http://doi.org/10.35784/iapgos.2455

OVERVIEW OF FEATURE SELECTION METHODS USED IN MALIGNANT MELANOMA DIAGNOSTICS

Magdalena Michalska
Lublin University of Technology, Department of Electronics and Information Technology, Lublin, Poland

Abstract. *Currently, a large number of trait selection methods are used. They are becoming more and more of interest among researchers. Some of the methods are of course used more frequently. The article describes the basics of selection-based algorithms. FS methods fall into three categories: filter wrappers, embedded methods. Particular attention was paid to finding examples of applications of the described methods in the diagnosis of skin melanoma.*

Keywords: feature selection methods, filter methods, wrappers methods, embedded methods

PRZEGLĄD METOD SELEKCJI CECH UŻYWANYCH W DIAGNOSTYCE CZERNIAKA

Streszczenie. *Obecnie stosuje się wiele metod selekcji cech. Cieszą się coraz większym zainteresowaniem badaczy. Oczywiście niektóre metody są stosowane częściej. W artykule zostały opisane podstawy działania algorytmów opartych na selekcji. Metody selekcji cech należące dzielą się na trzy kategorie: metody filtrowe, metody opakowujące, metody wbudowane. Zwrócono szczególnie uwagę na znalezienie przykładów zastosowań opisanych metod w diagnostyce czerniaka skóry.*

Słowa kluczowe: metody selekcji cech, metody filtrowania, metody opakowujące, wbudowane metody

Introduction

Early detection and classification of melanoma is extremely important for treatment and patient outcome. In order to classify selected features of the image, they must be properly selected. Important in the diagnostic processes is the selection of an appropriate set of data (dermatoscopic images), a classification method of skin lesions, the classification process and selection of features. This last stage is also not the easiest one. Figure 1 presents a diagram of the diagnostic process based, of course, on an appropriately selected method of selecting features.

There are many methods of selecting features. The feature selection methods are broken down into three basic categories: filters, wrappers and embedded methods [4]. In recent years, researchers have developed many methods to select features through IT tools [6, 11, 32]. Still new feature selection methods are being proposed.

The rapidly increasing number of features is a very serious problem to be solved. This increases the computational complexity of the algorithm, extends the learning process and increases multi-level classification method.

The best result of the classifier is given by a properly selected feature selection algorithm. Feature selection, reduction of the feature space dimensionality, reduces the number of free parameters in the classifier necessary for estimation. When collecting data again, you can focus only on the features important for the classification algorithm [31]. Filters mainly use the general characteristics of data sets. Wrappers and embedded methods build a subset of functions based on selected algorithms.

The most important algorithms for selecting the features of medical images include methods [21]: SBS (Sequential Backward Selection), SFS (Sequential Forward Selection) and its modifications (SFFS (Sequential Forward Floating Search)). The SFFS algorithm requires providing the algorithm's stop condition, the number of necessary operations does not have to be so large due to the removal of features previously selected from the subset.

Other methods are: method Plus-L-Minus-R, NNFP (Nearest Neighbor with Feature Projection), methods based on genetic algorithms, OSA (Oscillating Search Algorithm), methods based on the use of fractal dimension, methods based on information thery.

Figure 2 presents a summary of used groups of feature selection methods based on four categories such as classification, segmentation, annotation and retreval. Scientists use filter methods the most, followed by embeded methods. Filter methods are at the first place of use.

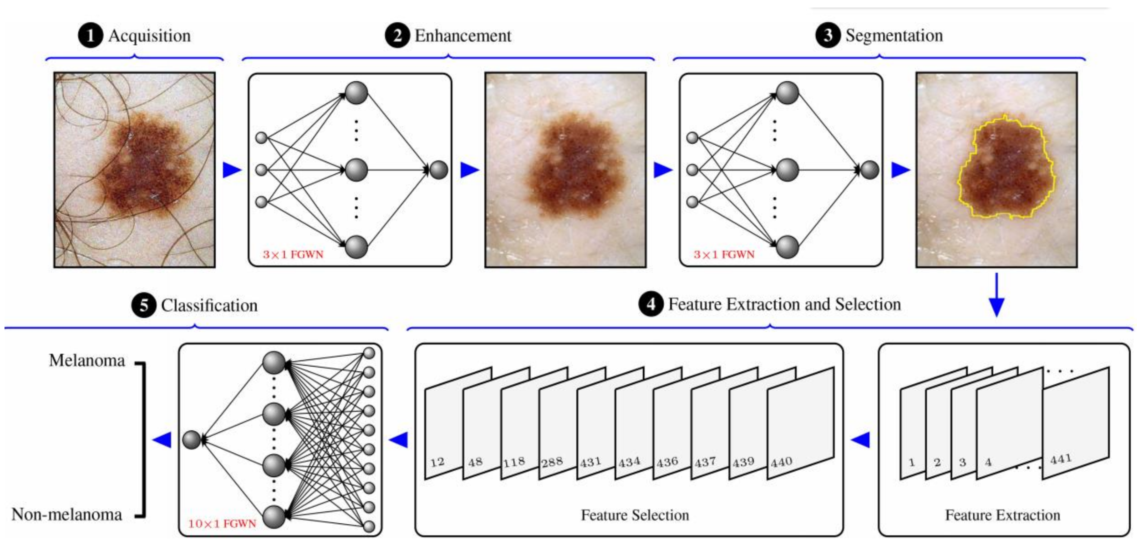

Fig. 1. Scheme of diagnosis method of skin lesions from dermoscopic images [29]

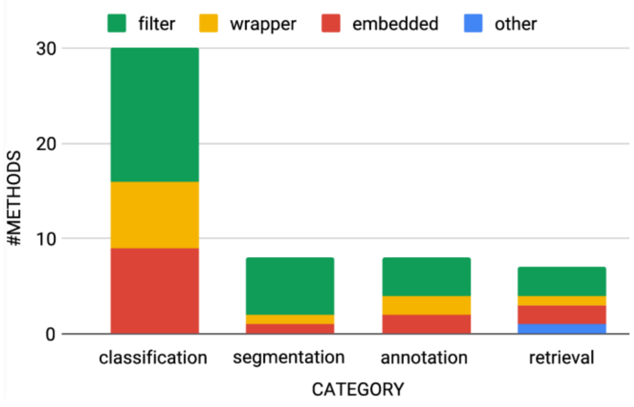

Fig. 2. Application of selected types of feature selection methods in numbers [4]

1. Filters methods

As a pre-processing process, the most frequently used by researchers are filter methods. The methods use a statistical measure and the functions are selected for retention or removal from the data. The methods are usually one-dimensional and take into account features independently or in relation to the dependent variable. The filters methods include: correlation-based (CFS) [13], consistency-based filter [7] and information gain [12, 13]. CFS then combines this evaluation formula with an appropriate measure of correlation and a heuristic filter search strategy, selecting subsets of attributes not correlated between them. It can showe a correlation with the all class.

Another very common filters method is ReliefF [18, 19]. Filter models according to developed sources are more computationally efficient [39]. The relief algorithm is effective in determining a given feature [20]. Figure 3 gives a detailed description of the reliefF algorithm. ReliefF randomly choses an instance R_i from class. It can find K for the nearest neighbors from the same class (nearest hits H) and from the different classes (nearest misses M), i.e.:

$$W_i = W_i - \frac{\sum_{k-1}^{K} D_H(k)}{n \cdot k} + \sum_{c=1}^{C-1} p_c \cdot \frac{\sum_{k=1}^{K} D_M(k)}{n \cdot k} \tag{1}$$

where W_i – quality measure for feature according to R_i values, hits H and misses M; $D_H(k)$ and $D_M(k)$ – distance between the selected instance and its nearest neighbors in H (or M); c_p – class probability c; n – repeats n times.

```
Input: Feature data matrix: D, repeat times: n, the number of the neighbors: K
Output: Vector W for the feature attributes ranking
Begin
  for j=1 to n do
    Randomly select an instance R_j;
    Find K nearest hits H and nearest misses M;
    for i=1 to all features do
      Updating estimation W_i by Equation(1);
    end
  end
End
```

Fig. 3. ReliefF algorithm [39]

In the study [23], several different feature selection algorithms were used to create subsets for classifiers. The algorithms are based on various bases, e.g. Pearson's correlation coefficient based on feature selection gain factor [34]. Relief-F, principal component analysis (PCA) and feature selection based on correlation (CFS) are also used in many works. These algorithms are commonly used, because they have a number of advantages. Computing performance is one of them. Additionally, they have become less time-consuming and do not result in excessive and independent evaluation criteria [28].

In most cases, the selected features are determined by the correlation results of statistical tests [16]. Common used defining correlation coefficients are [1]: pearson's correlation, LDA (Linear Discriminant Analysis), ANOVA stands, Chi-Square [29]. The data based on the dermoscopic images served as a test kit to evaluate the effectiveness of the classification.

Several feature selection algorithms were used in [29]: ReliefF algorithm, Fisher score [5], chi-square. Table 1 below shows a comparison of the results from [29].

Table 1. Accuracy for different feature selector [29]

Parameter	Feature selector				
	ReliefF	FCBF	FS	mRMR	Chi-squere
Mean of accuracy [%]	87.1	85.8	85.8	87	85.8

In filtering methods, class separation, error probability, and inter-class distance are used. Very common is correlation-based feature selection, entropy, consistency-based feature selection and filter methods do not remove multicollinearity, it should be fixed before training models [17].

2. Wrappers methods

The scheme of functioning has been re-colored in Figure 4 wrappers methods. It belongs to them set of all features. Next is selected the best subset to generate a subset and learnig algorithm is started. After all those the Performance is been done.

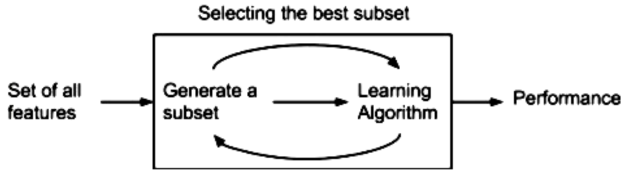

Fig. 4. Wrapper methods model [14]

In wrapper methods important is to use a subset of features and train a model. Based on the inferences from the previous model, features subset are added or removed [5, 9]. Very common for wrapper methods are forward feature selection, backward feature elimination, recursive feature elimination. These are usually computationally very expensive.

The study [24] adopted them. Greedy stepper search methods contain subsets in forward or backward direction. The selection stop when any feature is added or removed. This function degrades the result of the subset up to this point [37]. The best first method searches for subsets of functions. An empty feature set starts the selection forward, features compatible with the evaluation method are added to the data set. On the other hand, all features start backselection, and mismatched features are removed from the set [15].

The Support Vector Machine (SVM) recursive support vector machine-recursive feature elimination (SVM-RFE) method, which is a very typical wrapper selector. The method was first developed for the gene selection process using the SVM classifier [20, 25, 26]. The system of mobile applications [8] helps to classify skin nevi on dermatoscopic images as melanoma, benign nevi.

A type of machine learning technique is Genetic Programming (GP). It allows the use of the evolutionary algorithm for simple and understandable classifiers [22, 36]. GP is also used to diagnose tumor expression, c was used in the selection of features and classifiers [2, 37].

In order to use the selected skin lesion classification algorithm, first of all, reduce the size of the dermatological image on which it is located. With a large amount of data, it is useful to reduce features and design functions to reduce their size. This operation allows for greater efficiency of the used classifier. In the work [33], an innovative, two-step GP algorithm was developed to select the features and structure of features for the classification of the skin cancer picture. The local binary pattern helps to show gray and color characteristics from dermoscopic images.

Unlike wrapper and embedded methods, filter methods require more computational effort. In addition, they are less accurate in their selection. Wrappers are over-matched when the number of samples is smaller than the number of elements.

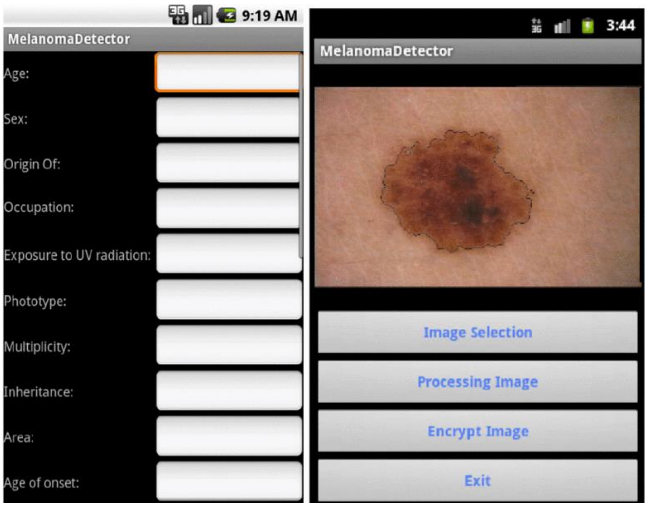

Fig. 5. Mobile application from melanoma detection [8]

3. Embedded methods

Embedded methods use internal representations of selected classifiers, which evaluate the usefulness of features in the learning process. In order to build a model, methods of selecting features are also used. [36]. They also usually give better results than filter methods. They are designed according to the selected classification algorithm. The methods are faster because the selection process does not require calling the classifier multiple times for each feature subset.

For detection melanoma the best is to find combination of different criteria. The lesion area was analyzed in terms of lesion area division parameters – Figure 6 [38].

Benign lesions differ from malignant ones in terms of selected characteristic attributes. The analyzed homogeneity and selected color characteristics (Figure 3a) usually have higher values in the case of benign skin lesions. in order to classify with the highest efficiency, a combination of selected attributes should be used [9].

One of the rapidly developing methods is LASSO (Least Absolute Shrinkage and Selection Operator) [3] and method RLS (Relaxed Linear Separability) [10]. The methods are especially used with bigger number of samples in the training set. The following are used to reduce the dimensions of the function: Sammon mapping, principal components analysis (PCA), decomposition of singular values. Very often a low variance filter and a high correlation filter are used. Useful for feature selection in addition to classifiers are Random forests [30, 35].

The diagnosis of melanoma is possible thanks to the visualization of the analysis of structured data [30]. The created data set made to measure significantly exceeds the limits of today's multi-dimensional and multi-dimensional visualization techniques. Visualization based on (PCA) [16] reduces dimensionality to a manageable range and provides better visualization. PCA may introduce errors, but the tolerance of error can be assessed and controlled [27].

In [1] using SVM based on the selected features from PCA, achieved accuracy of around 92% with 11 features. Figure 7 shows the selection results for 5 features using the PCA method. the developed methods allow to distinguish malignant from benign changes, becoming a fairly powerful diagnostic tool.

Sequential forward search algorithms SFS (ang. sequential forward selection) and sequential search backwards SBS (ang. Sequential Backward Selection) are examples of simple boxing methods. In the case of the first method, the algorithm adding a new feature in each subsequent step [2]. With both of these methods, the forward or reverse scanning step is followed by a reverse scanning step. This allows for the removal of a feature in the SFS algorithm that becomes redundant after adding others, and in the case of the SBS algorithm. It is possible to consider a given feature again, although it was removed in an earlier step of the algorithm [32]. Also very common is Backward Feature Elimination and inverse process Forward Feature Construction [35].

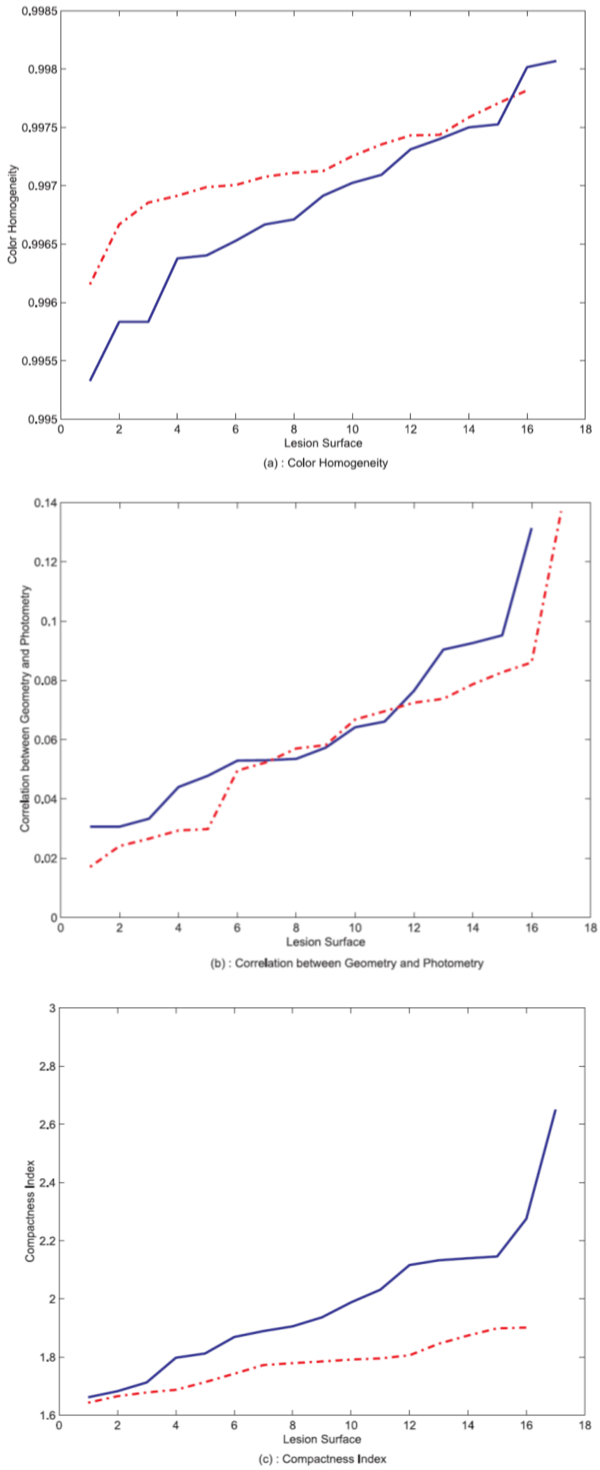

Fig. 6. Correlations and compactness in relation to benign lesions "red dotted line" and malignant lesions "blue solid line" [38]

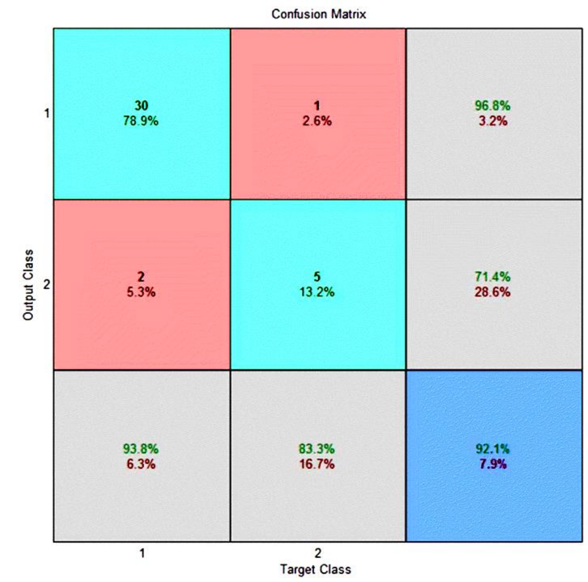

Fig. 7. The matrix of using SUV for features selction using PCA [1]

4. Conclusions

Function selection algorithms use many features of skin lesions. There are often a number of characteristics that need to be kept. The feature subset evaluator measures a features quantity and returns the search value. The choice of features and their design alone. Performance improvements can be achieved by selecting an appropriately selected feature, or by using a more extensive multi-level feature. Each of the methods mentioned have many advantages and disadvantages. Their features complement each other. Filter methods are less computationally expensive, embedded methods allow for a more precise selection. The use of additional neuron networks, SUVs, methods of feature elemination, decision trees gives the opportunity to obtain the best possible result, allowing for an accurate diagnosis.

References

[1] Alquran H., Qasmieh I. A., Alqudah A. M., Alhammouri S., Alawneh E., Abughazaleh A., Hasayen F.: The melanoma skin cancer detection and classification using support vector machine. IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Aqaba, Jordan, 2017, 1–5 [http://doi.org/10.1109/AEECT.2017.8257738].
[2] Al-Sahaf H., Al-Sahaf A., Xue B., Johnston M., Zhang M.: Automatically Evolving Rotation-Invariant Texture Image Descriptors by Genetic Programming. IEEE Transactions on Evolutionary Computation 21(1)/2017, 83–101.
[3] Andersen S. W., Runger G. C.: Automated feature extraction from profiles with application to a batch fermentation process. Journal of the Royal Statistical Society: Series C (Applied Statistics) 61(2)/2012, 327–344.
[4] Bolón-Canedo V., Remeseiro B.: Feature selection in image analysis: a survey. Artif Intell Rev 53/2020, 2905–2931.
[5] Celebi M. E., Aslandogan Y. A., Stoecker W. V., Iyatomi H., Oka H., Chen X.: Unsupervised border detection in dermoscopy images. Skin Res Technol. 13/2007, 1–9.
[6] Chmielnicki W.: Efektywne metody selekcji cech i rozwiązywania problemu wieloklasowego w nadzorowanej klasyfikacji danych. Rozprawa doktorska. Instytut Podstawowych Problemów Techniki PAN, Kraków 2012.
[7] Dash M., Liu H.: Consistency-based search in feature selection. Artificial Intelligence 151(1–2)/2003, 155–176 [http://doi.org/10.1016/S0004-3702(03)00079-1].
[8] Doukas C., Stagkopoulos P., Kiranoudis C. T., Maglogiannis I.: Automated skin lesion assessment using mobile technologies and cloud platforms. Engineering in Medicine and Biology Society (EMBC) – Annual International Conference of the IEEE, 2012.
[9] Ercal F., Chawla A., Stoecker W.V., Lee H., Moss R. H.: Neural Network diagnosis of malignant melanoma from color images. IEEE Transactions on Biomedical Engineering 41(9)/1994, 837–845.
[10] Gościk, J., Łukaszuk, T.: Application of the recursive feature elimination and the relaxed linear separability feature selection algorithms to gene expression data analysis. Advances in Computer Science Research 10/2013, 39–52.
[11] Guyon I., Elisseeff A.: An introduction to variable and feature selection. Journal of Machine Learning Research 3/2003, 1157–1182.
[12] Hall M., Smith Lloyd A.: Practical feature subset selection for machine learning. Springer 1998.
[13] Hall M.: Correlation-based feature selection for machine learning. Department of Computer Science 19/2000.
[14] https://moredvikas.wordpress.com/2018/10/09/machine-learning-introduction-to-feature-selection-variable-selection-or-attribute-selection-or-dimensionality-reduction/
[15] Huang J., Ling C. X.: Using AUC and accuracy in evaluating learning algorithms. IEEE Trans Knowledge Data Eng. 17(3)/2005, 299–310.
[16] Keerthi Vasan K., Surendiran B.,: Dimensionality reduction using Principal Component Analysis for network intrusion detection. Perspectives in Science 8/2016, 510–512.
[17] Khan M. A., Tallha A., Muhammad S., Aamir S., Khursheed A., Musaed A., Syed I. H., Abdualziz A.: An implementation of normal distribution based segmentation and entropy-controlled features selection for skin lesion detection and classification. BMC Cancer 18(1)/2018, 638.
[18] Kira K., Rendell L. A.: A practical approach to feature selection. Machine Learning Proceedings 1992, 249–256.
[19] Kononenko I.: Estimating attributes: Analysis and extensions of Relief. L. De Raedt, & F. Bergadano (Eds.): Machine Learning: ECML-94 1994, 171–182.
[20] Kuo B. C., Ho H. H., Li C. H., Hung C. C., Taur J. S.: A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification. IEEE J Sel Top Appl Earth Obs Remote Sens 7(1)/2014, 317–326.
[21] Liu H., Yu L.: Toward integrating feature selection algorithms for classification and clustering. IEEE Transactions on knowledge and data engineering 17(4)/2005, 491–502.
[22] Neshatian K., Zhang M., Andreae P.: A filter approach to multiple feature construction for symbolic learning classifiers using genetic programming. IEEE Trans. Evol. Comput. 16(5)/2012, 645–661.
[23] Oliveira R. B., Pereira A. S., Tavaresa J. M. R. S.: Skin lesion computational diagnosis of dermoscopic images: Ensemble models based on input feature manipulation. JMRS Tavares – Computer methods and programs Computer Methods and Programs in Biomedicine 149/2017, 43–53.
[24] Oliveira R. B., Pereira A. S., Tavaresa J. M. R. S.: Skin lesion computational diagnosis of dermoscopic images: Ensemble models based on input feature manipulation. JMRS Tavares – Computer methods and programs Computer Methods and Programs in Biomedicine 149/2017, 43–53.
[25] Pal M., Foody G. M.,: Feature selection for classification of hyperspectral data by SVM. IEEE Trans Geosci Remote Sens 48(5)/2010, 2297–2307.
[26] Qi C., Zhou Z., Sun Y., Song H., Hu L., Wang Q.: Feature selection and multiple kernel boosting framework based on PSO with mutation mechanism for hyperspectral classification. Neurocomputing 220/2017, 181–190.
[27] Ramezani M, Karimian A, Moallem P.: Automatic Detection of Malignant Melanoma using Macroscopic Images. J Med Signals Sens. 4(4)/2014, 281–290.
[28] Robnik-Šikonja M., Kononenko I.: Theoretical and empirical analysis of ReliefF and RReliefF. Machine learning 53(1–2)/2003, 23–69.
[29] Sadri A. R., Azarianpour S., Zekri M., Celebi M. E., Sadri S.: WN-based approach to melanoma diagnosis from dermoscopy images. IET Image Process. 11(7)/2017, 475–482.
[30] Shahid M., Khan S.: Dermoscopy Images classification based on color, texture and shape features using SVM. The 3rd International Conference on Next Generation Computing (INC GC2017b) 2017, 243–245.
[31] Stapor K., Automatyczna klasyfikacja obiektów. Akademicka Oficyna Wydawnicza EXIT, Warszawa 2005.
[32] UCI Machine Learning Repository [http://archive.ics.uci.edu/ml/datasets.html].
[33] Ul Ain B., Xue B., Al-Sahaf H., Zhang M.: Genetic programming for feature selection and feature construction in skin cancer image classification. Pacific Rim International Conference on Artificial Intelligence, Springer 2018, 732–745.
[34] Witten I. H., Frank E., Hall M. A.: Data mining: Practical machine learning tools and techniques. Morgan Kaufmann 2011.
[35] Xie F., Fan H., Li Y., Jiang Z., Meng R., Bovik A.: Melanoma classification on dermoscopy images using a neural network ensemble model, IEEE Transactions on Medical Imaging 36(3)/2017, 849–858.
[36] Xue B., Zhang M., Browne W. N., Yao X.: A survey on evolutionary computation approaches to feature selection. IEEE Trans. Evol. Comput. 20(4)/2016, 606–626.
[37] Yu J., Almal A. A., Dhanasekaran S. M.,Ghosh D., Worzel W. P., Chinnaiyan A., M.: Feature selection and molecular classification of cancer using genetic programming. Neoplasia 9(4)/2007, 292–303.
[38] Zagrouba E., Barhoumi W.: An accelerated system for melanoma diagnosis based on subset feature selection. Journal of Computing and Information Technology – CIT 13(1)/2005, 69–82.
[39] Zhou X., Wang J. J.: Feature selection for image classification based on a new ranking criterion. Journal of Computer and Communications 3/2015, 74–79 [http://doi.org/10.4236/jcc.2015.33013].

M.Sc. Magdalena Michalska
e-mail: magdalena.michalska@pollub.edu.pl

Ph.D. student at the Department of Electronics and Information Technology, Lublin University of Technology. Recent graduate of Warsaw University of Technology. Her research field covers medical image processing, 3D modelling, optoelectronics, and spectrophotometry. Author of more than 10 publications.

http://orcid.org/0000-0002-0874-3285

otrzymano/received: 19.01.2021 *przyjęto do druku/accepted: 15.03.2021*