Online Image Segmentation using Credibilistic Fuzzy Clustering

Yevgeniy Bodyanskiy*1*, Alina Shafronenko*1*, Diana Rudenko*1*, Anton Polubiekhin*1*, Dmytro Frolov*1*

1Kharkiv National University of Radio Electronics, Nauky ave 14, Kharkiv, 61166, Ukraine

Abstract

Computational intelligence methods are widely used to solve many complex problems, including, of course, traditional: Data Mining and such new directions as Dynamic Data Mining, Data Stream Mining, Big Data Mining, Web Mining, Text Mining, etc. In the paper was proposed new adaptive on-line methods of fuzzy robust clustering-segmentation of data streams based on probabilistic, possibilistic and credibilistic approaches. Using proposed approach, it's possible to solve clustering task in on-line mode when data are fed to processing sequentially, possible in real time.

Keywords

Segmentation, clustering, data stream, credibilistic approach, fuzzy image segmentation

1. Introduction

The current state of technological development is inextricably linked with the development of computerized tools, which, in turn, are dependent on the mathematical apparatus and practical algorithms that use it. The development of computer tools, in particular hardware, acts as a catalyst for the development of existing and the emergence of new scientific fields, such as Data Science. Modern capabilities of computing environments allow the implementation of algorithmically sufficiently complex methods that are the basis of intellectual analysis. And this should become an impetus for the development of new hardware and software systems based on the theoretical principles of artificial intelligence.

Recently, in the tasks of analyzing and processing non-stationary signals of an arbitrary nature under conditions of uncertainty, computational intelligence methods are increasingly being used, among which hybrid neural networks can be distinguished.

By the task of data segmentation, we will understand the division of the data sample into homogeneous homomorphic segments based on the analysis of changes in the internal properties of the data. Currently, several segmentation methods are known, namely using wavelet analysis [1], fractal-wavelet technologies [2], neuro-fuzzy technologies [3-5], etc.

Depending on the specifics of the problem being solved, two main types of forecasting and segmentation methods can be applied: real-time and batch.

Many neural network architectures, including hybrid structures, are used to solve this kind of problems, but these systems are either cumbersome in their architecture or not sufficiently adapted for real-time learning. In most cases, the activation functions of such networks are sigmoidal functions, splines, polynomials, and radial basis functions.

2. Credibilistic fuzzy clustering

Traditionally, the initial information for the clustering problem is a sample of observations consisting of N n - dimensional feature vectors:

$$X = \{x(1), x(2), ..., x(k), ..., x(N)\}, \ x(k) = (x_1(k), ..., x_n(k))^T \in R^n, \ k = 1, 2, ..., N,$$

and the result of the algorithm is the distribution of the initial data set into m classes with a certain level $w_j(k)$ belonging to the k-th feature vector of the j-th cluster.

At the same time, there is a wide class of problems when the initial information comes not in vector, but in matrix form, i.e.

$$x(k) = \{x_{i_1 i_2}(k)\};$$

where $i_1 = 1, 2, ..., n_1, i_2 = 1, 2, ..., n_2, k = 1, 2, ..., N$. Such situation is characteristic, for example, in image processing [6], when the initial $(N_1 \times N_2)$-matrix is divided into $N = N_1 N_2 (n_1 n_2)^{-1}$ $(n_1 \times n_2)$ fragment matrices that are subject to clustering, because of which the homogeneous in some sense segments of this image. Traditionally, this problem is solved by preliminary vectorization of fragments and the use of already known procedures, the most popular of which is the method of clustering fuzzy C-means [6, 7].

To process matrix data, it is necessary to introduce matrix methods of clustering-segmentation, for which it is advisable to introduce into consideration the matrix method of fuzzy C - means, which is a generalization of FCM. This method will avoid unnecessary vectorization-devectorization operations when processing data given in the form of two-dimensional arrays and provides information processing in online mode.

So, let the sample of observations be given

$$x(k) = \{x_{i_1 i_2}(k)\} \in R^{n_1 \times n_2}, \ k = 1, 2, ..., N,$$

at the same time, for the convenience of further processing, these data are pre-centered relative to the average:

$$\bar{x} = \frac{1}{N} \sum_{k=1}^{N} x(k) \tag{1}$$

and normalized to its spherical norm (Frobenius norm):

$$\|x(k)\| = \sqrt{Tr\left(x(k)x^T(k)\right)}. \tag{2}$$

The matrix probabilistic criterion is used as the objective function of clustering:

$$E(w_j(k), c_j) = \sum_{k=1}^{N} \sum_{j=1}^{m} w_j^\beta(k) D^2(x(k), c_j) = \sum_{k=1}^{N} \sum_{j=1}^{m} w_j^\beta(k) Tr\left((x(k) - c_j)(x(k) - c_j)^T\right), \tag{3}$$

in presence of constraints

$$\sum_{j=1}^{m} w_j(k) = 1 \ \lor \ \sum_{j=1}^{m} w_j(k) - 1 = 0, k = 1, 2, ..., N, \ 0 < \sum_{j=1}^{m} w_j(k) < N, j = 1, 2, ..., m.$$

By introducing the Lagrange function:

$$L(w_j(k), c_j, \lambda(k)) = \sum_{k=1}^{N} \sum_{j=1}^{m} w_j^\beta(k) D^2(x(k), c_j) + \sum_{k=1}^{N} \lambda(k) \left(\sum_{j=1}^{m} w_j(k) - 1\right) =$$

$$= \sum_{k=1}^{N} \left(\sum_{j=1}^{m} w_j^\beta(k) D^2(x(k), c_j) + \lambda(k) \left(\sum_{j=1}^{m} w_j(k) - 1\right)\right), \tag{4}$$

where $\lambda(k)$ – uncertain Lagrange multiplier, and solving the system of Kuhn-Tucker equations:

$$\begin{cases} \dfrac{\partial L(w_j(k),c_j,\lambda(k))}{\partial w_j(k)} = \beta w_j^{\beta-1}(k)D^2(x(k),c_j) + \lambda(k) = 0; \\[2ex] \dfrac{\partial L(w_j(k),c_j,\lambda(k))}{\partial \lambda_j(k)} = \sum\limits_{j=1}^{m} w_j(k) - 1 = 0; \\[2ex] \left\{ \dfrac{\partial L(w_j(k),c_j,\lambda(k))}{\partial c_j(k)} \right\} = -2\sum\limits_{k=1}^{N} w_j^{\beta}(k)(x(k)-c_j) = O, \end{cases}$$

where $\left\{ \dfrac{\partial L(w_j(k),c_j,\lambda(k))}{\partial c_j(k)} \right\}$ – $(n_1 \times n_2)$ - matrix formed from partial derivatives $\dfrac{\partial L(w_j(k),c_j,\lambda(k))}{\partial c_{ji_1i_2}}$; O – matrix of the same dimension formed by zeros, thus, we arrive at the final form of the algorithm:

$$\begin{cases} w_j(k) = \dfrac{(D^2(x(k),c_j))^{\frac{1}{1-\beta}}}{\sum\limits_{l=1}^{m}(D^2(x(k),c_l))^{\frac{1}{1-\beta}}}; \\[3ex] \lambda(k) = -\left(\sum\limits_{l=1}^{m}\left(\beta D^2(x(k),c_l)^{\frac{1}{1-\beta}} \right) \right)^{1-\beta}; \\[3ex] c_j = \dfrac{\sum\limits_{k=1}^{N} w_j^{\beta}(k)x(k)}{\sum\limits_{k=1}^{N} w_j^{\beta}(k)}. \end{cases} \qquad (5)$$

The resulting system gives rise to a wide class of clustering procedures. Thus, if we set $\beta = 2$, we get a simple and effective matrix clustering algorithm [8], which is a generalization of the popular procedure of J. Bezdek [6]:

$$\begin{cases} w_j(k) = \dfrac{(Tr(x(k)-c_j)(x(k)-c_j)^T)^{-1}}{\sum\limits_{l=1}^{m}(Tr(x(k)-c_l)(x(k)-c_l)^T)^{-1}}; \\[3ex] c_j = \dfrac{\sum\limits_{k=1}^{N} w_j^2(k)x(k)}{\sum\limits_{k=1}^{N} u_j^2(k)}, \end{cases} \qquad (6)$$

where Tr – matrix trace symbol.

The main difference between probabilistic and possibilistic approaches is that probabilistic algorithms use relative similarities between objects and clusters, while probabilistic algorithms use absolute similarities.

Instead of the fuzzy partition matrix in the fuzzy C-means algorithm, the possible C-means algorithm uses a $(N \times m)$ - matrix of possibilities or typicality matrix $T = \{t_j(k)\}$, where $t_j(k) \in [0, 1]$ – the possibility that the object $x(k)$ belongs to cluster j.

The possibilistic matrix has only the following limitations:

$$0 < \sum_{j=1}^{m} t_j(k) \le m, \ k = 1, 2, \ldots, N. \tag{7}$$

This means that an object can have a feature vector that contains only values close to zero (usually such objects are considered noise) or only ones.

Krishnapuram, Keller et al proposed the probabilistic *C*-means (PCM) algorithm and two algorithms that combine probabilistic and possibilistic approaches: the probabilistic-possibilistic *C*-means algorithm (FPCM) and the possibilistic-probabilistic *C*-means algorithm (PFCM) [9-11].

In the PCM algorithm, formula (6) was replaced by the expression:

$$\begin{cases} t_j(k) = \cfrac{1}{1 + \left(\cfrac{Tr\left(x(k) - c_j\right)\left(x(k) - c_j\right)^T}{\gamma_j} \right)^{\frac{1}{\beta-1}}}; \\[4em] c_j = \cfrac{\sum\limits_{k=1}^{N} t_j^\beta(k) x(k)}{\sum\limits_{k=1}^{N} t_j^\beta(k)}, \end{cases} \tag{8}$$

where $\gamma_j > 0$ – a constant determined empirically. It can be seen that the calculation of the cluster prototype in formulas (6) and (8) is identical, with the only difference that the matrix of fuzzy partitioning is changed to the matrix of possibilities. The calculation of the possibility of an object belonging to a cluster in formula (8) can be justified as a bell-shaped function presented in Figure 1.
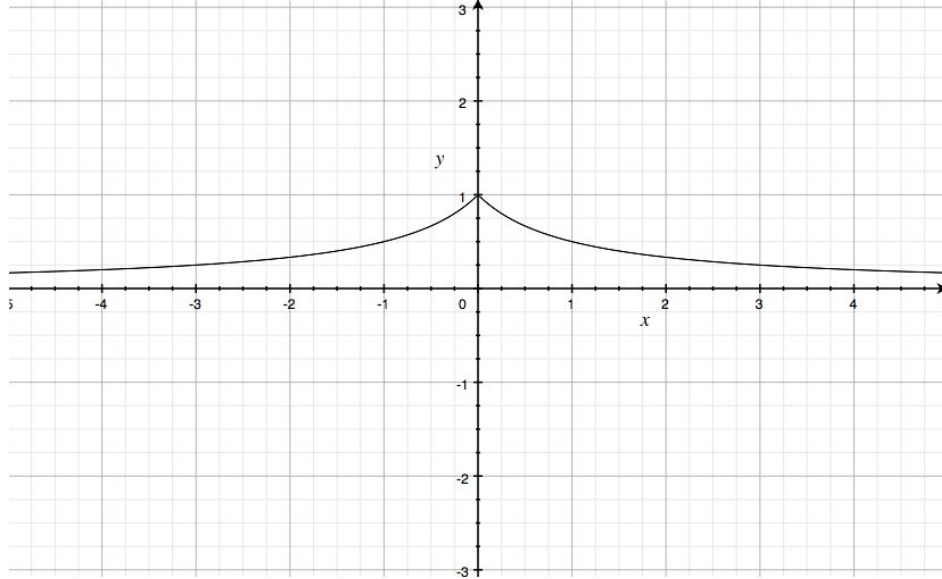

Figure 1: A bell function showing the dependence between membership distances

Keller and Krishnapuram suggested choosing the parameter γ_j in the form:

$$\gamma_j = K \frac{\sum\limits_{k=1}^{N} w_j^\beta(k) Tr\left(x(k) - c_j\right)\left(x(k) - c_j\right)^T}{\sum\limits_{k=1}^{N} w_j^\beta(k)}, \tag{9}$$

where $K > 0$ (most often $K = 1$). But calculations γ_j by formula (9) requires memory to store the

fuzzy partition matrix, as well as time for its use.

The PCM algorithm does a good job of suppressing interference and can usually be applied when it is necessary to improve the results obtained with the help of other algorithms. Also, this algorithm can merge close clusters into one, from which it follows that the initial number of clusters that was set in advance is too large (at the same time, the PCM algorithm can merge clusters that should be separated).

The FPCM and PFCM algorithms use both a fuzzy partition matrix and a feature matrix, trying to take advantage of both approaches.

The FPCM algorithm uses the following formulas:

$$
\begin{cases}
w_j(k) = \dfrac{\left(Tr\left(x(k)-c_j\right)\left(x(k)-c_j\right)^T\right)^{\frac{1}{1-\beta}}}{\sum\limits_{l=1}^{m}\left(Tr\left(x(k)-c_l\right)\left(x(k)-c_l\right)^T\right)^{\frac{1}{1-\beta}}}; \\[4ex]
t_j(k) = \dfrac{\left(Tr\left(x(k)-c_j\right)\left(x(k)-c_j\right)^T\right)^{\frac{1}{1-\eta}}}{\sum\limits_{l=1}^{N}\left(Tr\left(x(l)-c_j\right)\left(x(l)-c_j\right)^T\right)^{\frac{1}{1-\eta}}}; \\[4ex]
c_j = \dfrac{\sum\limits_{k=1}^{N}\left(w_j^{\beta}(k)+t_j^{\eta}(k)\right)x(k)}{\sum\limits_{k=1}^{N}\left(w_j^{\beta}(k)+t_j^{\eta}(k)\right)},
\end{cases}
\tag{10}
$$

where $\eta > 0$ (in most cases $\eta = 2$).

The FPCM algorithm uses the standard procedure for calculating the fuzzy partition matrix, but the possibility matrix is calculated using a new formula. Cluster prototypes are calculated using the sum of both matrices.

The PFCM method uses a standard procedure for calculating the fuzzy partition matrix (as in formula (6)). The procedure for calculating the possibility matrix was taken from PCM (8) and slightly modified. Centroids are calculated as in the FPCM algorithm, but both matrices have their own weights:

$$
\begin{cases}
w_j(k) = \dfrac{\left(Tr\left(x(k)-c_j\right)\left(x(k)-c_j\right)^T\right)^{\frac{1}{1-\beta}}}{\sum\limits_{l=1}^{m}\left(Tr\left(x(k)-c_l\right)\left(x(k)-c_l\right)^T\right)^{\frac{1}{1-\beta}}}; \\[4ex]
t_j(k) = \dfrac{1}{1+\left(b\dfrac{Tr\left(x(k)-c_j\right)\left(x(k)-c_j\right)^T}{\gamma_j}\right)^{\frac{1}{\beta-1}}}; \\[4ex]
c_j = \dfrac{\sum\limits_{k=1}^{N}\left(aw_j^{\beta}(k)+bt_j^{\eta}(k)\right)x(k)}{\sum\limits_{k=1}^{N}\left(aw_j^{\beta}(k)+bt_j^{\eta}(k)\right)},
\end{cases}
\tag{11}
$$

where $a > 0$, $b > 0$. The constants a and b determine the relative importance of the fuzzy

partition matrix and the capability matrix in the centroid calculation function. By setting a = 0, algorithm (11) goes to PCM, and by setting b = 0, algorithm (11) goes to FCM.

Analyzing all the presented methods, several conclusions can be drawn. First, the membership function of the FCM algorithm with its limitations is too "strong", allowing outlier objects to be assigned to one or more clusters, which, in turn, can greatly affect the underlying structure of the data set. On the other hand, the PCM method's constraint on the features is too weak – it allows to refer to a cluster independently of the rest of the data. Also, PCM is very sensitive to the initialization of the capability matrix. The PFCM method is an efficient combination of the two approaches, and the clustering results depend on the parameter setting a, b, β, η.

Algorithm (6) can be extended to the case when data for processing are received sequentially in on-line mode. To do this, by applying the Arrow-Hurwitz-Uzawa saddle point search procedure to the Lagrangian (4), when the $(k+1)^{th}$ observation is received, the estimates of the membership levels and centroids can be refined using recurrence relations [12]

$$\begin{cases} w_j(k+1) = \dfrac{(D^2(x(k+1),c_j(k))^{\frac{1}{1-\beta}})}{\sum\limits_{l=1}^{m}(D^2(x(k+1),c_l(k))^{\frac{1}{1-\beta}})}; \\[4mm] c_j(k+1) = c_j(k) - \eta(k)\left\{\dfrac{\partial L(w_j(k+1),c_j,\lambda(k+1))}{\partial c_j}\right\} = \\[4mm] = c_j(k) + \eta(k)w_j^{\beta}(k+1)(x(k+1)-c_j(k)), \end{cases} \tag{12}$$

for an arbitrary value of the fuzzifier β and

$$\begin{cases} w_j(k+1) = \dfrac{(Tr(x(k+1)-c_j(k))(x(k+1)-c_j(k))^T)^{-1}}{\sum\limits_{l=1}^{m}(Tr(x(k+1)-c_l(k))(x(k+1)-c_l(k))^T)^{-1}}; \\[4mm] c_j(k+1) = c_j(k) + \eta(k)w_j^2(k+1)(x(k+1)-c_j(k)), \end{cases} \tag{13}$$

for β = 2.

It is easy to see that the expression (13) is an adaptive version of the procedure (4), and (13) is, accordingly, (6).

The matrix credibility criterion is used as the objective function of clustering:

$$E\left(Cred_j(k),c_j\right) = \sum_{k=1}^{N}\sum_{j=1}^{m}Cred_j^{\beta}(k)D^2\left(x(k),c_j\right) = \sum_{k=1}^{N}\sum_{j=1}^{m}Cred_j^{\beta}(k)Tr\left((x(k)-c_j)(x(k)-c_j)^T\right) \tag{14}$$

in presence of constraints

$$\begin{cases} 0 \leq Cred_j(k) \leq 1 \forall j,k, \\ \sup Cred_j(k) \geq 0,5 \forall k, \\ Cred_j(k) + \sup Cred_l(k) = 1 \end{cases} \tag{15}$$

where $Cred_j(k)$ - level of observation $x(k)$ credibility.

In the procedures of credibilistic fuzzy clustering, the level of membership is determined by the membership functions [13]:

$$w_j(k) = \varphi_j\left(D\left(x(k),c_j\right)\right) = \varphi_j Tr\left((x(k)-c_j)(x(k)-c_j)^T\right) \tag{16}$$

where φ_j - decreases monotonically on the interval $[0,\infty]$ and with condition $\varphi_j(0)=1, \varphi_j(\infty)\to 0$.

It is easy to see that membership level (16) using the distance is based on similarity measure [14]. As such a measure in [15], it was proposed to use a function:

$$w_j(k) = \frac{1}{1 + Tr\left((x(k)-c_j)(x(k)-c_j)^T\right)}. \tag{17}$$

Thus, if the fuzzy clustering algorithm in a batch form can be written as [16]:

$$\begin{cases} w_j(k) = \dfrac{1}{1 + Tr\left((x(k)-c_j)(x(k)-c_j)^T\right)}, \\[2mm] w_j^*(k) = \dfrac{w_j(k)}{\sup w_l(k)}, \\[2mm] Cred_j(k) = \dfrac{w_j^*(k) + 1 - \sup w_l^*(k)}{2}, \\[2mm] c_j = \dfrac{\sum\limits_{k=1}^{N} Cred_j^{\beta}(k)x(k)}{\sum\limits_{k=1}^{N} Cred_j^{\beta}(k)}, \end{cases} \tag{18}$$

in the online mode this procedure (18) has the form (19):

$$\begin{cases} \sigma_j^2(k+1) = \sum\limits_{\substack{l=1 \\ l\neq j}}^{m}\left(Tr\left((x(k+1)-c_j)(x(k+1)-c_j)^T\right)^{\frac{1}{1-\beta}}\right)^{-1}, \\[4mm] w_q(k+1) = \left(1 + \dfrac{\left(Tr\left((x(k+1)-c_j)(x(k+1)-c_j)^T\right)\right)^{\beta-1}}{\sigma_j^2(k+1)}\right)^{-1}, \\[4mm] w_j^*(k+1) = \dfrac{w_j(k+1)}{\sup w_l(k+1)}, \\[2mm] Cred_j(k+1) = \dfrac{1}{2}\left(w_j^*(k+1) + 1 - \sup w_l^*(k+1)\right), \\[2mm] c_j(k+1) = c_j(k) + \eta(k+1)Cred_j^{\beta}(k+1)\left(x(k+1)-c_j(k)\right) \end{cases} \tag{19}$$

or in case when $\beta = 2$

$$\begin{cases}
\sigma_j^2(k+1) = \left(\sum_{\substack{l=1 \\ l \neq j}}^{m} Tr\left((x(k+1)-c_j)(x(k+1)-c_j)^T \right)^{-2} \right)^{-1}, \\[4mm]
w_q(k+1) = \left(1 + \dfrac{Tr\left((x(k+1)-c_j)(x(k+1)-c_j)^T \right)}{\sigma_j^2(k+1)} \right)^{-1}, \\[4mm]
w_j^*(k+1) = \dfrac{w_q(k+1)}{\sup w_l(k+1)}, \\[4mm]
Cred_j(k+1) = \dfrac{1}{2}\left(w_j^*(k+1) + 1 - \sup w_l^*(k+1) \right), \\[4mm]
c_j(k+1) = c_j(k) + \eta(k+1)Cred_j^2(k+1)\left(x(k+1) - c_j(k) \right).
\end{cases} \qquad (20)$$

It is easy to see that the recurrent fuzzy clustering algorithm is not more complex than the online modifications of probabilistic, possibilistic, and robust procedures [17, 18].

3. Experimental research

Digital images, including satellite images of the city of Kharkiv, were used to test the implemented matrix credibilistic modifications of the clustering algorithm. Samples have no missing attributes and are numeric.

The result of the algorithm is the final fuzzy partition matrix for all sample objects and class prototypes.

When processing digital images, objects (matrices or vectors of the same dimensions) are formed from fragments of this image, and each pixel from the RGB (Red-Green-Blue) color model is converted to the Grayscale model, where the brightness of a pixel is expressed as a scalar value from the interval [0,1]. The conversion from the RGB model to the Grayscale model is performed according to the formula:

$$Y = (0.299R + 0.587G + 0.114B)/255,$$

where Y is the brightness of the pixel glow, R, G, B are the brightness of the glow of red, green, and blue tones, respectively, the values of which are in the interval [0, 255].

Observation sets formed from digital images are processed according to the same principle as standard quantitative samples. After image processing, each cluster is assigned the colors of the Grayscale model, and each object is colored in the color of the nearest cluster.

To evaluate the quality of the algorithm, the following criteria were used: Partition Coefficient (PC), Classification Entropy (CE), Partition Index (PI) with the same initialized fuzzy partition matrix U0.

Table 1 shows the results of the accuracy and speed of the clustering algorithms on the Iris sample, and Table 2 shows the results of the satellite digital image of the city of Kharkiv. The time given is an average for one iteration, considering the vectorization-devectorization operation.

Table 1
Results of cluster analysis on the Iris sample

Methods	PC	CE	PI	Time (c)
FCM	0.531	0.811	12.19	0.003
Matrix method of FCM	0.531	0.811	12.19	0.0025
Matrix method of credibilistic clustering	0.530	0.811	12.18	0.0022

Table 2
The results of the cluster analysis on the digital image

Methods	PC	CE	PI	Time (c)
FCM	0.697	0.419	8.23	1.9
Matrix method of FCM	0.697	0.419	8.23	1.8
Matrix method of credibilistic clustering	0.695	0.419	8.22	1.8

On Figure 2 show the initial image, the resampled sample (20% of objects) and the result of the cluster analysis and the process of the algorithm.
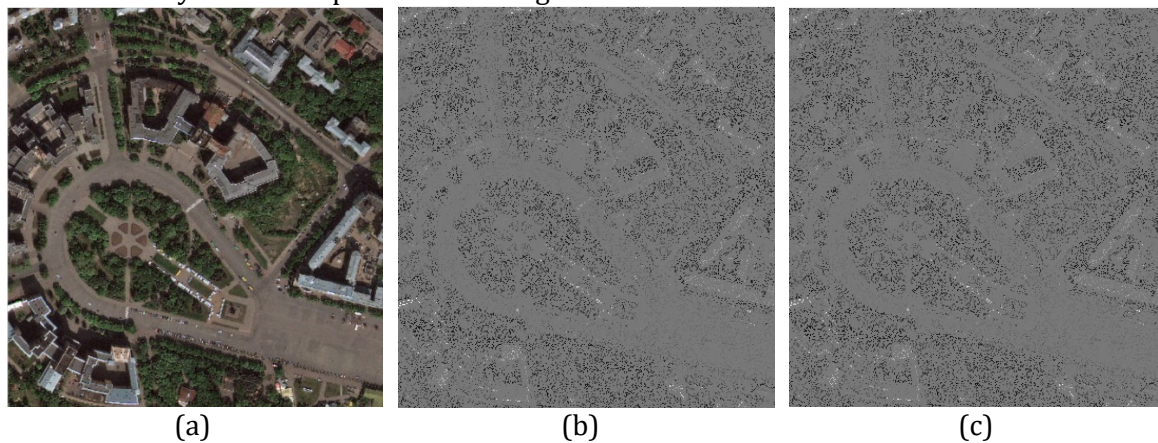

(a)   (b)   (c)

Figure 2: The result of clustering: (a) - Initial digital image for clustering; (b) - Refined sample (20% of objects) for clustering; (c) - Output image of cluster analysis

On Figure 3 shows the result of digital image clustering by the adaptive matrix mrthod of fuzzy credibilistic clustering.
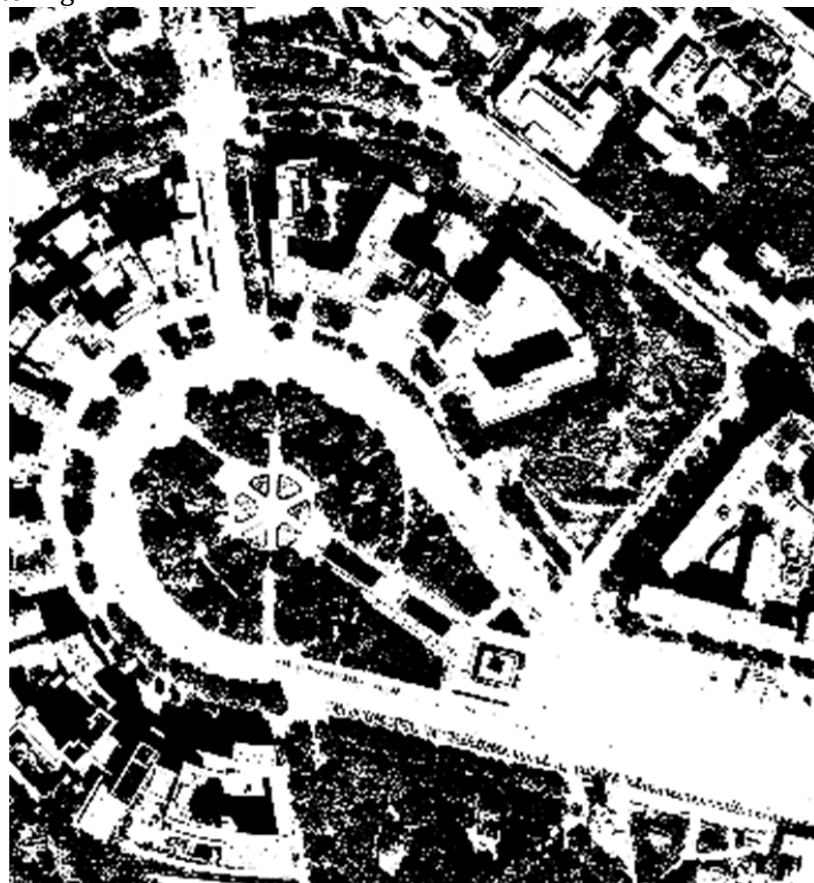

Figure 3: Output image after adaptive cluster analysis

Acknowledgements

References

[1] Chan, K. Efficient time series matching by wavelets, in: Proceedings of 15th IEEE Int. Conf. on Data Engineering, 1999, pp. 126-133.

[2] Ruspini, Enrique H., James C. Bezdek, James M. Keller, Fuzzy clustering: A historical perspective, in: IEEE Computational Intelligence Magazine 14.1, 2019, pp. 45-55.

[3] J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, N.Y.: Plenum Press, 1981.

[4] Oyewole, Gbeminiyi John, and George Alex Thopil, Data clustering: application and trends, Artificial Intelligence Review 56.7, 2023, pp. 6439-6475.

[5] Agrawal, Anurag Vijay, et al. "A probability-based fuzzy algorithm for multi-attribute decision-analysis with application to aviation disaster decision-making." Decision Analytics Journal 8 (2023): 100310.

[6] Wongkhuenkaew, Ritipong, et al. "Fuzzy K-nearest neighbor based dental fluorosis classification using multi-prototype unsupervised possibilistic fuzzy clustering via cuckoo search algorithm." International Journal of Environmental Research and Public Health 20.4 (2023): 3394.

[7] Bodyanskiy, Yevgeniy, Valentyna Volkova, and Mark Skuratov. "Matrix Neuro-Fuzzy Self-Organizing Clustering Network." Computer Science (1407-7493) 50 (2011).

[8] Zhang, Yong, et al. "Possibilistic c-means clustering based on the nearest-neighbour isolation similarity." Journal of Intelligent & Fuzzy Systems 44.2 (2023): 1781-1792.

[9] Hashemi, Seyed Emadedin, Fatemeh Gholian-Jouybari, and Mostafa Hajiaghaei-Keshteli. "A fuzzy C-means algorithm for optimizing data clustering." Expert Systems with Applications 227 (2023): 120377.

[10] Zhang, Yong, et al. "Possibilistic c-means clustering based on the nearest-neighbour isolation similarity." Journal of Intelligent & Fuzzy Systems 44.2 (2023): 1781-1792.

[11] Hussain, Ishtiaq, Kristina P. Sinaga, and Miin-Shen Yang. "Unsupervised multiview fuzzy c-means clustering algorithm." Electronics 12.21 (2023): 4467.

[12] Yang, Miin-Shen, and Josephine BM Benjamin. "Sparse possibilistic c-means clustering with Lasso." Pattern Recognition 138 (2023): 109348.

[13] Bodyanskiy Ye, Shafronenko A., Mashtalir S., Online robust fuzzy clustering of data with omissions using similarity measure of special type, Lecture Notes in Computational Intelligence and Decision Waking-Cham: Springer, 2020, pp.637-646.

[14] Young F.W., Hamer R.M. Theory and Applications of Multidimensional Scaling-Hillsdale, N.J.: Erlbaum, 1994.

[15] Dahiya, Sonika, and Anjana Gosain. "DOIFCM: An Outlier Efficient IFCM." Computational Intelligence in Analytics and Information Systems. Apple Academic Press, 2023. 135-149

[16] Kyrychenko, O., Ostapov, S., & Malyk, I., Cluster Analysis of Information in Complex Networks, in: International Journal of Computing, 22(4), 2023, pp. 515-523. https://doi.org/10.47839/ijc.22.4.3360

[17] Linan, M. N., B. Gerardo, and R. Medina. "Modified weight initialization in the self-organizing map using Nguyen-Widrow initialization algorithm." Journal of Physics: Conference Series. Vol. 1235. No. 1. IOP Publishing, 2019. https://doi.org/10.47839/ijc.19.1.1694

[18] Sirola, Martti, and John Einar Hulsund. "Machine-learning methods in prognosis of ageing phenomena in nuclear power plant components." International Scientific Journal of Computing 20.1 (2021): 11-21.https://doi.org/10.47839/ijc.20.1.2086