

COMPARISON OF THE INFLUENCE OF STANDARDIZATION AND NORMALIZATION OF DATA ON THE EFFECTIVENESS OF SPONGY TISSUE TEXTURE CLASSIFICATION

Róża Dzierżak

Lublin University of Technology, Institute of Electronics and Information Technology, Lublin, Poland

Abstract. The aim of this article was to compare the influence of the data pre-processing methods – normalization and standardization – on the results of the classification of spongy tissue images. Four hundred CT images of the spine (L1 vertebra) were used for the analysis. The images were obtained from fifty healthy patients and fifty patients with diagnosed with osteoporosis. The samples of tissue (50×50 pixels) were subjected to a texture analysis to obtain descriptors of features based on a histogram of grey levels, gradient, run length matrix, co-occurrence matrix, autoregressive model and wavelet transform. The obtained results were set in the importance ranking (from the most important to the least important), and the first fifty features were used for further experiments. These data were normalized and standardized and then classified using five different methods: naive Bayes classifier, support vector machine, multilayer perceptrons, random forest and classification via regression. The best results were obtained for standardized data and classified by using multilayer perceptrons. This algorithm allowed for obtaining high accuracy of classification at the level of 94.25%.

Keywords: texture analysis, standardization, normalization, classification

PORÓWNANIE WPŁYWU STANDARYZACJI I NORMALIZACJI DANYCH NA SKUTECZNOŚĆ KLASYFIKACJI TEKSTURY TKANKI GĄBCZASTEJ KRĘGOSŁUPA

Streszczenie. Celem niniejszego artykułu było porównanie wpływu metod wstępnego przetwarzania danych - normalizacji i standaryzacji - na wyniki klasyfikacji obrazów tkanki gąbczastej. Do analiz wykorzystano czterysta obrazów tomografii komputerowej kręgosłupa (kręgi L1). Obrazy pochodziły od pięćdziesięciu zdrowych pacjentów oraz pięćdziesięciu pacjentów ze zdiagnozowaną osteoporozą. Uzyskane próbki tkanki (50×50 pikseli) poddano analizie tekstury w wyniku czego otrzymano deskryptory cech oparte na histogramie poziomów szarości, macierzy gradientu, macierzy RL, macierzy zdarzeń, modelu autoregresji i transformacji falkowej. Otrzymane wyniki ustawiono w rankingu ważności (od najistotniejszej do najmniej ważnej), a pięćdziesiąt pierwszych cech wykorzystano do dalszych eksperymentów. Dane zostały poddane normalizacji oraz standaryzacji, a następnie klasyfikowane przy użyciu pięciu różnych metod: naiwny klasyfikator Bayesa, maszyna wektorów wspierających, wielowarstwowe perceptrony, las losowy oraz klasyfikacji poprzez regresję. Najlepsze wyniki uzyskano dla danych na których przeprowadzono standaryzację i poddano klasyfikacji za pomocą wielowarstwowych perceptronów. Taki algorytm postępowania pozwolił na uzyskanie wysokiej skuteczności klasyfikacji na poziomie 94,25%.

Słowa kluczowe: analiza tekstury, standaryzacja, normalizacja, klasyfikacja

Introduction

Continuous technical development entails the development of medicine, which increases the effectiveness of diagnosing many diseases. Currently, medical imaging techniques, including computed tomography, play one of the main roles. The use of modern computed tomographs (CT) allows to obtain a monochrome image of the patient's body section in very good quality [1]. Depending on the needs, one can adjust the appropriate exposure parameters. Modern CT scanners are systems with intelligent X-ray dosing. During scanning, the system changes the lamp current parameters and significantly reduces them after passing through the areas requiring higher values (such as pelvis, shoulder girdle) [17]. The image from the CT scanner consists of the so-called voxels, which are the three-dimensional equivalent of pixels in a two-dimensional image. In the images of individual layers of the examined organ, each pixel has its value determined in Hounsfield units, which correspond to the x-ray absorption coefficients [23].

The key issue in computer image processing is clearly identifying the areas of interest (ROI) [23]. The right choice of such an area increases the chances of obtaining diagnostically effective results. A valuable source of information about the condition of the tissue being examined is the texture of the image [2,3]. This property may include image granularity, pattern orientation, homogeneity, local contrast, or average brightness level of a given image area. On this basis, it is possible to distinguish two images from each other, as well as designate areas in a given image that meet certain conditions. The image texture can be symbolically described by providing the values for the finite feature vector [21]. In order to characterize the texture mathematically, a number of parameters calculated based on the properties of the digital image were introduced. In the literature, the following types of parameters can be found to describe the texture: – statistical [8], – structural [8], – using signal processing techniques [12], – morphological ([11, 15, 17]). The high quality of CT images positively affects the possibilities of interpreting the

texture of the areas of interest to us. This allows classification and segmentation, among others, in liver images [5, 13], detection of lung diseases [24] and evaluation of the effectiveness of chemotherapy in rectal cancer, classification of brain tumors and gastrointestinal cancers [22].

The image texture analysis methods combined with appropriate pre-processing and classification algorithms have found wide application in the diagnosis of internal organ diseases imaged by various methods. Examples of such applications include the diagnosis of benign and malignant microcalcifications on mammographic images of the breast (X rays) [10], classification of atherosclerotic plaques in coronary arteries (endovascular ultrasound) [22], identification of malignant brain tumor types (magnetic resonance imaging) [25], detection of focal lesions in the liver (computed tomography) [6] and identification of Hashimoto's disease (intravascular ultrasound) [18, 19].

Due to the high efficiency of the use of texture analysis in the tissue diagnostic process, an attempt was made to use this method in the detection of osteoporosis [7, 16, 20]. Osteoporosis is a skeletal disease which leads to bone fractures that can occur even after a minor injury. Most often they relate to the spine, but they can also occur in other locations. Excessive bone susceptibility to osteoporosis damage results from a decrease in bone mineral density and disturbances in its structure and quality. Osteoporosis is often asymptomatic. Only the fractures of the vertebral bodies often cause chronic back pain syndrome that prevents normal functioning [4]. Therefore, it is important to regularly monitor the condition of bone tissue. A standard procedure in the diagnosis of osteoporosis is densitometry, which is used to assess the bone mineral density. The test result is expressed by means of indicators comparing the bone density of the examined person with the bone density of young healthy persons (T-score) and peers (Z-score) [14]. However, this does not allow for accurately determining the area of the tissue in which the defects occur, which is possible in the case of analyzing the texture of specific areas.

The following article presents the use of spongy tissue texture analysis on the diagnosis of osteoporosis and the impact of data pre-processing – normalization and standardization – on the results of tissue classification.

1. Material

The CT scans of spine from a hundred patients were used to conduct the experiment. Each patient was examined on a GE 32-row tomograph in the standard L-S spine examination protocol. Fifty of them belonged to the control group, without diagnosis of osteoporosis and osteopenia. The same number of patients was also found in the group diagnosed as having osteoporosis. Four samples were obtained from each patient, and therefore four hundred spongy tissue images were used in the study.

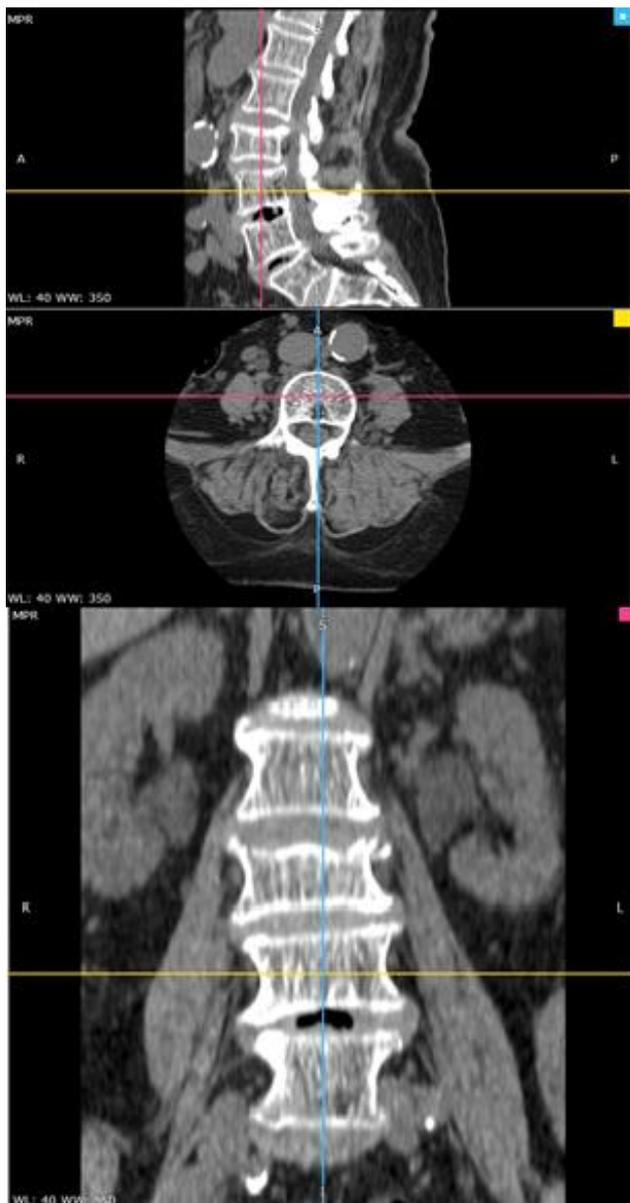


Fig. 1. Image of the spine in three planes

The samples representing the spongy tissue of the spine were selected from the image of the cross-section of the L1 vertebra in its central part. The sample size was adjusted to obtain the maximum possible tissue area. As a result of using this approach, 50×50 pixel samples were obtained.



Fig. 2. Selection of tissue sample area.

2. Method

The tissue samples obtained from the images were subjected to texture analysis. As a result, 290 features described by specific numerical values were obtained. Due to the large divergence of numerical intervals and the need to compare them with each other, the pre-processing operations, i.e. normalization and data standardization, were performed. On the basis of the obtained results, five types of classifiers were built and their effectiveness evaluated by using five parameters commonly used in descriptions of medical experiments.

2.1. Texture analysis

Image analysis was carried out with the MaZda program (version 4.6) [26]. This program allows to analyse the grey cardboard images and determine the numerical values of image features. The set of features has been obtained on the basis of histogram (9 features: histogram's mean, histogram's variance, histogram's skewness, histogram's kurtosis, percentiles 1%, 10%, 50%, 90% and 99%), gradient (5 features: absolute gradient mean, absolute gradient variance, absolute gradient skewness, absolute gradient kurtosis, percentage of pixels with nonzero gradient), run length matrix (5 features × 4 various directions: run length nonuniformity, grey level nonuniformity, long run emphasis, short run emphasis, fraction of image in runs), co-occurrence matrix (11 features × 4 various directions × 5 between-pixels distances: angular second moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy), autoregressive model (5 features: parameters $\theta_1, \theta_2, \theta_3, \theta_4$, standard deviation) and Haar wavelet (24 features: wavelet energy – the features are computed at 6 scales within 4 frequency bands LL, LH, HL, and HH) [26].

2.2. Distribution of significance of features

In the research, 290 features were obtained for each sample. Among them, the features with constant values for each sample were eliminated and 267 features remained after reduction. They are ranked in the ranking of the importance of features from the most to the least important. Fifty features occupying subsequent, initial places in the ranking were used for further experiments.

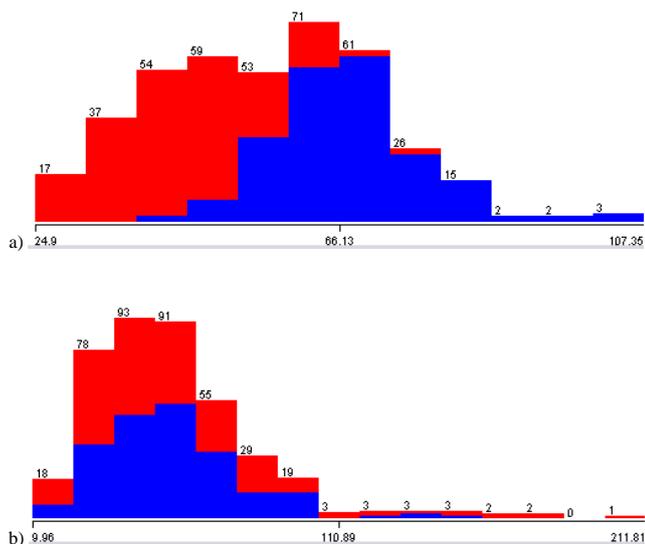


Fig. 3. Visualization of the distribution of values of features a – first in the ranking, b – last in the ranking

2.3. Normalization

Data normalization is the scaling of original data (e.g. input data) to a small specific range. This method performs a linear transformation of the original data usually to the interval [0,1] according to the formula:

$$x' = \frac{x - \min}{\max - \min} \cdot (\text{new}_{\max} - \text{new}_{\min}) + \text{new}_{\min}$$

where [min, max] is the range in which the input data falls, while [new_min, new_max] is the new range of data [9].

As a result of the transformation, the range of the first feature in the ranking (209) changed from 29.9 to 107, 35 to a range from 0 to 1.

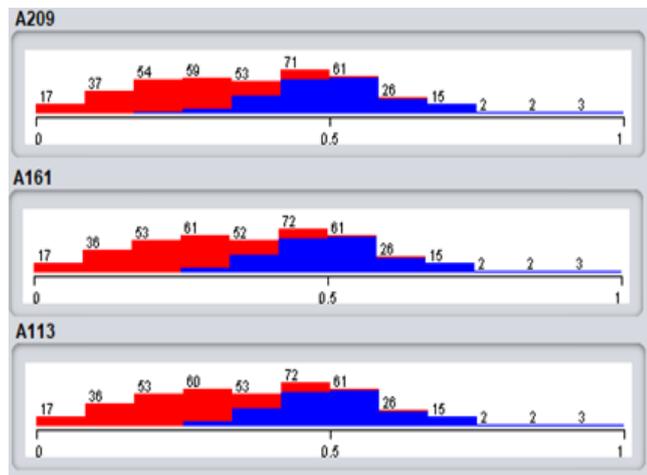


Fig. 4. Visualization of the distribution of feature values after normalization

2.4. Standardization

Standardization is the central preprocessing step in data mining, to standardize the values of features or attributes from different dynamic range into a specific range [9]

Standardization is a type of normalization of a random variable, as a result of which the variable obtains the average expected value 0 and standard deviation 1 [9]. This operation is performed according to the Z test formula:

$$Z = \frac{x - \mu}{\sigma}$$

where:

x – observed variable value,

μ – expected value, average,

σ – standard deviation.

One important restriction of the Z-score standardization is that it must be applied in global standardization [9].

As a result of standardization, the range of values for the first feature in the ranking (209) changed from 29.9 to 107, 35 to a range from -2.08 to 3.39.

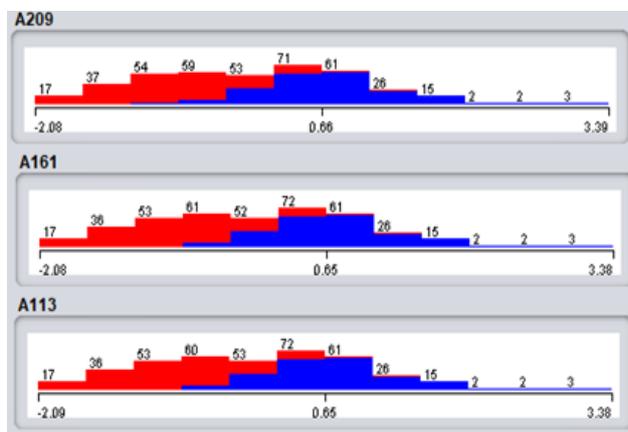


Fig. 4. Visualization of the distribution of feature values after standardization

2.5. Classification

Five types of classifiers were built based on the data obtained in the normalization and standardization process:

- Naive Bayes classifier (NBC),
- Support Vector Machine (SVM),
- Multilayer Perceptron (MP),
- Random Forest (RF),
- Classification via regression.

In order to assess the accuracy of classifiers, the following were used: general classification accuracy (ACC), true positive rate (TPR), true negative rate (TNR), positive predictive value (PPV) and negative predictive value (NPV).

3. Results

The obtained classification results (presented in the tables below) allow to clearly determine the effectiveness of the classifiers depending on the type of pre-processing that were used.

Support Vector Machine turned out to be the most effective classifier for the data after normalization. It obtained the highest values of all indicators used to assess the effectiveness of classifiers. The ACC value was 88.5%. Other TPR and PPV parameters reached 86.5%, TNR and NPV 90.5%. Similar results were achieved for the Naive Bayes classifier. They only differed in the ACC value, which in this case amounted to 88.25% and was 0.25% lower than in the case of Support Vector Machine. Random Forest turned out to be the least effective classifier. The ACC value here was 86.25 and was 2.25% lower than the highest. Moreover, for the remaining parameters, Random Forest showed the lowest values. TPR and PPV obtained 84.5%, TNR and NPV 88.0%.

Table 1. Classification results after data normalization

Classifier	ACC	TPR	TNR	PPV	NPV
NaiveBayes	88.25	86.5	90.0	86.5	90.0
SVM	88.50	86.5	90.5	86.5	90.5
Multilayer Perceptron	88.00	85.5	90.5	85.5	90.5
Random Forest	86.25	84.5	88.0	84.5	88.0
Classification Via Regression	87.75	85.0	90.5	85.0	90.5

Multilayer Perceptron turned out to be the best classification method for standardized data. The highest values for ACC = 94.25% as well as TPR and PPV equal to 95.5% were obtained for this classifier. For TNR and NPV, the values were 1% lower than for Support Vector Machine and Classification via Regression, which were equal to 94%. Naive Bayes turned out to be the least effective classifier, achieving an accuracy of 87.25%. Other parameters were also the lowest among those obtained and amounted to 86.0% for TPR and PPV and 88.5% for TNR and NPV.

Table 2. Classification results after data standardization

Classifier	ACC	TPR	TNR	PPV	NPV
NaiveBayes	87.25	86.0	88.5	86.0	88.5
SVM	93.75	93.5	94.0	93.5	94.0
Multilayer Perceptron	94.25	95.5	93.0	95.5	93.0
Random Forest	91.00	91.5	90.5	91.5	90.5
Classification Via Regression	91.50	89.0	94.0	89.0	94.0

4. Conclusions

The results presented above clearly indicate the higher efficiency of classification of standardized data. The accuracy of results is 5.75% better, which is a significant difference in diagnostic tests. Moreover, other parameters are much higher in the case of the data after standardization – TPR and PPV by 9% and TNR and NPV by 3.5% from the highest values.

The above-mentioned results clearly indicate the best algorithm for data analysis which is their standardization and then the use of the Multilayer Perceptron classifier. The classification results obtained in this way allow obtaining a relatively high accuracy of 94.25%. In terms of medical diagnostics, the result is the basis for using this method in creating automatic image analysis systems.

References

- [1] Budzik G., Dziubek T., Turek P.: Podstawowe czynniki wpływające na jakość obrazów tomograficznych. *Problemy Nauk Stosowanych* 2015, 77–84.
- [2] Chen Y, Dougherty E.R.: Gray-scale morphological granulometric texture classification. *Optical Engineering* 33 (8)/1994, 2713–2722.
- [3] Cichy P.: Analiza tekstury obrazów cyfrowych – zastosowanie do wybranej klasy obrazów biomedycznych. *Rozprawa doktorska, Politechnika Łódzka, Wydział Elektrotechniki i Elektroniki, Instytut Elektroniki, Łódź* 2001.
- [4] Downey P.A., Siegel M.I.: Bone Biology and the Clinical Implications for Osteoporosis. *Phys Ther* 86/2006, 77–91.
- [5] Duda D., Krtowski M., Bęzy-Wendling J.: Klasyfikacja tekstur w rozpoznawaniu nowotworów wątroby na podstawie serii obrazów tomograficznych. *Obrazowanie Medyczne, tom 1, 2005.*
- [6] Duda D., Krętowski M., Bęzy-Wendling J.: Ekstrakcja cech teksturalnych w klasyfikacji obrazów tomograficznych wątroby. *Zeszyty Naukowe Politechniki Białostockiej, Informatyka, 2007.*
- [7] Dzierżak R., Omiołek Z., Tkacz E., Kępa A.: The Influence of the Normalisation of Spinal CT Images on the Significance of Textural Features in the Identification of Defects in the Spongy Tissue Structure. *IBE 2018 Innovations in Biomedical Engineering*, 2019, 55–66.
- [8] Giannakopoulos X., Karhunen J., Oja E.: An Experimental Comparison of Neural ICA Algorithms. *Proc. Int. Conf. on Artificial Neural Networks ICANN'98, 1998, 651–656.*
- [9] Ismail Bin M., Dauda U.: Standardization and Its Effects on K-Means Clustering Algorithm. *Research Journal of Applied Sciences, Engineering and Technology* 6(17)/ 2013, 3299–3303.
- [10] Lazarek J.: Metody analizy obrazu – analiza obrazu mammograficznego na podstawie cech wyznaczonych z tekstury. *Informatyka, Automatyka Pomiary w Gospodarce i Ochronie Środowiska* 4/2013, 10–13.
- [11] Lee T.W., Lewicki M.S.: Unsupervised Image Classification, Segmentation and Enhancement Using ICA Mixture Models. *IEEE Transactions on Image Processing* 11(3)/2002, 270–279.
- [12] Lygeros J.: A Formal Approach to Fuzzy Modelling. *Proceedings of ACC, 1995, 3740–3744.*
- [13] Mala K., Sadasivam V.: Automatic Segmentation and Classification of Diffused Liver Diseases using Wavelet Based Texture Analysis and Neural Network. *Annual IEEE INDICON Conference, 2005, 216–219.*
- [14] Marcus R., Feldman D., Dempster D., Luckey M., Cauley J.: *Osteoporosis*, 4th ed. Elsevier Academic Press, 2013.
- [15] Matheron G.: *Random sets and integral geometry*. Wiley, New York 1975.
- [16] Nasser Y., Hassouni M., Brahim A., Toumi H., Lespessailles E., Jennane R.: Diagnosis of osteoporosis disease from bone X-ray images with stacked sparse autoencoder and SVM classifier. *Proceedings of the 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), 2017, 1–5.*
- [17] Nieniewski M., Serneels R.: Extraction of the Shape of Small Defects on the Surface of Ferrite Cores. *Machine Graphics and Vision* 9 (1/2)/2000, 453–462.
- [18] Omiołek Z.: Improvement of the classification quality in detection of Hashimoto's disease with a combined classifier approach. *Journal of Engineering in Medicine* 231(8)/ 2017, 774–782.
- [19] Omiołek Z., Wójcik W.: The use of Hellwig's method for dimension reduction in feature space of thyroid ultrasound images. *Informatyka, Automatyka, Pomiary* 3/2014, 14–17 [DOI: 10.5604/20830157.1121333].
- [20] Reshmalakshmi C., Sasikumar M.: Trabecular bone quality metric from X-ray images for osteoporosis detection. *Proceedings of the 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), India, 2017, 1694–1697.*
- [21] Snitkowska E.: *Analiza tekstur w obrazach cyfrowych i jej zastosowanie do obrazów angiograficznych, Rozprawa doktorska, Politechnika Warszawska, 2004.*
- [22] Strzelecki M., Materka A.: *Tekstura obrazów biomedycznych. Metody analizy komputerowej. Wydawnictwo PWN, Warszawa* 2017.
- [23] Tadeusiewicz R., Śmietaniński J.: Pozyskiwanie obrazów medycznych oraz ich przetwarzanie, analiza, automatyczne rozpoznawanie i diagnostyczna interpretacja. *Wydawnictwo Studentckiego Towarzystwa Naukowego, Kraków* 2011.
- [24] Titus A., Nehemiah H., Kannan A.: Classification of interstitial lung disease using particle swarm optimized support vector machines. *International Journal of Soft Computing* 10 (1)/2015, 25–36.
- [25] Usman, K., Rajpoot, K.: Brain tumor classification from multi-modality MRI using wavelets and machine learning. *Pattern Analysis and Applications* 20(3)/2017, 871–881.
- [26] www.elelel.p.lodz.pl/programy/cost/progr_mazda.html [06.05.2018].

M.Sc. Róża Dzierżak
e-mail: r.dzierzak@pollub.pl

Ph.D. student at the Electrical Engineering and Computer Science Faculty. Assistant in the Institute of Electronics and Information Technology of Lublin University of Technology. Recent graduate Biomedical Engineering and Computer Science. Her research interests include medical image processing.



ORCID ID: 0000-0001-5640-0204

otrzymano/received: 15.05.2019

przyjęto do druku/accepted: 15.06.2019