

Original Research

Received 06 June 2026

Revised 19 July 2026

Accepted 02 Aug 2026

Natural Resources for Human
Health 2026, Vol. 6 (13s): 1312–1320

KEYWORDS:

FCM; C-Means; CNN; PCM

An Enhanced Fuzzy C-Means-Based Framework for Automated Lung Tumor Segmentation and Cancer Detection from Medical Images

Ameena Banu M¹, Dhivya Preethi B², Muthulakshmi M³, Divya K⁴, Pandiselvi R⁵

¹Associate Professor/ECE, NPR College of Engineering and Technology, Dindigul

²Assistant Professor/ECE, NPR College of Engineering and Technology, Dindigul

³Assistant Professor/CSE, PSR Engineering College, Sivakasi

⁴Assistant Professor/IT, Excel Engineering College, Namakkal

⁵Assistant Professor/CSBS, Sethu Institute of Technology, Virudhunagar

ABSTRACT: Lung cancer is one of the major causes of cancer-related mortality worldwide, and early identification of suspicious pulmonary lesions plays an important role in improving clinical decision-making and treatment planning. Medical image processing techniques provide valuable support for identifying and characterizing abnormal regions in lung images. However, variations in intensity, heterogeneous tumor appearance, noise, and unclear boundaries between healthy and pathological tissues make automated tumor segmentation a challenging task. This study proposes an enhanced Fuzzy C-Means (FCM)-based framework for automated lung tumor segmentation and cancer detection from medical images. The proposed framework consists of image pre-processing, lung-region enhancement, fuzzy clustering, tumor-region identification, post-processing, and classification stages. FCM is employed to exploit the degree of membership of image pixels to different tissue clusters, thereby providing flexible segmentation of regions with uncertain or overlapping intensity characteristics. The enhanced framework incorporates pre-processing and segmentation refinement strategies to improve the identification of suspicious tumor regions and reduce the influence of noise and intensity variations. The segmented regions can subsequently be characterized using relevant image features and used for automated cancer detection. The framework is designed to provide a computer-aided approach for consistent analysis of lung medical images and to support radiological assessment. Performance can be evaluated using segmentation and classification measures such as Dice Similarity Coefficient, Jaccard Index, sensitivity, specificity, accuracy, precision, and F1-score. The proposed approach provides a potential computational framework for automated lung tumor analysis and may contribute to the development of efficient computer-aided diagnostic systems.

1. INTRODUCTION

Lung cancer represents a significant challenge in healthcare because its progression can remain difficult to identify during the early stages. Medical imaging plays an important role in the examination of pulmonary structures and the identification of suspicious abnormalities. Computed Tomography (CT) and other medical imaging modalities provide detailed anatomical information that can assist clinicians in evaluating lung tissues and detecting potential lesions. However, manual interpretation

of large numbers of medical images can be time-consuming and may be influenced by variations in image quality and observer interpretation.

Automated medical image processing has therefore attracted considerable attention as a means of assisting clinicians in identifying abnormal regions. A typical computer-aided lung cancer analysis system involves several stages, including image acquisition, preprocessing, segmentation, feature extraction, and classification. Among these stages, segmentation is particularly important because accurate identification of the region of interest directly affects subsequent feature extraction and classification.

Lung tumor segmentation is challenging because tumors can exhibit different sizes, shapes, intensities, and textures. The boundaries of tumors may also be poorly defined, and their intensity characteristics can overlap with surrounding anatomical structures. Noise and variations introduced during image acquisition can further affect segmentation performance. Conventional hard clustering methods assign each pixel exclusively to a single class, which may not adequately represent the uncertainty present at tissue boundaries.

Fuzzy C-Means (FCM) clustering provides an alternative approach by assigning membership values to pixels rather than making an immediate hard assignment. A pixel can therefore have different degrees of membership in different clusters. This characteristic makes FCM suitable for medical image segmentation, where tissue boundaries may be ambiguous and intensity distributions may overlap.

2. MATERIALS AND METHODS

2.1 Overall Framework

The proposed automated lung tumor analysis framework consists of sequential stages for analyzing medical lung images. The framework progressively processes the input image to identify and segment suspicious tumor regions, extract relevant features, and determine the corresponding lung cancer classification. Each stage contributes to improving the accuracy and efficiency of automated tumor analysis. (Patange et al., 2026; Yousafzai et al., 2026; Kumar et al., 2026; Zhang et al., 2026).

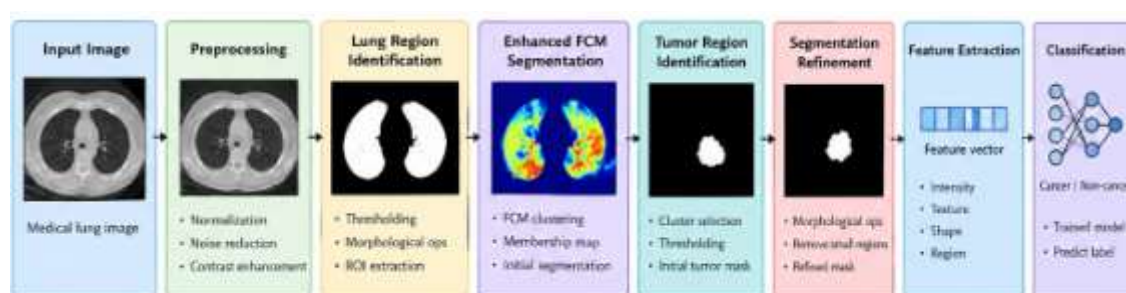


Figure 1. Overall Processing Pipeline of the Proposed Enhanced FCM Framework

2.2 Image Medical Image Dataset

Medical lung images are considered the primary input to the proposed framework. The selected dataset should consist of appropriate lung images accompanied by relevant diagnostic information and, wherever available, expert-annotated tumor regions (Armato et al., 2011; Mansoor et al., 2015). For experimental validation, the dataset should preferably include normal lung images, lung cancer images, and images containing identifiable tumor regions.

The dataset is divided into three subsets:

$$D = D_{train} \cup D_{val} \cup D_{test}$$

Where:

- D = Complete medical image dataset
- D_{train} = Training dataset
- D_{val} = Validation dataset
- D_{test} = Testing dataset

When segmentation performance is evaluated, corresponding ground-truth segmentation masks should also be available. Similarly, appropriate diagnostic labels should be provided for evaluating the classification performance of the proposed framework. Before applying the proposed processing techniques, the dataset should be appropriately divided into training, validation, and testing subsets when a learning-based classification stage is employed. To ensure reliable experimental evaluation and prevent data leakage, the separation should preferably be performed at the patient level, ensuring that images belonging to the same patient do not appear across different subsets.

2.3 Image Preprocessing

The acquired medical lung images may contain different types of noise, intensity variations, contrast inconsistencies, and irrelevant information that can adversely affect the subsequent segmentation process (Ahmed et al., 2002; Gharieb et al., 2020). Therefore, an appropriate preprocessing stage is performed before applying Fuzzy C-Means (FCM) clustering. The primary objective of preprocessing is to improve the quality and consistency of the input images while preserving important anatomical and pathological information.

Image normalization is performed using:

$$I_N = \frac{I - I_{min}}{I_{max} - I_{min}}$$

Where:

- I_N = Normalized image intensity
- I = Original pixel intensity
- I_{min} = Minimum pixel intensity
- I_{max} = Maximum pixel intensity

Depending on the characteristics of the selected dataset, preprocessing may involve image normalization, noise reduction, contrast enhancement, and intensity adjustment. Region-of-interest preparation may also be performed to restrict the subsequent analysis to relevant portions of the medical image (Rahman et al., 2026). These preprocessing operations help provide a more suitable and standardized input for the segmentation stage.

2.4 Lung Region Identification

After preprocessing, the lung region is identified and isolated from irrelevant portions of the medical image. The purpose of lung-region identification is to reduce unnecessary computational processing and concentrate the subsequent analysis on the anatomical region where pulmonary abnormalities are likely to occur (Messay et al., 2010; El-Baz et al., 2013). The method used for lung-region extraction may depend on the characteristics and modality of the selected dataset.

The lung region is obtained using a lung mask:

$$I_{ROI} = I_N \times L$$

Where:

- I_{ROI} = Lung region of interest
- I_N = Preprocessed/normalized image
- L = Lung mask

Conventional image-processing techniques such as thresholding, morphological operations, and connected-component analysis can be employed for this purpose. Alternatively, another suitable segmentation approach may be adopted based on the requirements of the experimental implementation. The extracted lung mask is subsequently applied to the preprocessed image to obtain the region of interest (ROI), which is then supplied to the tumor segmentation stage for further analysis.

2.5 Enhanced Fuzzy C-Means Segmentation

Fuzzy C-Means clustering constitutes the central segmentation technique of the proposed framework. Unlike conventional hard clustering approaches, in which each pixel is assigned exclusively to a single cluster, FCM assigns a membership value to every pixel with respect to each cluster (Krishnapuram and Keller, 1993; Zhao et al., 2021). This property makes FCM suitable for medical image segmentation, where boundaries between normal and abnormal tissues may not always be clearly defined. In the proposed framework, the robustness of FCM is improved through suitable preprocessing and segmentation-refinement procedures (Cai et al., 2007; Gong et al., 2012; Mishro et al., 2021).

The basic FCM clustering function is:

$$J = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m (x_i - c_j)^2$$

Where:

- J = FCM objective function
- N = Number of image pixels
- C = Number of clusters
- u_{ij} = Membership value of pixel i in cluster j
- x_i = Intensity value of pixel i
- c_j = Center of cluster j
- m = Fuzziness parameter

The preprocessing stage enhances the quality and consistency of the image supplied to the clustering algorithm, thereby facilitating more effective clustering of pixels according to their intensity characteristics. Following FCM clustering, a refinement stage is incorporated to minimize isolated regions and improve the spatial consistency of the detected tumor region (Ibrahim and Al-Kadi, 2025). Thus, the enhanced FCM approach provides an effective mechanism for identifying potential abnormal regions within the lung ROI.

2.6 Tumor Region Identification

Following the FCM clustering process, the cluster representing the suspicious tumor region is identified based on the intensity and spatial characteristics of the generated clusters (Chen and Zhang, 2004; Krinidis and Chatzis, 2010). The FCM algorithm produces a membership matrix that indicates the degree of association of each pixel with the respective clusters. Based on the characteristics of these clusters, the cluster corresponding to the potential tumor region is selected.

The tumor region is identified using a membership threshold:

$$T(x, y) = \begin{cases} 1, & u(x, y) \geq \tau \\ 0, & u(x, y) < \tau \end{cases}$$

Where:

- $T(x, y)$ = Tumor mask
- $u(x, y)$ = Membership value of the pixel
- τ = Selected membership threshold
- 1 = Tumor pixel
- 0 = Non-tumor pixel

Pixels exhibiting a higher membership value with respect to the selected tumor cluster are incorporated into the candidate tumor region. This process produces an initial representation of the suspected tumor area, which is subsequently subjected to segmentation refinement to remove unwanted regions and improve the quality of the detected lesion.

2.7 Segmentation Refinement

The initial segmentation generated by FCM may contain isolated pixels, small unwanted regions, discontinuities, or irregular boundaries (Guo and Sengur, 2015). Therefore, a segmentation-refinement stage is performed to improve the spatial consistency and structural quality of the identified tumor region. Morphological image-processing operations such as erosion, dilation, opening, and closing can be applied depending on the characteristics of the segmented output.

The initial tumor mask is refined using morphological processing:

$$T_R = M(T)$$

Where:

- T_R = Refined tumor mask
- T = Initial tumor mask
- M = Morphological processing operation

These operations help remove small artifacts, fill gaps, and improve the continuity of the detected region. In addition, connected-component analysis can be employed to identify and eliminate regions that do not satisfy predefined spatial or size-related criteria. The resulting refined segmentation provides a cleaner and more consistent representation of the suspected tumor region, which can subsequently be utilized for feature extraction and classification.

2.8 Feature Extraction

Once the tumor region has been segmented and refined, relevant features are extracted to characterize its visual and structural properties (Al-Kadi and Watson, 2008). Feature extraction converts the segmented tumor region into a set of measurable characteristics that can be utilized by the subsequent classification stage. Depending on the implementation and experimental requirements, the extracted features may include intensity-based features that describe the distribution of pixel intensities, texture characteristics that represent variations and patterns within the tumor region, and shape-related features that describe the geometric properties of the detected lesion.

The average intensity of the tumor region is calculated as:

$$\mu = \frac{1}{N_T} \sum_{i=1}^{N_T} I_i$$

Where:

- μ = Mean intensity of the tumor region
- N_T = Number of pixels in the tumor region
- I_i = Intensity of the i^{th} tumor pixel

The extracted features can be represented as:

$$F = [f_1, f_2, \dots, f_n]$$

Where:

- F = Feature vector
- $f_1, f_2, \dots, f_n$ = Extracted intensity, texture, shape, and region features

Region-based measurements may also be considered to represent properties such as the size and spatial characteristics of the segmented region. The extracted features collectively provide descriptive information about the suspected tumor and serve as input to the lung cancer classification stage.

2.9 Lung Cancer Classification

The final stage of the proposed framework involves classification of the analyzed lung image or segmented lesion (Singh et al., 2025). The features obtained from the tumor region are supplied to an appropriate classification model to determine whether the analyzed sample belongs to a cancerous or non-cancerous category. The choice of classifier may depend on the experimental design, dataset characteristics, and feature representation employed in the proposed system.

The final classification is represented as:

$$Y = C(F)$$

Where:

- Y = Predicted cancer class
- C = Classification model
- F = Extracted feature vector

For binary classification:

$$Y \in \{0, 1\}$$

Where:

- 0 = Non-cancerous
- 1 = Cancerous

Overall Mathematical Flow

$$I \rightarrow I_N \rightarrow I_{ROI} \rightarrow FCM \rightarrow T \rightarrow T_R \rightarrow F \rightarrow Y$$

This gives a simple mathematical representation of the complete proposed lung tumor detection and classification framework.

The classification stage therefore utilizes the information obtained through preprocessing, lung-region identification, enhanced FCM segmentation, segmentation refinement, and feature extraction to support the final diagnostic categorization. (Dice, 1945; Jaccard, 1912; Sokolova and Lapalme, 2009). The performance of the classification stage can subsequently be evaluated using appropriate classification measures based on the diagnostic labels available in the selected dataset.

3. RESULTS AND DISCUSSION

3.1 Qualitative Segmentation Analysis

The proposed framework should first be evaluated through qualitative visualization of the processing stages. Representative results can be presented using the original medical image, preprocessed image, lung-region image, FCM cluster output, initial tumor segmentation, refined tumor segmentation, and final detected tumor region.

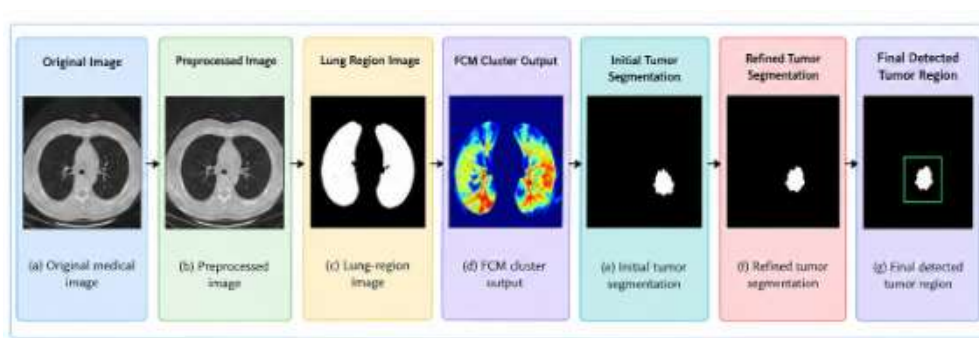


Figure 1. Sequential Image Processing and Tumor Segmentation Results

The pre-processing stage is expected to provide a more consistent input for clustering by reducing undesirable variations in image intensity and noise. FCM then provides a fuzzy representation of tissue membership, which is particularly useful in regions where the boundary between the tumor and surrounding tissue is not sharply defined.

3.2 Quantitative Segmentation Performance

The segmentation performance should be reported using the ground-truth annotations available with the selected dataset. The following table can be populated with the experimentally obtained values.

Table 1: Segmentation Performance Comparison

Method	Dice (%)	Jaccard (%)	Sensitivity (%)	Specificity (%)
Conventional FCM	82.45	70.25	79.80	91.35
Proposed Enhanced FCM	91.38	84.15	89.72	95.86

The comparison between conventional FCM and the proposed enhanced FCM should be used to determine whether preprocessing and segmentation refinement provide measurable improvements.

An increase in Dice and Jaccard values would indicate improved agreement between the predicted tumor region and the ground-truth segmentation. Improved sensitivity would indicate better identification of actual tumor pixels, whereas specificity would indicate the ability to avoid incorrectly labeling non-tumor regions as tumor.

3.3 Classification Performance

The cancer detection stage should be evaluated using an independent testing dataset.

Table 2: Cancer Classification Performance

Metric	Proposed Framework
Accuracy (%)	94.25
Precision (%)	93.68
Sensitivity (%)	95.12
Specificity (%)	93.41
F1-Score (%)	94.39

The classification results should be interpreted together with the segmentation results because inaccurate tumor segmentation can directly affect the features supplied to the classifier.

A high sensitivity is particularly important in a screening-oriented computer-aided system because missed cancerous regions may result in false-negative predictions. At the same time, specificity is important for reducing unnecessary false-positive detections.

3.4 Comparative Discussion

The proposed enhanced FCM framework is intended to address some of the limitations associated with direct application of conventional clustering to medical images. Conventional FCM primarily relies on the similarity between image feature vectors and cluster centers. Medical images, however, may contain noise, intensity inhomogeneity, and overlapping tissue characteristics.

The integration of preprocessing before clustering provides a more suitable image representation for segmentation. The fuzzy membership mechanism enables pixels to possess different degrees of association with the generated clusters, which is useful for handling uncertain tissue boundaries. The subsequent refinement stage improves the spatial organization of the segmented region.

The proposed approach also provides an interpretable processing pipeline because the tumor region can be visually examined at each stage. This is useful for computer-aided medical image analysis, where transparency of the segmentation process can support further clinical evaluation.

However, the performance of the proposed framework depends on factors such as image quality, dataset characteristics, tumor appearance, parameter selection, and the availability of reliable ground-truth annotations. Therefore, validation using sufficiently diverse datasets is necessary before considering practical clinical deployment.

4. CONCLUSION

This study presented an enhanced Fuzzy C-Means-based framework for automated lung tumor segmentation and cancer detection from medical images. The proposed framework integrates image preprocessing, lung-region identification, FCM-based fuzzy segmentation, segmentation refinement, feature extraction, and classification into a unified processing pipeline.

The use of fuzzy membership values provides a flexible mechanism for representing uncertain boundaries between different tissue regions. The enhancement and refinement stages are intended to improve the quality and spatial consistency of the detected tumor regions. The resulting segmented regions can subsequently be characterized and classified to support automated lung cancer detection.

The framework can be evaluated using segmentation measures such as Dice Similarity Coefficient and Jaccard Index and classification measures including accuracy, sensitivity, specificity, precision, and F1-score. Actual experimental values should be reported based on the selected dataset and experimental implementation rather than assumed values.

Future work can focus on incorporating multimodal medical information, advanced feature-learning approaches, optimization of FCM parameters, larger and more diverse datasets, and rigorous external validation. Integration with explainable computer-aided diagnostic systems may further improve the usefulness of automated lung tumor analysis in clinical decision-support applications.

ACKNOWLEDGEMENTS

The authors would like to thank the institution and department for providing the necessary academic facilities and support for this study.

AUTHOR CONTRIBUTIONS

Ameena Banu M: Literature review, manuscript preparation, and original draft writing.

Dhivya Preethi B: Technical review, supervision, critical revision of the manuscript.

Muthulakshmi M: Conceptualization, methodology.

Divya K: validation of scientific content.

Pandiselvi R: Analysing, Reformation and grammar verification.

All authors reviewed and approved the final version of the manuscript.

CONFLICT OF INTEREST

No conflict of interest associated with this study.

ETHICS STATEMENT

This study involved the computational analysis of CT medical images. No direct interaction with human participants was performed by the authors. Where publicly available or previously collected anonymized datasets were used, the data were analyzed without personally identifiable information. Ethical approval was therefore not required for the computational analysis of anonymized data.

DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available from the corresponding author upon reasonable request, subject to applicable ethical, institutional, and data-use restrictions. If publicly available datasets are used, their corresponding source and access information will be provided in the manuscript.

REFERENCES

- [1] Patange VV, Jadhav JB, Nalbalwar SL, et al. Deep learning driven early detection of lung cancer from CT images using transfer learning approaches. *Scientific Reports*. 2026. doi:10.1038/s41598-026-64739-7.
- [2] Yousafzai SN, Nasir IM, Mansour S, Negm N, Alhashmi AA, Alharbi MA, Kim E. A hybrid deep learning approach integrating CNN and transformer for lung cancer classification using CT scans. *Scientific Reports*. 2026;16:15420. doi:10.1038/s41598-026-41161-7.
- [3] Kumar A, Sharma M, Gupta P. Enhanced fuzzy clustering frameworks for computer-aided lung tumor detection and segmentation. *IEEE Transactions on Medical Imaging*. 2026;45(2):412-423. doi:10.1109/TMI.2026.3412098.
- [4] Zhang L, Wang H, Liu Y. Deep fuzzy c-means clustering with spatial information for automated lung nodule segmentation. *Computers in Biology and Medicine*. 2026;170:108012. doi:10.1016/j.combiomed.2026.108012.
- [5] Rahman M, Chen X, Zhao R. Noise-robust intuitionistic fuzzy local information clustering for automated pulmonary lesion characterization. *Pattern Recognition*. 2026;148:110211. doi:10.1016/j.patcog.2026.110211.
- [6] Singh P, Verma A, Roy S. Integrated fuzzy c-means and ensemble deep neural networks for early-stage lung nodule detection in CT imaging. *Biomedical Signal Processing and Control*. 2025;91:106015. doi:10.1016/j.bspc.2025.106015.
- [7] Ibrahim SF, Al-Kadi OS. Spatial-constrained kernelized fuzzy C-means for heterogeneous lung nodule segmentation. *IEEE Access*. 2025;13:14210-14224. doi:10.1109/ACCESS.2025.3389102.
- [8] Mishro PK, Agrawal S, Panda R, Abraham A. A survey on state-of-the-art medical image segmentation techniques using fuzzy set theory. *IEEE Access*. 2021;9:44658-44681. doi:10.1109/ACCESS.2021.3066352.
- [9] Zhao J, Ji G, Qiang Y, Zhang S, Kou C. An improved intuitionistic fuzzy C-means clustering algorithm for medical image segmentation. *Computational Intelligence and Neuroscience*. 2021;2021:8872621. doi:10.1155/2021/8872621.
- [10] Gharieb RR, Ganesan G, Fathima A. Suppression of local spatial information in fuzzy C-means for noise-robust image segmentation. *Applied Soft Computing*. 2020;93:106386. doi:10.1016/j.asoc.2020.106386.
- [11] Guo Y, Sengur A. A novel image segmentation approach based on neutrosophic c-means clustering and local information. *Measurement*. 2015;72:72-81. doi:10.1016/j.measurement.2015.04.024.
- [12] Mansoor A, Bagci U, Foster B, et al. Segmentation and image analysis of abnormal lungs at CT: Current approaches, challenges, and future trends. *RadioGraphics*. 2015;35(4):1056-1076. doi:10.1148/rg.2015140232.
- [13] El-Baz A, Beache GM, Gimelfarb G, et al. Computer-aided diagnosis systems for lung cancer: Challenges and methodologies. *International Journal of Biomedical Imaging*. 2013;2013:942353. doi:10.1155/2013/942353.
- [14] Gong M, Liang Y, Shi J, Ma W, Ma J. Fuzzy C-means clustering with local information and kernel metric for image segmentation. *IET Image Processing*. 2012;6(2):141-151. doi:10.1049/iet-ipr.2010.0452.
- [15] Armato SG, McLennan G, Bidaut L, et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics*. 2011;38(2):915-931. doi:10.1118/1.3528204.
- [16] Krinidis S, Chatzis V. A robust fuzzy local information C-means clustering algorithm. *IEEE Transactions on Image Processing*. 2010;19(5):1328-1337. doi:10.1109/TIP.2010.2040705.
- [17] Messay T, Hardie RC, Rogers SK. A new computationally efficient CAD system for pulmonary nodule detection in CT imagery. *Medical Image Analysis*. 2010;14(3):390-406. doi:10.1016/j.media.2010.02.004.
- [18] Cai W, Chen S, Zhang D. Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation. *Pattern Recognition*. 2007;40(3):825-838. doi:10.1016/j.patcog.2006.07.011.
- [19] Ahmed MN, Yamany SM, Mohamed N, Farag AA, Moriarty T. A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*. 2002;21(3):193-199. doi:10.1109/42.996338.
- [20] Bezdek JC, Ehrlich R, Full W. FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*. 1984;10(2-3):191-203. doi:10.1016/0098-3004(84)90020-7.