Density Weighted K-Nearest Neighbors Algorithm for Outliers in the Training Set Are So Close to the Test Element

Selahaddin Batuhan Akben¹ and Ahmet Alkan²

1. Bahce Vocational School, Osmaniye Korkut Ata University, Bahce/Osmaniye 80500, Turkey

2. Electrical and Electronics Engineering Department, Kahramanmaraş Sutcu Imam University, Kahramanmaraş 46100, Turkey

Abstract: In KNN (K-Nearest Neighbour) method, the distance-weighted algorithm is applied in order to reduce the effect of noisy data. However, as this weighting algorithm will be insufficient when the outliers in the training set are so close to the test element, a new weighting algorithm is required for KNN. In this study, instead of distance-weighting, a new density-weighted KNN algorithm is proposed for reducing the effect of noisy data. In the first stage of the proposed method, the coefficient of density of each element in the training set was obtained by Parzen window method. And then, the membership of each test element was determined according to the total of density coefficients (weights) of neighbours belonging to the same class. As for the last stage, the performance results of the frequently used KNN methods and the proposed method (Density-weighted KNN, Classical KNN and Distance-weighted KNN) were compared. The obtained results have shown that the proposed method is more successful by almost 1% than classical KNN method, by 9% than distance weighted KNN method. Moreover, it is more successful than other KNN methods when the test element is so close to the training set elements.

Key words: Density-Based KNN algorithm, KNN, DWKNN, Parzen Windows, classification.

1. Introduction

The KNN (K-Nearest Neighbour) algorithm is used widely as it is a simple and effective classification algorithm. However, KNN has got some disadvantages. One of the most important of these disadvantages is being sensitive to noisy data [1]. This occurs especially when the test element is so close to the noisy elements (or outliers) of training set [2]. In order to resolve this problem, the commonly used feature is the weighting of training data within the neighbourhood boundary according to their distances to the test element [3]. However, in this case, one or several of training data which is too close to the test element may also cause wrong classification [4]. Because although neighbours which are too close to the test element prevail other class element in terms of

distance, they may be a lot fewer in number [5, 6]. A sample data distribution for this kind of problems is shown in Fig. 1.

There are two classes in training set as seen in the example data set in Fig. 1 and the outlier of one of the classes is too close to the test element. This outlier will classify wrongly the test element in the distance-weighted classification. In order to avoid this case, the distance of elements in the training set to their class centre is used in this weighting. However, if the outliers are as near as being above of the test element, this weighting based on double distance is useless. Moreover, the method based on weighting by the double distance extends the process time. Apart from all of these, the distance unit and the weighting degree should be selected rightly. In other words, in the distance-weighted classification, problems such as distance unit selection and the need of classification based on which distances occur. Even if all selections

Corresponding author: Selahaddin Batuhan Akben, Ph.D., research fields: signal processing, biomedical signals, data mining, pattern recognition and artificial intelligence.

Fig. 1 Outliers disadvantage for DWKNN (distance-weighted KNN) algorithm.

Fig. 2 Disadvantages of the elements in local density region which is away from their set centre for KNN algorithm.

were made rightly, the success rate may be low due to extreme closeness of the outliers to the test element.

The density-weighted KNN algorithm was proposed against these problems of the DWKNN (distance-weighted KNN) method. In this density-weighted KNN method proposed in literature, the coefficient of density forming according to the training elements within the neighbourhood boundary of the test element is used. In this method, if the test element has a higher coefficient of density when included in which class within the neighbourhood boundary, it is included in that class. However, if it is

better as for success, the density coefficient should be calculated for each test element in this method. Moreover, the calculation of coefficient of the test element made separately for each class. For this reason, the process time is very long and it is difficult to use in big data set. In this method, although the outlier or the effect of noises is reduced, the problems arising from extreme closeness of the test element to any training element were not avoided completely. Because this method is very sensitive to the selection of neighbour number [7].

Ultimately there are lots of methods proposed in literature in order to avoid disadvantages relating to noisy data in the KNN method. However, nothing was proposed completely for failures arising from extreme closeness of noisy training data to the test element to be classified in any of these methods. Because this problem occurs in all algorithms based on the distance between the test elements and neighbours [8]. Therefore a new KNN method which will increase the success and will be affected less from noises is required by taking into account both the data' intraclass scatters and distances to the test element. In order to meet this requirement, the advantages of methods proposed for the distance and density in literature can be used.

In the weighting of elements with their distances to their set centres, the noisy elements will have low weight value. Thus, the significance level of noises can be reduced. However, in this case, a problem may occur as the effect of elements in local density regions which are away from the set centre will go down. This kind of sample problem is shown in Fig. 2.

As seen in Fig. 2, there is a local density region which is away from its set centre in one of the classes. If the significance of the elements in this local density region is reduced depending upon the distance to its own set centre, the classification can be wrong. Because the test element can belong to the class of the elements in this local density region.

Therefore, in the proposed method, the distances of

training elements both to their own set centre and to the elements belonging to their class around them were considered together. In the calculation of the coefficient of density, instead of making calculation for each test element, the calculation of the coefficients for training elements for once will be better. Because both the problem related to the distances of the test element to its neighbours will go away and the relationship of the neighbours with their own class elements will be considered.

In that case, if the coefficients of density of the elements are calculated separately for each class in the training set, the disadvantages of distance-weighted and previous density-weighted methods can be overcome. Because noises in data set have low density as they are more distant to set centre than other elements. They can be easily distinguished due to this low density [9]. Moreover, while calculating the coefficients of density of the noisy elements, their distances to the elements close to them in the same class around them will be considered. Thus, the noisy data will be weighted more accurately by their location. The effect of noises and outliers on neighbourhood can be easily reduced. Likewise, the effect of local density regions including noisy data around the test element will also be taken into account. Thus, a KNN method can be formed, in which the noises are evaluated optimally but the disadvantages relating to their closeness to the test element are eliminated.

In this new method, the coefficients of density of each element in the training set which are obtained according to their own class elements can be specified as weights [10]. In this sense, it is aimed to calculate the coefficients of density based on the elements of the training data in their own class. The calculated coefficients of density are used as weights in KNN algorithm. The membership of a test element can be specified by the sum total of weights (coefficients of density) of neighbour training elements belonging to the same class. Additionally, the block diagram of this

study can be seen in Fig. 3.

The method proposed in this study was applied to real data sets and the results of two commonly used KNN methods and the proposed KNN method were compared.

2. Materials

In the calculation of the density coefficients, intraclass standard deviations become important. If intraclass standard deviations are low, the determination of the density-based noise detection will become difficult. Because, the density coefficients will be nearly same. Because intraclass standard deviation gives information relating to the amount of noisy data and outliers that will be very close to test element. Such that, the datasets having high intraclass standard deviation value (class elements scattered inhomogeneous), it is expected that the success of proposed method will increase in comparison with DWKNN method.

Also, very closeness or superposition of some elements belonging to different classes (the closeness of different classes to each other) becomes important in the density-weighted KNN method. Because closeness of different classes gives information relating to the closeness of the test elements to neighbour elements belonging to different classes. Depending upon the number of elements of different classes which are very close to each other (depending upon superposition of classes), it is expected that the proposed method will be more successful than the DWKNN method.

Therefore, data used in this study should have intraclass standard deviation value changing gradually (different from each other in ascending sort). Additionally, datasets must contain the different classes which are very close to each other in an amount changing gradually. So that advantages and disadvantages of the proposed method can reveal.

For this reason, 4 data sets—Seeds, Wine, Balance-Scale and Heart- including these characteristics

Fig. 3 Block diagram of this study.

Table 1 Scatter of data sets.

	Seeds	Wine	Balance scale	Heart
Standard deviation	0.1435	0.2946	0.4739	0.7646

Units are synchronized by (standard-deviation)/(mean).

Table 2 Distances of the elements in data sets to other class elements.

	Seeds	Wine	Balance scale	Heart
Number of data at the same position belonging to different classes	15.36	12.12	5.56	2.29

Units are synchronized by (distance between the classes)/(the mean distance between the elements).

were used by taken from UCI database [11]. Basic characteristics of these data sets are that intraclass scatters (standard deviations) of data sets have gradually changing rates and include gradually changing amount of data belonging to different classes which are at the same or very close position intraclass standard deviation values of data sets used in the study are shown in Table 1.

As seen in Table 1, the data set of seeds has a homogeneous distribution. However, Heart Data Set

has scattered classes. Namely, the higher the values in Table 1 are, the lower the distance of the test elements to the noisy data and outliers.

Information on the distances (telescoping of the classes) between different classes in data sets are shown in Table 2. These values in Table 2 give information relating to the closeness of the test elements to neighbour elements belonging to different classes. Namely, the lower the values in Table 2 are, the lower the distance of the test elements to the noisy

Table 3 General characteristics of data sets.

Data sets	Number of elements	Number of features	Number of classes
Seeds	178	13	3
Wine	210	7	3
Balance scale	625	4	5
Heart	270	14	2

Fig. 4 Sample data set for the proposed method. **Fig. 5** Calculation of the density coefficients of the elements (Cm-Pn means the density coefficient of nth element belonging to mth class).

neighbours belonging to different classes. In other words, it is expected that low values in Table 2 will reduce the success of distance-weighted KNN method. As for Table 3, general characteristics of data sets are shown.

3. Methods

The logic of the proposed method for sample data

set in Fig. 4 is as follows:

The density coefficients of the elements belonging to the same class in the training set are formed according to their distances to each other. In other words, the density coefficients are formed separately according to the distances of C1 elements in Fig. 4 between each other and the distances of C2 elements between each other. As for Fig. 5, the calculation of the density coefficients of data set in Fig. 4 is shown.

As seen in Fig. 5, the density coefficients of the elements in each class were calculated according to other elements in their own class. In other words, for the formation of the coefficient of density of an element, only other elements in its class are determinants. Thus, a separate calculation of the density coefficient is required for each class in the training set.

Following the calculation of the density coefficient, the class of the test element is determined in second stage. This process is carried out according to the sum total of the density coefficients of the elements belonging to each class within the neighbour boundary. In other words, if the neighbour boundary is 5, the total of the density coefficients of the class elements is found for each class in 5 closest neighbours. Then, the test element is included in a class where the total value is greater. How to make the density-weighted KNN classification proposed in the study is given for sample data set shown in Figs. 4, 5 and 6. In this sample data set, the neighbours “k” boundary was selected as 5. Cm-Pn value means the density coefficient of nth element belonging to mth class. In other words, C2-P3 value means the density coefficient of the third element belonging to the class-2.

As seen in Fig. 6, there are 2 elements belonging to the class-1 within the neighbour boundary. The density coefficients of these elements are C1-P2 and C1-P3. In that case, the value of test element belonging to the class-1 is $[(C1-P2) + (C1-P3)]$. In Fig. 6, the elements belonging to the class-2 within the neighbour boundary are C2-P1, C2-P2 and C2-P3. In that case, the value of test element belonging to the secondary will be $[(C2-P1) + (C2-P2) + (C2-P3)]$. Which value is greater from both these values or which sum total of $[(C1-P2) + (C1-P3)]$ and $[(C2-P1) + (C2-P2) + (C2-P3)]$ is greater, the test element belongs to that class.

In this method, the membership of the test element is determined depending upon both its distance to the neighbours and the distances of these neighbours to their own class. If the neighbour elements belonging to any class are away from their own class, the sum total of weights will be low. Accordingly, the possibility of the membership of the test element to that class will decrease. In this way, the proposed method does not lose its advantage in the weighting of elements with their distance to their class centre. On the other hand, the test element uses the advantage of the distance-weighted KNN as the class elements will be within the neighbour boundary of that class to which they are closer. While doing all of these, the extreme closeness between the test element and the neighbours become less important. Because the importance related to the distance of neighbours to the test element has decreased. Moreover, as the training set is weighted for once, there will not be too much waste of time. Additionally, by virtue of the logic of study, the proposed method will be steady against the change of neighbour boundary. How the proposed method eliminate the failure occurring due to neighbours which are very close to the element to be classified and the steadiness of the proposed method against the change of neighbour boundary can be seen in Fig. 7.

As seen in Fig. 7, one of the elements of class-1 is

Fig. 7 A sample data distribution for the relationship between very close neighbours and KNN methods.

Fig. 6 Application of the density-weighted KNN method for 5 neighbours in a sample data set.

very close to the test element. If it is assumed that this test element is a member of class-2, a failure occurs in the distance-weighted KNN method due to the element of class-1 which is very close. In order to fix the failure, the degree of distance weighting should be selected correctly. In other words, correct prediction of weighting degree is required. As for the classical KNN method, there will be different results in 1 to 4 selections of neighbour boundary. For this reason, correct neighbour boundary prediction is required for the correct classification. However, in the proposed method, as the element of class-1 which is very close to the test element is away from the elements of its own class, its weight (density

coefficient) will be low. Moreover, even if the neighbour boundary changes, the result for the proposed method will not change. Because in the proposed method, the evaluation of the test element is made according to all element in the training set. However, it's not the same for the distance-weighted KNN method. In the distance-weighted KNN method, as intraclass standard deviation increases, the importance of the very close neighbours to the test element will increase and the success rate will decrease. In that case, depending upon intraclass standard deviation value, it is expected to be more successful of proposed method than the distance-weighted KNN method.

In this study, the coefficients of density were obtained by Parzen Window Method. According to Parzen Window Method, there is a window "R" selected for finding the density. The size of this window is determined according to the size of data. The window is positioned in a way that element whose coefficient will be calculated will be in the centre of window and the number of other elements within the window is calculated. The application of these procedures is as follows: If the density function is as P in Eq. 1,

$$P(x) = \frac{k/n}{V} \quad (1)$$

In this Eq. 1, x is a sample, n is the number of sample and V is volume (Size). If the window function in Eq. 2 is put into its place in this density function, the density function will be as in Eq. 3 and the coefficients can be determined for each sample.

$$\varphi(u) = \begin{cases} 1, & |u_j| < \frac{1}{2}, j = 1, \dots, d \\ 0, & \text{Otherwise} \end{cases} \quad (2)$$

$$P_\varphi(x) = \frac{1}{n} \sum_{i=1}^{i=n} \frac{1}{h^d} \varphi\left(\frac{x - x_i}{h}\right) \quad (3)$$

d value in Eq. 3 is the size of data set [12-14]. While calculating the density function, the windowing results multiply by the density function, but as the objective is to obtain coefficient for data, this multiplication is not required. In other words, the

calculation of density coefficient is not required in position where there is no data. Additionally, as the coefficients are calculated separately for classes having all features, d value may be neglected. In this case, the density function forming the coefficients is as in Eq. 4.

$$P_\varphi(x) = \frac{1}{h \cdot n} \sum_{i=1}^{i=n} \varphi\left(\frac{x - x_i}{h}\right) \quad (4)$$

As a result, dividing the numerical total of the elements within the window to the number of elements and the size of window will form the coefficients. Namely the coefficient is proportional to elements within the window. The critical selection for this is the size of window. If the size of window is selected too small, the coefficients of density will become peak value consisting of zeros in general. Otherwise, there will be the coefficients having the same values. Both cases are not proper for KNN classification. Because each element should be evaluated by its own characteristic and different coefficients should be taken in the KNN classification. For this reason, the size of window should be adjusted in a way that it gives coefficient to each element and makes different coefficients of elements according to their characteristics as much as possible. In other words, the size of window should be adjusted in a way that the coefficient of elements in the regions where data set is dense (in the centre of data set) will be great and the coefficient of elements away from the centre will be small (Low). If the greatest gap between elements in a data set scattered in space is selected as two times window size, the required conditions are provided. Thus, the coefficients are great in the centre and smaller as becoming distant from the centre. Accordingly, the calculation of the coefficients of density by Parzen Window in the proposed method is as in sample data set in Fig. 8.

As seen in Fig. 8, each element was positioned so as to be in the centre of window. Calculations were made according to the number of other elements within the window positioned.

Fig. 8 Calculation of the coefficients of density in the proposed method.

Table 4 Success rates of KNN methods.

Success rate (%)	Seeds	Wine	Balance scale	Heart
A: KNN	90.5 ± 0.5 (19)	76.2 ± 1.2 (1)	89.5 ± 0.2 (21)	67.4 ± 1.7 (21)
B: Distance-Weighted KNN	90.1 ± 0.7 (19)	75.3 ± 1.2 (1)	83.4 ± 0.2 (19)	62.3 ± 0.8 (21)
C: Density-Weighted KNN	91.6 ± 0.6 (17)	76.9 ± 1.3 (1)	90.4 ± 0.3 (21)	68.0 ± 1.2 (21)
C/B Increase	% 1.19	% 1.0	% 0.96	% 0.91
C/A Increase	% 1.61	% 2.20	% 8.37	% 9.07

4. Experimental Results

In this study, data sets were subjected to the Cross-Validation process in order to evaluate better the performance of classifier. Total 10 cycles were applied and in each cycle, 10% of data which are different from the previous are selected for the test and 90% of them are selected for the training. The results of classifier arising from the application of the proposed method to data sets are as in Table 4. The parenthetical values in Table 4 are the number of neighbours (boundary) giving the highest result, “±” signs are the standard deviation values arising from the random selection of test and training data. These standard deviation values for the performances were generally found close for both the proposed and other KNN methods.

As seen in Table 4, the density-weighted KNN method is more successful than both the classical KNN method and the distance-weighted KNN method. While the contribution of the proposed method to success (5th and 6th column in Table 4) in comparison

with the classical KNN method is inversely proportional to intraclass standard deviations of data sets, it is directly proportional to its contribution to success in comparison with the distance-weighted KNN method. In other words, as intraclass distributions become homogeneous (standard deviation is low), the contribution of the proposed method to success increases in comparison with the classical KNN, but decreases in comparison with the distance-weighted KNN. While the contribution of the proposed method to success is directly proportional to closeness of classes (closeness of test elements to the training elements belonging to different classes), it is inversely proportional to the distance-weighted KNN. In other words, as the closeness of test elements to the training elements of different classes increases, the success of the proposed method decreases in comparison with the classical KNN but increases in comparison with the distance-weighted KNN. In order to understand this better, the relationship between intraclass scatter values of data sets (standard deviations in Table 1) and the contribution of the

proposed method to the classification success (5th and 6th columns in Table 4) can be seen in Fig. 9. While one of axis in Fig. 9 is intraclass standard deviation value of data sets (values in Table 1), other axis is how many times the proposed method is more successful by percent than other KNN methods (5th and 6th columns in Table 4).

As seen in Fig. 9, as intraclass standard deviation value in data sets decreases (as classes become homogeneous), the proposed method provides higher performance by 1.19% than the classical KNN methods. In other words, as intraclass standard deviation increases (as the homogeneity of classes decreases), contribution to success is lower by 0.28%. Because it is more difficult to determine the noises or outliers by the coefficients of density of data distributed (density coefficients will be close or same) [15]. However, it cannot be said that the success of the proposed method in comparison with the classical KNN method is directly associated with intraclass standard deviation. Because the contribution of the proposed method to success in comparison with the classical KNN method does not change too much depending upon the intraclass standard deviation (0.28% value can be neglected). However, the contribution of the proposed method and the standard deviation to success is total opposite for the distance-weighted KNN method and the success change rate is so great that it cannot be neglected. In other words, as intraclass standard deviation increases (as intraclass homogeneities decrease), the proposed method is even more successful than the distance-weighted KNN method. Because scattered intraclass data will increase more the effect of elements which are very close to the test element and decrease the success of the distance-weighted KNN method [16, 17]. Therefore, the proposed method was more successful by 9% than the distance-weighted KNN method.

In addition to all of these information, the relationship of the proposed method between the

closeness of different class elements to each other (information in Table 2) and the classifier success rates can be seen in Fig. 10.

While one of axis in Fig. 10 is the value of closeness of different class elements of data sets (values related to telescoping of classes in Table 2), other axis is how many times the proposed method is more successful by percentage than other KNN methods (5th and 6th columns in Table 4).

In the KNN classification, as training elements belonging to different classes become distant from the test element, the distances of neighbours instead of their number become important [18]. Because the number of training elements belonging to different class very close to the test element decreases. For this reason, the distance-weighted KNN becomes more useful than classical KNN. As seen in Fig. 10, as the neighbour elements belonging to different classes become distant from the test element, the success of the distance-weighted KNN method becomes closer to the success of the proposed method (the gap changed by 9.07% to 1.61%). However, this is total opposite for the classical KNN. This is because the density increases excessively due to more closeness of the neighbour data away from the test element to its own class. Additionally, the contribution of the proposed method to success in comparison with the classical KNN method is not so much associated with the distance of different class elements to each other. Because the distance of different class elements to each other did not affect much the success of the proposed method in comparison with the classical KNN method.

In the light of these informations, in brief, the proposed method is more successful than commonly used current KNN methods. However, the proposed method becomes more important for data sets in which intraclass data is scattered (high intraclass standard deviation) and the distance between different class elements is low (telescoping of classes) especially than the distance-weighted KNN method.

Fig. 9 Superiority of the proposed method to other KNN methods depending upon the standard deviation.

Fig. 10 Superiority of the proposed method to other KNN methods depending upon the interclass distance.

Table 5 Operation time of the proposed method and other KNN methods.

Operation time (second)	KNN	Distance-Weighted KNN	Density-Weighted KNN
Seeds	0.091	0.105	0.108
Wine	0.092	0.097	0.101
Balance scale	0.456	0.493	0.528
Heart	0.184	0.210	0.214
Average	0.205	0.226	0.237

Table 5 should be examined in order to see the operation time of the proposed method.

As seen in Table 5, the operation time of the

proposed method takes longer than other KNN methods. However, the gap is almost 8.5% more for the classical KNN, almost 3.9% more for the

distance-weighted KNN and has acceptable level. In other words, the operation time of the proposed method is not long enough to reduce the efficiency (performance).

5. Discussion and Conclusion

One of the greatest problems encountered in the KNN classification is how to evaluate the elements within the neighbour boundary. Especially the noisy data are very important for this problem. Because the noisy data may cause failure in the distance-weighted KNN method in data sets in which intraclass deviation is high and the neighbour elements are so close to the test element. In the total opposite situation, they may cause failure in the classical KNN in data sets in which intraclass deviation is low and the neighbour elements are so away from the test element.

For this reason, in this study, the new density-weighted KNN method was suggested. The proposed method is affected less by both intraclass standard deviation value and telescoping classes than other methods. Also proposed method is not affected from neighbour boundary. However, as there are not too many failures arising from this kind of data sets in the classical KNN method, the contribution of the proposed method remains about 1%. But as the distance-weighted KNN method is affected more by this kind of data, the contribution of the proposed method to success rises up to 9% in case of high intraclass standard deviation and low distance between different class elements (where the classes telescope). The proposed method is more successful (by 1%) than the classical KNN method and even more successful (up to 9%) than the distance-weighted KNN method according to the structure of data sets.

As a result, the proposed method is offered to data sets having low intraclass standard deviation and low distance between the classes. Moreover, the proposed method is proper for producing higher performance by combining with other density-weighted KNN method in literature.

References

- [1] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z., Steinbach, M., Hand, D. J., and Steinberg, D. 2008. "Top 10 Algorithms in Data Mining." *Knowl. Inf. Syst.* 14: 1-37.
- [2] Lu, C. T., Chen, D., and Kou, Y. 2008. "Detecting Spatial Outliers with Multiple Attributes: Tools with Artificial Intelligence." *Proceedings 15th IEEE International Conference on 2008*, 122-8.
- [3] Wilson, D. R., and Martinez, T. R. 2000. "Reduction Techniques for Instance-Based Learning Algorithms." *Machine Learning* 38: 257-86.
- [4] Gou, J., Du, L., and Xiong, T. 2012. "Weighted K-Nearest Centroid Neighbor Classification." *Journal of Comp. Inf. Sys.* 8: 851-60.
- [5] Bhatia, N. 2010. "Survey of Nearest Neighbor Techniques." *Computer Vision and Pattern Recognition* 8: 302-5.
- [6] Tan, J., Zhang, J., and Li, W. 2010. "An Improved Clustering Algorithm Based on Density Distribution Function." *Computer and Information Science* 3: 23-9.
- [7] Shi, K., Li, L., Liu, H., and He, J. 2011. "An Improved KNN Text Classification Algorithm Based on Density." *Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference*, 113-7.
- [8] Kulkarni, S. G., and Babu, M. V. 2013. "Introspection of Various K-Nearest Neighbor Techniques." *UACEE International Journal of Advances in Computer Science and Its Applications* 3: 103-6.
- [9] Silverman, B. W. 1986. *Density Estimation for Statistics and Data Analysis*. Florida: Chapman & Hall/CRC.
- [10] Duda, R. O., Hart, P. E., and Stork, D. G. 2001. *Pattern Classification*. New York: A Willey-Interscience Publication, John WileySons, INC.
- [11] Charytanowicz, M., Niewczas, J., Kulczycki, P., Kowalski, P. A., Lukasik, S., and Zak, S. 2010. *A Complete Gradient Clustering Algorithm for Features Analysis of X-ray Images, Inf. Tech. in Biom.* EwaPietka and JacekKawa, Springer-Verlag, Berlin-Heidelberg.
- [12] Babich, G. A. 2002. "Weighted Parzen Windows for Pattern Classification." *Pattern Analysis and Machine Intelligence IEEE Transactions* 18: 567-70.
- [13] Wanga, S., Chungb, F., and Xionga, F. 2008. "A Novel Image Thresholding Method Based on Parzen Window Estimate." *Pattern Recognition* 41: 117-29.
- [14] Erdogmus, D., Hild, K. E., Principe, J. C., Lazaro, M., and Santamaria, I. 2004. "Adaptive Blind Deconvolution of Linear Channels Using Renyi's Entropy with Parzen Window Estimation." *IEEE Transactions on Signal*

Processing 52: 1489-98.

- [15] Ester, M., Kriegel, H. P., Sander, J., and Xu, X. 1996. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise." In *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD96)*, 226-31.
- [16] Wilson, D. R., and Martinez, T. R. 1997. "Improved Heterogeneous Distance Functions." *Journal of Artificial Intelligence Research* 6: 1-34.
- [17] Liu, W. and Chawla, S. 2011. "Class Confidence Weighted KNN Algorithms for Imbalanced Data Sets." *Advances in Knowledge Discovery and Data Mining* 6635: 345-56.
- [18] Taneja, S., Gupta, C., Goyal, K., and Gureja, D. 2014. "An Enhanced K-Nearest Neighbor Algorithm Using Information Gain and Clustering." *Fourth International Conference on Advanced Computing & Communication Technologies*, 325-9.