|B\|_2$	0.6	0.5	0.5	0.54	0.7	0.62	0.34	0.64
$\|B\|_{max}$	0.62	0.7	0.62	0.62	0.62	0.44	0.62	0.62
$\|B\|_\infty$	0.66	0.42	0.58	0.56	0.68	0.62	0.62	0.62
$\|B\|_1$	0.6	0.68	0.44	0.48	0.46	0.62	0.58	0.5
$rk(B)$	0.62	0.62	0.64	0.64	0.78	**0.8**	0.52	0.48

Table 16
The (U)LBP-based RQA results for Everolimus using the vector-based descriptors

Index	$VB_{58}^{\mathcal{UR}}$	$VB_{256}^{\mathcal{UR}}$	$VB_{58}^{cg\mathcal{UR}}$	$VB_{256}^{cg\mathcal{UR}}$	$VB_{58}^{\mathcal{UCR}}$	$VB_{256}^{\mathcal{UCR}}$	$VB_{58}^{cg\mathcal{UCR}}$	$VB_{256}^{cg\mathcal{UCR}}$
$\langle VB \rangle$	0.66	0.5	0.42	0.56	0.6	0.56	0.62	0.54
$Var(VB)$	0.62	0.66	0.56	0.56	0.52	0.52	0.46	0.58
$CV(VB)$	0.62	0.54	0.6	0.56	0.5	0.68	0.52	0.56
$Gi(VB)$	0.78	0.7	0.54	0.5	0.6	0.58	0.5	0.48
$I(VB)$	0.62	0.52	0.68	0.56	0.5	0.54	0.38	0.66
$\bar{I}(VB)$	0.42	0.4	0.64	0.58	0.52	0.7	0.54	0.48

Table 17
The (U)LBP-based RQA results for Everolimus using the matrix-based descriptors

Index	$B_{58}^{\mathcal{UR}}$	$B_{256}^{\mathcal{UR}}$	$B_{58}^{cg\mathcal{UR}}$	$B_{256}^{cg\mathcal{UR}}$	$B_{58}^{\mathcal{UCR}}$	$B_{256}^{\mathcal{UCR}}$	$B_{58}^{cg\mathcal{UCR}}$	$B_{256}^{cg\mathcal{UCR}}$
$\|B\|_F$	0.64	0.5	0.52	0.56	0.4	0.6	0.5	0.62
$\|B\|_2$	0.6	0.5	0.5	0.54	0.7	0.62	0.34	0.64
$\|B\|_{max}$	0.62	0.7	0.62	0.62	0.62	0.44	0.62	0.62
$\|B\|_\infty$	0.62	0.38	0.54	0.5	0.66	0.6	0.6	0.62
$\|B\|_1$	0.6	0.72	0.54	0.48	0.42	0.64	0.56	0.54
$rk(B)$	0.62	0.62	0.64	0.64	0.8	**0.82**	0.56	0.52

Table 18
The (U)LBP-based RQA results for Ipatasertib using the vector-based descriptors

Index	$VB_{58}^{\mathcal{UR}}$	$VB_{256}^{\mathcal{UR}}$	$VB_{58}^{cg\mathcal{UR}}$	$VB_{256}^{cg\mathcal{UR}}$	$VB_{58}^{\mathcal{UCR}}$	$VB_{256}^{\mathcal{UCR}}$	$VB_{58}^{cg\mathcal{UCR}}$	$VB_{256}^{cg\mathcal{UCR}}$
$\langle VB \rangle$	0.449	0.367	0.592	0.592	0.612	0.633	0.714	0.51
$Var(VB)$	0.592	0.531	0.408	0.633	0.571	0.653	0.449	0.612
$CV(VB)$	0.571	0.429	0.571	0.612	0.347	0.531	0.469	0.592
$Gi(VB)$	0.612	0.49	0.449	0.612	0.592	0.571	0.571	0.551
$I(VB)$	0.551	0.429	0.449	0.49	0.51	0.49	0.388	0.571
$\bar{I}(VB)$	0.551	0.551	0.612	0.571	0.429	0.531	0.469	0.449

Table 19

The (U)LBP-based RQA results for Ipatasertib using the matrix-based descriptors

Index	$B_{58}^{\mathcal{UR}}$	$B_{256}^{\mathcal{UR}}$	$B_{58}^{cg\mathcal{UR}}$	$B_{256}^{cg\mathcal{UR}}$	$B_{58}^{\mathcal{UCR}}$	$B_{256}^{\mathcal{UCR}}$	$B_{58}^{cg\mathcal{UCR}}$	$B_{256}^{cg\mathcal{UCR}}$
$\|B\|_F$	0.429	0.408	0.633	0.653	0.49	0.714	0.531	0.714
$\|B\|_2$	0.531	0.429	0.551	0.449	0.449	0.531	0.429	0.571
$\|B\|_{max}$	0.633	**0.735**	0.633	0.633	0.633	0.673	0.633	0.633
$\|B\|_\infty$	0.551	0.551	0.612	0.551	0.551	0.653	0.592	0.633
$\|B\|_1$	0.633	0.612	0.49	0.551	0.449	0.449	0.673	0.653
$rk(B)$	0.633	0.633	0.633	0.633	0.531	0.551	0.571	0.49

Table 20

The (U)LBP-based RQA results for ECG200 using the vector-type descriptors

Index	$VB_{58}^{\mathcal{UR}}$	$VB_{256}^{\mathcal{UR}}$	$VB_{58}^{cg\mathcal{UR}}$	$VB_{256}^{cg\mathcal{UR}}$	$VB_{58}^{\mathcal{UCR}}$	$VB_{256}^{\mathcal{UCR}}$	$VB_{58}^{cg\mathcal{UCR}}$	$VB_{256}^{cg\mathcal{UCR}}$
$\langle VB \rangle$	0.565	0.59	0.61	0.645	0.61	0.52	0.54	0.61
$Var(VB)$	0.605	0.615	0.61	0.66	**0.74**	0.635	0.63	0.605
$CV(VB)$	0.62	0.64	0.6	0.61	0.6	0.58	0.62	0.64
$Gi(VB)$	0.69	0.57	0.58	0.635	0.58	0.61	0.6	0.73
$I(VB)$	0.655	0.675	0.59	0.595	0.61	0.635	0.585	0.64
$\bar{I}(VB)$	0.655	0.6	0.59	0.625	0.62	0.55	0.54	0.585

Table 21

The (U)LBP-based RQA results for ECG200 using the matrix-type descriptors

Index	$B_{58}^{\mathcal{UR}}$	$B_{256}^{\mathcal{UR}}$	$B_{58}^{cg\mathcal{UR}}$	$B_{256}^{cg\mathcal{UR}}$	$B_{58}^{\mathcal{UCR}}$	$B_{256}^{\mathcal{UCR}}$	$B_{58}^{cg\mathcal{UCR}}$	$B_{256}^{cg\mathcal{UCR}}$
$\|B\|_F$	0.62	0.51	0.55	0.595	0.675	0.585	0.515	0.605
$\|B\|_2$	0.61	0.57	0.59	0.61	0.645	0.535	0.62	0.645
$\|B\|_{max}$	0.665	0.665	0.665	0.665	0.665	0.665	0.665	0.665
$\|B\|_\infty$	0.72	0.59	0.665	0.56	0.645	0.58	0.605	0.61
$\|B\|_1$	0.6	0.615	0.54	0.62	0.585	0.625	0.535	0.535
$rk(B)$	0.66	0.66	0.66	0.66	0.64	0.67	0.68	0.68

Table 22

The parameters of the traditional RQA algorithm
and the improvement attained by the novel (U)LBP descriptors

Time series	Radius	Mean RR coefficient	Improvement
Abemaciclib	0.07	0.1	6.7 %
Alpelisib	0.13	0.1	5.82 %
Bleomycin	0.13	0.11	15.42 %
Cediranib	0.11	0.1	7.95 %
Dasatinib	0.12	0.1	9.72 %
Etoposide	0.11	0.1	10 %
Everolimus	0.11	0.1	12.2 %
Ipatasertib	0.13	0.11	16.73 %
ECG200	0.11	0.11	9.46 %

From the above results, it follows that the newly proposed ULBP-based and LBP-based complexity descriptors are more or significantly more efficient in the classification tasks of pharmacokinetical and physiological data than their original counterparts. This regularity is observed for the relatively short dose-response time series as well as for the longer ECG time series. Specifically, in our numerical experiments, the improvement achieved by the newly suggested (U)LBP-based indices ranges from 5.82% to 16.73%.

In the second measurement, we perform the *principal component analysis* (PCA) [5] on the ECG200 dataset in order to visualize the mutual relationships between the tested complexity descriptors. In Fig. 6.1 and 6.2, the analyzed indices are labeled as follows: 1 – RR, 2 – $RATIO$, 3 – DET, 4 – DIV, 5 – L_{max}, 6 – L_{mean}, 7 – L^{-}_{mean}, 8 – $ENTR$, 9 – $TREND$, 10 – LAM, 11 – V_{max}, 12 – V_{mean}, 13 – $\langle VB \rangle$, 14 – $Var(VB)$, 15 – $CV(VB)$, 16 – $Gi(VB)$, 17 – $I(VB)$, 18 – $\bar{I}(VB)$, 19 – $\|B\|_F$, 20 – $\|B\|_2$, 21 – $\|B\|_{max}$, 22 – $\|B\|_\infty$, 23 – $\|B\|_1$, 24 – $rk(B)$ where the instantiations of the VB vector and the B matrix are listed by the conditions (23) and (24), respectively. Moreover, in both .s, the complexity descriptors derived from the coarse-grained derivative arrays are labeled as cgULBP or as cgLBP.

From eight PCA projections presented in Fig. 1 and 2, it follows that all standard RQA complexity measures are concentrated almost in one point. Also, the PCA projections from Fig. 1 demonstrate that four matrix-based descriptors, i.e., $\|B\|_F$, $\|B\|_2$, $\|B\|_\infty$ and $\|B\|_1$ as well as one vector-based index, i.e., $Var(VB)$ carry very specific information. Moreover, from four two-dimensional projections presented in Fig. 2, it can be seen that two vector-based central tendency metrics, i.e., $\langle VB \rangle$ and $Var(VB)$ as well as two vector-based information-theoretical metrics, i.e., $Gi(VB)$ and $I(VB)$ together with one matrix-based coefficient, i.e., $\|B\|_2$ do not duplicate the structural information contained in the original RQA metrics.

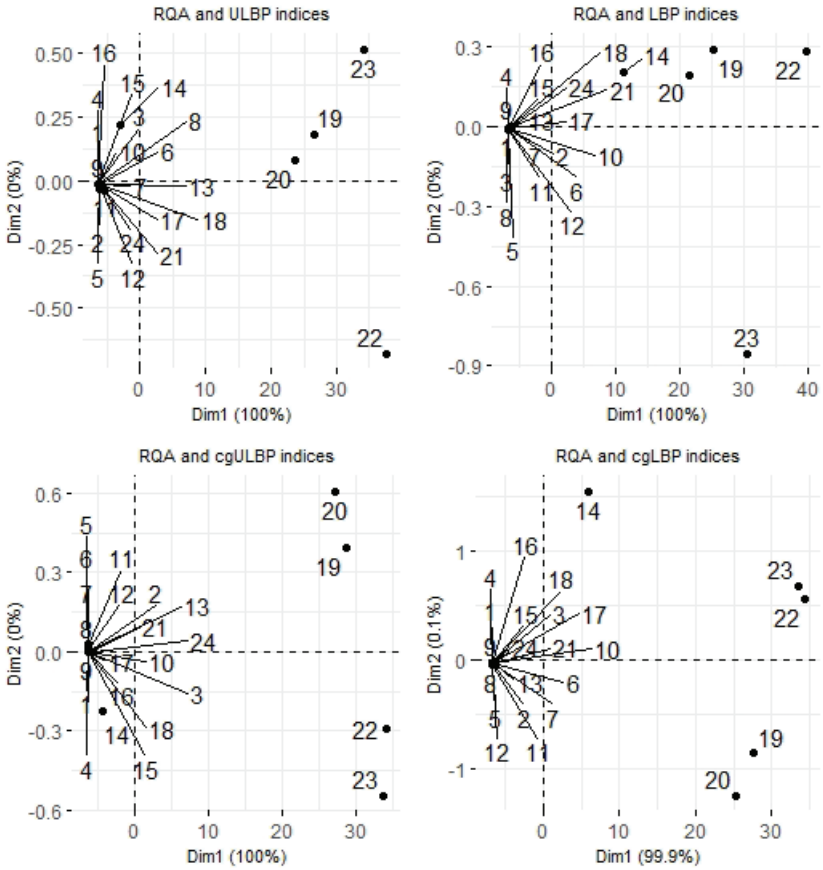

Fig. 1. The PCA projection results of the novel (U)LBP-based descriptors derived from the matrix (21)

Rys. 1. Wyniki analizy głównych składowych nowych (U)LBP współczynników otrzymanych z macierzy nr 21

Accordingly, it can be concluded that the collection of the newly suggested (U)LBP-based RQA complexity descriptors is overwhelmingly more informative than the set of the original RQA indices. In our opinion, the above phenomena can (to some extent) elucidate the enhanced performance of the novel (U)LBP complexity descriptors in comparison with the standard RQA metrics.

10.6. Conclusions

Recent pharmacological and biochemical technologies have permitted scholars to collect a large number of pharmacokinetical data in the medical domain. For this reason, the development of reliable methods for automated analysis of pharmacologically and physiologically important time series extracted from medical databases is a very important research topic. Moreover, in the recent years, several authors tried to use the LBP methodology to study medical images [11, 17, 20]. The

present contribution proposes to apply the local binary pattern algorithm as well as its uniform variant to analyze the physiologically and pharmacologically significant time series. Our preliminary results demonstrated that the LBP and ULBP methodologies applied to the unthresholded recurrence and cross-recurrence matrices are very promising as sensitive metrics enabling to characterize the structural patterns in the medical data under consideration.

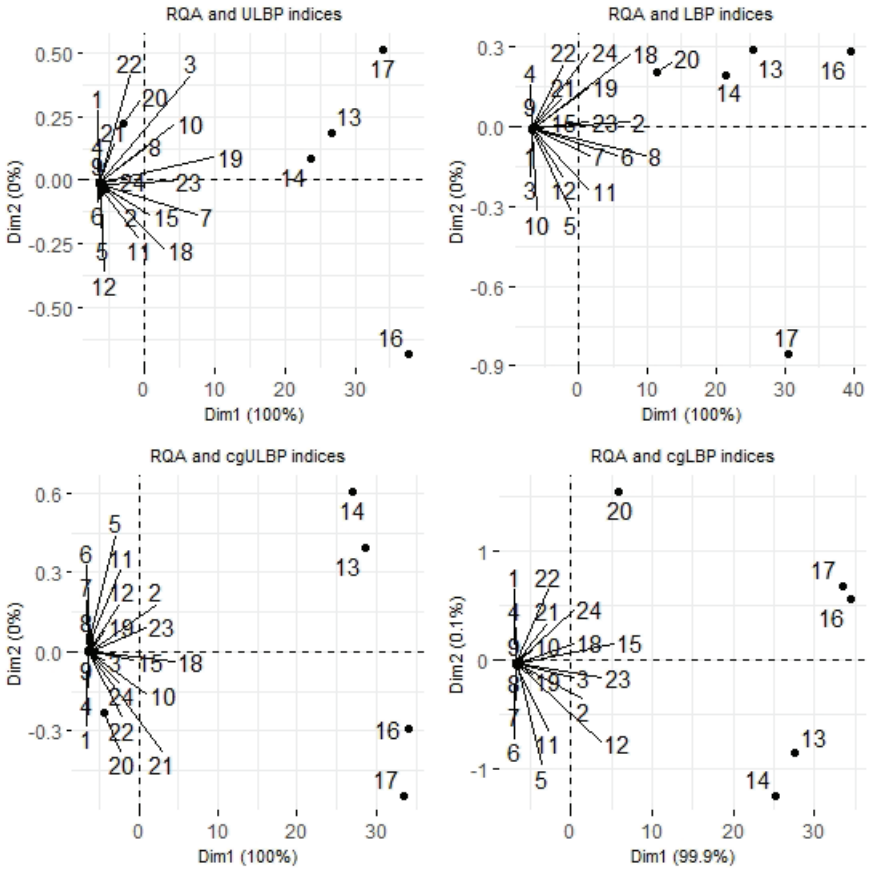

Fig. 2. The PCA projection results of the novel (U)LBP-based descriptors derived from the matrix (22)

Rys. 2. Wyniki analizy głównych składowych nowych (U)LBP współczynników otrzymanych z macierzy nr 22

Therefore, it can be conjectured that not only medical images but also medical numerical data can be analyzed by the (U)LBP techniques.

Bibliography

1. S. Brahnam, L.C. Jain, L. Nanni, A. Lumini: Local Binary Patterns: New Variants and Applications, Springer – Verlag, Berlin, Heidelberg 2014.
2. J.-P. Eckmann, S.O. Kamphorst, D. Ruelle: Recurrence plots of dynamical systems, Europhys. Lett. (1987) **5**: 973–977.

3. D. Eroglu, T.K. DM. Peron, N. Marwan, F.A. Rodrigues, L. da F. Costa, M. Sebek, I.Z. Kiss, J. Kurths: Entropy of weighted recurrence plots, Phys. Rev. E (2014) **90**: 042919.
4. C.A. Garcia, nonlinearTseries: Nonlinear Time Series Analysis. R package version 0.2.11. (2021), https://CRAN.R-project.org/package=nonlinearTseries.
5. R.A. Johnson, D.W. Wichern: Applied Multivariate Statistical Analysis, Pearson Education, Inc., Upper Saddle River 2007.
6. A. Kassambara, ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0 (2020), https://CRAN.R-project.org/package=ggpubr
7. A. Kassambara, F. Mundt, factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. (2020), https://CRAN. R- project. org/package=factoextra
8. E.J. Leavline, D.A.A.G. Singh, P. Maheswari: Local binary pattern family descriptors for texture classification, IJIGSP (2018) **10**: 40–45.
9. A.L.R. Lubbock, L.A. Harris, V. Quaranta, D.R. Tyson, C.F. Lopez, Thunor: visualization and analysis of high-throughput dose-response datasets, Nucleic Acid Res. (2021) **49**: W633-W640.
10. N. Marwan, M.C. Romano, M. Thiel, J. Kurths: Recurrence plots for the analysis of complex systems, Phys. Rep. (2007) **438**: 237–329.
11. L. Nanni, A. Lumini, S. Brahnam: Local binary patterns variants as texture descriptors for medical image analysis, Artif. Intell. Med. (2010) **49**: 117–125.
12. T. Ojala, M. Pietikäinen, D. Harwood: Performance evaluation of texture measures with classification based on Kullback discrimination of distributions, Proceedings of the 12th IAPR International Conference on Pattern Recognition vol. 1 (1994): 582–585.
13. R Core Team, R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2020), https://www.R-project.org
14. K. Slowikowski, ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'. R package version 0.8.2. (2020), https://CRAN.R-project.org/ package=ggrepel
15. J. Sugiyama, K. Kobayashi, wvtool: Image Tools for Automated Wood Identification. R package version 1.0. (2016), https://CRAN.R-project.org/package=wvtool
16. R. Todeschini, V. Consonni, *Molecular Descriptors for Chemoinformatics*, *Volume I & II*, Wiley – VCH Verlag GmbH & Co. KGaA, Weinheim 2009.

17. O.A. Vătămanu, M. Ionescu, G.-I. Mihalaş: Analysis and classification of ultrasound medical images using the Local Binary Pattern operator, Stud. Health Technol. Inform. (2013) **190**:175–178.
18. H. Wickham, ggplot2: Elegant Graphics for Data Analysis, Springer – Verlag, New York 2016, https://ggplot2.tidyverse.org
19. P. Wilczek: Novel centrality measures and distance-related topological indices in network data mining, Silesian J. Pure Appl. Math. (2017) **7**, 21–63.
20. X. Xianchuan, Q. Zhang: Medical image retrieval using local binary patterns with image Euclidean distance, 2009 International Conference on Information Engineering and ComputerScience (2009): 1–4.
21. J.P. Zbilut, J.-M. Zaldívar-Comenges, F. Strozzi: Recurrence quantification based Liapunov exponents for monitoring divergence in experimental data, Phys. Lett. A (2002) **297** (3–4): 173–181.

AN APPLICATION OF THE LOCAL BINARY PATTERN ALGORITHM AND ITS UNIFORM VARIANT TO IMPROVE THE RECURRENCE AND CROSS-RECURRENCE QUANTIFICATION ANALYSES OF THE PHARMACOLOGICALLY AND PHYSIOLOGICALLY IMPORTANT TIME SERIES

Abstract

The *recurrence* and *cross-recurrence matrices* are recognized as a very efficient tool allowing to visualize the recurrences of nonlinear dynamical systems. The complexity of the structural patterns detected in these matrices is quantified by the RQA descriptors. The purpose of the present work was to introduce several novel RQA complexity descriptors based on the *local binary pattern* (LBP) operator and its *uniform* version (ULBP). These operators label the pixels of an image by thresholding the neighborhood of each pixel and report the results as binary numbers. In our approach, we consider the (*coarse-grained*) *unthresholded* recurrence and cross--recurrence matrices as images on which the LBP and ULBP operators can be defined. The usefulness of the newly suggested RQA complexity descriptors in the *time series classification tasks* is tested on the selected 393 time-dependent cell proliferation measurements extracted from the (publicly available) HTS007 dataset as well as on one (also publicly available) ECG dataset containing 200 time series. In all exemplary

cases, our results showed that, in comparison to the traditional RQA indices, the classification accuracy attained by the novel (U)LBP-based RQA descriptors is higher or significantly higher. Moreover, by using the *multivariate technique* (PCA), we demonstrated that some of the newly proposed complexity descriptors carry very specific structural information which can not be obtained by means of the original RQA procedure. Accordingly, it can be conjectured that, in the realm of pharmacokinetics and physiology, the novel (U)LPB-based RQA complexity indices will be regarded as a valuable statistical tool in analyzing nonlinear time series.

Keywords: nonlinear time series, pharmacokinetics, RQA methodology, LBP and ULBP algorithms