*Retraction*

# Retracted: Artificial Intelligence-Guided Subspace Clustering Algorithm for Glioma Images

## Journal of Healthcare Engineering

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope

(2) Discrepancies in the description of the research reported

(3) Discrepancies between the availability of data and the research described

(4) Inappropriate citations

(5) Incoherent, meaningless and/or irrelevant content included in the article

(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

In addition, our investigation has also shown that one or more of the following human-subject reporting requirements has not been met in this article: ethical approval by an Institutional Review Board (IRB) committee or equivalent, patient/participant consent to participate, and/or agreement to publish patient/participant details (where relevant).

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

## References

[1] Y. Zhang, Y. Zhou, Z. Liao, G. Liu, and K. Guo, "Artificial Intelligence-Guided Subspace Clustering Algorithm for Glioma Images," *Journal of Healthcare Engineering*, vol. 2021, Article ID 5573010, 9 pages, 2021.

*Research Article*

# Artificial Intelligence-Guided Subspace Clustering Algorithm for Glioma Images

**Yong Zhang** [iD]**,[1] Yu-mei Zhou,[2] Zhen-hong Liao,[1] Gao-yuan Liu,[1] and Kai-can Guo[1]**

[1]*Department of Radiology, People's Hospital of Deyang City, Deyang City, Sichuan 618000, China*
[2]*Outpatient Department, People's Hospital of Deyang City, Deyang City, Sichuan 618000, China*

Correspondence should be addressed to Yong Zhang; yzhang210@cumt.edu.cn

In order to improve the accuracy of glioma segmentation, a multimodal MRI glioma segmentation algorithm based on superpixels is proposed. Aiming at the current unsupervised feature extraction methods in MRI brain tumor segmentation that cannot adapt to the differences in brain tumor images, an MRI brain tumor segmentation method based on multimodal 3D convolutional neural networks (CNNs) feature extraction is proposed. First, the multimodal MRI is oversegmented into a series of superpixels that are uniform, compact, and exactly fit the image boundary. Then, a dynamic region merging algorithm based on sequential probability ratio hypothesis testing is applied to gradually merge the generated superpixels to form dozens of statistically significant regions. Finally, these regions are postprocessed to obtain the segmentation results of each organization of GBM. Combine 2D multimodal MRI images into 3D original features and extract features through 3D-CNNs, which is more conducive to extracting the difference information between the modalities, removing redundant interference information between the modalities, and reducing the original features at the same time. The size of the neighborhood can adapt to the difference of tumor size in different image layers of the same patient and further improve the segmentation accuracy of MRI brain tumors. The experimental results prove that it can adapt to the differences and variability between the modalities of different patients to improve the segmentation accuracy of brain tumors.

## 1. Introduction

Glioma is the most common primary brain tumor which originates from glial cells. Because of its characteristic of infiltrating the surrounding tissues, it is difficult to be completely removed by surgery [1]. According to its degree of malignancy, it can be further divided into low-grade gliomas (LGG) and high-grade gliomas (HGG). HGGs are highly aggressive and usually lead to poor prognostic treatment, while LGGs are less aggressive and tend to have a better prognosis than HGGs [2]. Gliomas are common and aggressive in adults, and their five-year survival rate is only 10% among the highest grade glioblastomas. The first-line treatment strategy for glioma is to remove the tumor as much as possible and then supplement it with radiotherapy or adjuvant chemotherapy [3]. Among them, radiotherapy occupies a core position in the treatment of brain tumors.

However, this conventional treatment method often leads to the most common side effect of glioma patients within two years, that is, radiation necrosis of glioma [4]. Unfortunately, the recurrence of glioma will also appear at this time. The recurrence and necrosis of glioma appear similar on conventional images, and it is difficult to distinguish them. It is very important to distinguish between recurrence and necrosis of glioma at an early stage because the treatment strategies of the two are completely different. In the clinic, radiologists manually outline the target area of the lesion on medical images and then conduct targeted research and treatment [5]. As we all know, manually delineating the lesion is a time-consuming and labor-intensive task and relies on the doctor's experience. Therefore, the requirements for semiautomatic or automatic glioma clustering methods are very high in clinical practice. Patients with gliomas need frequent follow-up examinations. MRI is a

more commonly used imaging technique. Different MRI imaging sequences such as $T_2$-weighted fluid attenuated inversion recovery, $T_1$-weighted contrast-enhanced ($T_1C$) and $T_2$-weighted ($T_2$) can provide complementary information for the diagnosis of glioma. Nevertheless, the clustering of glioma lesions is still a challenging task [6]. Gliomas come in different sizes and shapes and may appear in different locations in the brain. The gray level of the MRI image of glioma changes unevenly, and due to the aggressive growth of glioma, its edge is blurred. In addition, artifacts and noise in brain MRI images increase the difficulty of clustering gliomas. Clinically, the method to distinguish the recurrence and necrosis of glioma is usually a follow-up, biopsy, and surgical operation. Because pathological diagnosis will bring huge economic pressure and physical and mental harm to patients with glioma, some scholars use MRI images to study glioma [7]. Magnetic resonance spectroscopy, $T_1C$ and weighted images are used to identify the necrosis and recurrence of gliomas. However, most previous studies have used MRI image information of a single modality. In addition, during the follow-up examination of glioma patients, the most commonly used MRI modalities are $T_1$, $T_1C$, $T_2$, and FLAIR [8]. Single-modal MRI images can represent part of the tumor information, while multimodal MRI images can represent the overall information of the tumor. Therefore, combining different MRI modal images can improve tumor discrimination and better reflect the degree of tumor invasion.

In this paper, compared with traditional pixels and voxels, superpixels and super voxels have the advantages of high computational efficiency, conform to human visual perception, and can efficiently represent image information. In order to make full use of the advantages of superpixels and improve the accuracy and robustness of GBM clustering, this paper proposes a multimodal MRI-GBM clustering algorithm based on superpixels. First, through the local k-means clustering algorithm with weighted distance, the multimodal MRI image is overclustered into a series of uniform, compact superpixels that can fit the edge of the image well. Then, the dynamic region merging algorithm is used to merge the superpixels step by step. Finally, the final clustering result of GBM is obtained through postprocessing. Experiments show that this algorithm can achieve better clustering results.

## 2. Related Work

Glioblastoma multiforme (GBM) is the most common type of malignant tumor with the highest mortality rate among brain tumors. Statistics show that 40% of brain tumor patients are gliomas. The median survival time of patients with glioma is only 8 months, and the survival rate of more than 5 years is almost zero [9]. GBM presents a heterogeneous tumor area in the multimodal MRI image. This area usually includes 3 parts: necrosis, enhanced tumor, and edema formed by the tumor squeezing the surrounding normal brain tissue. Due to the complexity and particularity of GBM tumor tissue morphology, monomodal MRI cannot clearly reflect the different tissue structures of GBM. In contrast,

multimodal MRI images contain rich tissue structure information and are widely used in the diagnosis and treatment of GBM. The multimodal MRI in this article mainly includes $T_2$ ($T_2$-weighted imaging), $T_1PRE$ ($T_1$-weighted imaging), $T_1POST$ ($T_1$ strong imaging), and FLAIR (liquid attenuation inversion imaging) [10]. Under different modalities, GBM tissue images showed different characteristics: active tumors showed high signals at $T_1POST$, necrotic parts showed low signals at $T_1POST$, and edema parts showed high signals at $T_2$ and FLAIR.

The clustering of GBM refers to marking or clustering GBM tissues from normal brain tissues based on these characteristics. In the literature, clustering algorithms based on pixels or voxels are widely used in GBM clustering. The basic idea of this kind of algorithm is to classify the pixel into the corresponding category according to the brightness information and texture information of each pixel on the multimodal image. Classification algorithms include unsupervised clustering and supervised learning. For example, Latha and Surya [11] proposed a fuzzy mean clustering algorithm based on FCM (fuzzy C-means algorithm). The algorithm uses the gray level of the multimodal image as the feature vector. First, it uses FCM to cluster all voxel points to obtain the initial classification, and then optimizes the initial classification based on prior knowledge such as symmetry and gray distribution to obtain the final clustering results. Since FCM clustering does not consider the spatial neighborhood information, and the gray distribution of GBM organization will overlap, it is easy to cause misclustering. Algorithms based on graph clustering [12, 13] are also very popular now. This type of algorithm uses the vertices of the graph to describe the pixels of the image and uses the edges of the graph to describe the similarity of two pixels, thereby forming a network graph. Then, by solving the energy minimization problem, the graph is clustered into subnetwork graphs to make different subnetwork graphs. The difference between network diagrams and the similarity within the same subnetwork diagram reach the maximum. This type of algorithm usually needs to solve a generalized feature vector problem. When the image is relatively large, this type of algorithm will encounter the problem of high computational complexity. In addition to the above two types of algorithms, the clustering algorithm based on level set is also widely used in GBM clustering. However, due to the uneven gray level of GBM tissue and the lack of obvious boundary between GBM organizations, it is easy to use this type of algorithm. Recently, superpixels or supervoxels have attracted a lot of attention, and clustering algorithms based on superpixels and supervoxels have also become a research hotspot.

## 3. Multimodal 3D-CNNs Research Methods

*3.1. Algorithm Framework.* The algorithm consists of 4 steps: (1) Image preprocessing, including coregistration of multimodal images and deheading; (2) Using the local k-means clustering algorithm with weighted distance proposed in this paper to cluster multimodal images into a series of brightness. The superpixels are uniform and can fit the edge

of the image well; (3) Using the dynamic area merging algorithm to merge the superpixels generated by (2). Generate several uniform and meaningful regions through region merging (different regions after merging are represented by different colors; (4) Postprocess the merged regions according to brightness distribution, etc., and finally complete the clustering of GBM. Figure 1 shows the algorithm framework of this paper.

### 3.2. Pretreatment.
Preprocessing includes registration of multimodal images and heading. The common registration of images ensures that the pixels at the same position in the multimodal image correspond to the same brain tissue. Select $T_2$ as the reference image and register the images of other modalities to $T_2$. The registration adopts rigid transformation and uses mutual information as the measure of image similarity. Deheading is a common step in brain image processing. On the one hand, it can reduce the amount of calculation for subsequent processing. On the other hand, it can also reduce the impact of nonbrain parenchymal tissue on subsequent processing. This paper uses a template image and the image after the skull is removed to realize the operation of removing the skull through the registration method. First, the template brain image is registered to $T_2$ through affine transformation. Then use the affine transformation matrix generated in the first step of registration to transform the template image after the skull is removed to realize the operation of removing the skull from $T_2$.

### 3.3. Basic Principles of CNNs.
Since CNNs [8] was first proposed in 1998, it has been widely valued by researchers as an efficient identification method. With the concept of deep learning [9] in 2006, CNNs, as a representative of supervised deep learning, once again became one of the research hotspots in many disciplines. In the field of image recognition, the network directly inputs the original image without the need. The complex preprocessing of the image has been widely used. CNNs use three methods to achieve feature extraction [10]: local receptive field, weight sharing, and subsampling. The local receptive field means that the neurons in each network layer are only connected to the neural unit in a small neighborhood of the previous layer. Through the local receptive field, each neuron can extract primary visual features, such as direction line segments and endpoints. Weight sharing makes convolutional neural networks have fewer parameters and require relatively little training data. Subsampling can reduce the resolution of features and achieve invariance to displacement, scaling, and other forms of distortion. In the convolutional layer, the feature map of the previous layer is convolved with a learnable kernel, and the result of the convolution passes the activation function to obtain the feature map of this layer. Commonly used activation functions are the Sigmoid function and hyperbolic tangent function [11]. The hyperbolic tangent function is as in formula (1), when $\alpha = 1.715\,9$ and $\beta = 2/3$, $f(\pm 1) = \pm 1$. Each output feature map may establish a relationship with the convolution of several

feature maps in the previous layer. Generally, the form of the convolutional layer is shown in the following formula:

$$\alpha = \frac{f(x)}{\tan g(\beta x)}, \tag{1}$$

$$f\left(\alpha_k^j\right)^{-1} = \sum \alpha_i^{j-1} {\gamma'}_{i,k}^{\,j} + \beta_i^j. \tag{2}$$

In the formula, $k$ represents the number of layers, $\gamma'$ is the convolution kernel, and $i$ represents a choice of the input feature map. Each output graph has an offset $\beta$. The subsampling layer performs sampling operations on the input. If there are $n$ feature maps in and out, then the number of feature maps after the subsampling layer is still $n$, but the output feature maps should be smaller. The general form of the subsampling layer is as follows:

$$f\left(\alpha_k^j\right)^{-1} = \lambda_k^j \mathrm{sample}\left(\alpha_k^{j-1}\right) + \beta_k^j. \tag{3}$$

In formula (3), sample( )represents the subsampling function. The subsampling function is generally to sum an $n \times n$ area of the input image of this layer, so the size of the output image is $1 \times n$ of the input image size. Each output feature map has its own $\beta$ and $\lambda$.

### 3.4. Multimodal 3D-CNNs Research Methods.
If the original input layer is $32 \times 32$, the original input is, respectively, subjected to $65 \times 5$ neighborhood convolutions to obtain the C1 layer, which contains 6 feature map; $C_1$ layer is downsampled to obtain $S_2$ layer, its size is $28 \times 28$; after two convolutions and downsampling, one-dimensional feature $F_5$ layer is obtained, and then the radial basis function is used to classify the features, and the classification result is reversed. Propagate to modify the convolution weights and biases of each layer to form a supervised deep learning algorithm. Classical 2D-CNNs are mainly used for digital recognition. The input image can always have a fixed size and then extract features from the entire image and apply it to MRI brain tumor clustering. The following problems will arise: First, the brain tumor image for clustering, a single pixel, must be classified, so the original input can only be the neighborhood of a single pixel, and the size of this neighborhood is difficult to grasp; secondly, different patients have different brain tumors and there are different image layers of brain tumors in the same patient [14]. Even if the neighborhood value of the original input layer is determined through the training layer, it is difficult to ensure that this neighborhood is suitable for all tumor points of this patient; third, how to make full use of the multimodal information of MRI to achieve the higher classification of accuracy. In order to solve the above problems, the following improvements are made to the classic 2D-CNNs.

Multimodal 3D-CNNs are shown in Figure 2, taking small neighborhoods of four modals at the same position, such as $14 \times 14$ (the specific neighborhood size is obtained by optimizing the training data grid) to form a 3D ($14 \times 14 \times 4$). For the original input layer, convolve the original input layer with a convolution template with a size of $3 \times 3 \times 2$ shared by 6 weights to obtain $612 \times 12 \times 3$ feature maps $C_1$; 6 features of
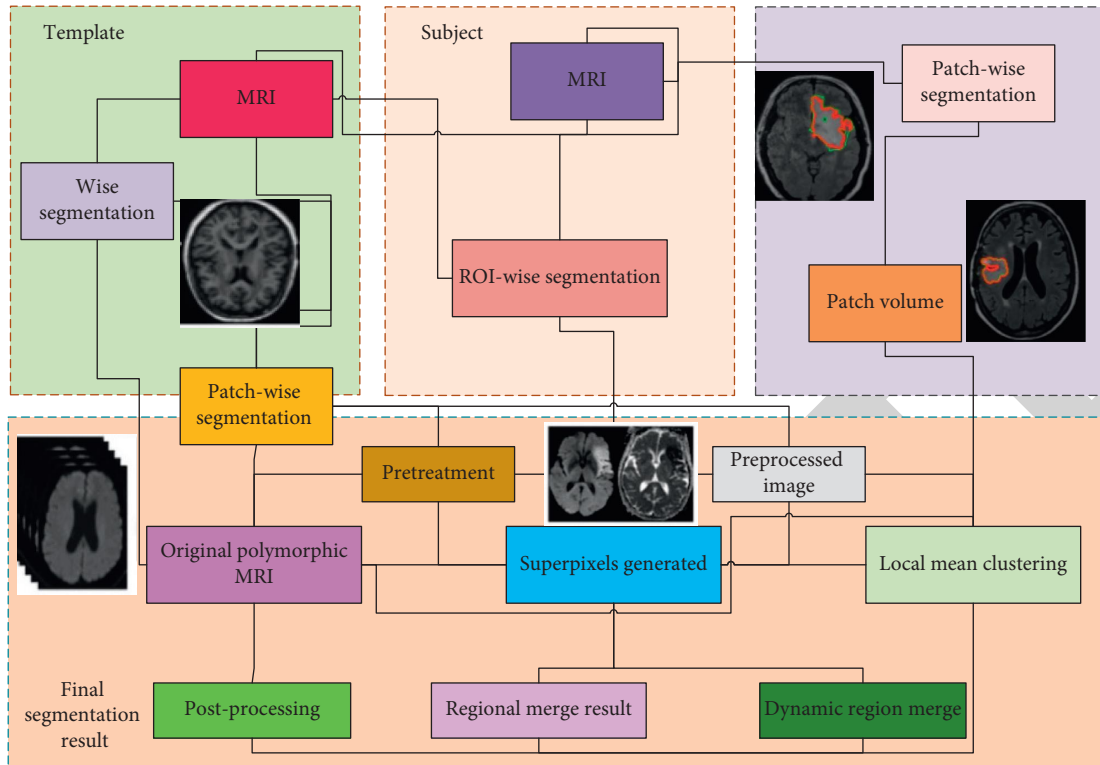
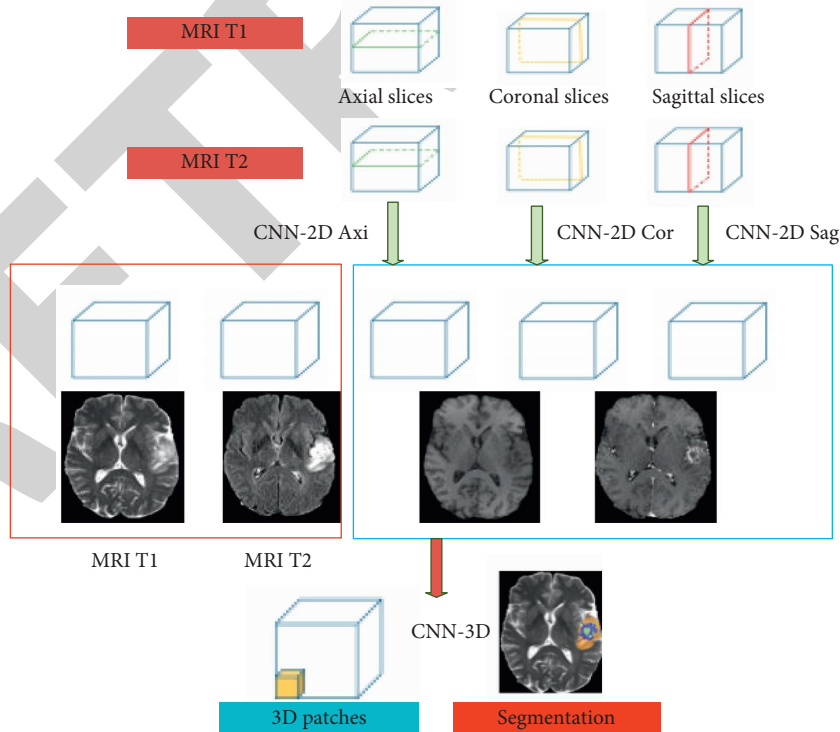FIGURE 1: Algorithm framework.



FIGURE 2: Multimodal 3D-CNNs.

the $C_1$ layer. The images are downsampled by 2D average to obtain the $S_2$ layer; after all the features of the $S_2$ layer are summed, $123 \times 3 \times 2$ convolution templates are used to obtain 12 feature maps of $6 \times 6 \times 2$ $C_3$ layer; $C_3$ layer is averaged down after sampling, the $S_4$ layer is obtained; the $S_4$ layer is normalized by column to obtain the 96-dimensional feature vector $F_5$.

The original input layer of multimodal 3D-CNNs is composed of four modalities. Through 3D convolution, the difference information between each modal is automatically extracted; the supervised learning method realizes the extraction of different classification features for different patient difference information; downsampling makes feature extraction contain more structural edge information while eliminating redundant information and noise; multimodal common input makes the original input require less neighborhood information to adapt to tumor points in different image layers and improve brain tumors clustering accuracy.

## 4. MRI Brain Tumor Subspace Clustering Algorithm Based on Multimodal 3D-CNNs Feature Extraction

In terms of feature extraction, although multimodal 3D-CNNs can extract the different information between the modalities that are more conducive to classification, deep learning will cause a partial loss of the original input information, and Haar wavelet transform is a simple and effective signal. The processing method is the preferred feature extraction method in pixel-based MRI brain tumor clustering [14]. Therefore, while acquiring the features of multimodal 3D-CNNs, reference [15] uses the 3D neighborhood gray information, neighborhood mean, standard deviation, Haar wavelet low-frequency coefficients, and multimodal 3D CNNs of each modal MRI image.

MRI brain tumor clustering based on multimodal 3D-CNNs feature extraction is divided into image preprocessing, feature extraction, feature selection, classifier training, image clustering, and other processes. The specific clustering process in this paper is shown in Figure 3.

The features together constitute the initial features of the clustering method in this paper, in which the 3D neighborhood is $5 \times 5 \times 5$, and the principal component analysis (PCA) method is used to select the features of the initial feature set to achieve the purpose of dimensionality reduction and elimination of redundant information [16]. A support vector machine (SVM) based on a radial basis kernel function is selected as the pixel classifier. In the training of the classifier, a layer of the patient with a tumor image is randomly selected, and 60 points inside and outside the tumor are taken as training samples [17].

In the training of multimodal 3D-CNNs parameters, a layer of images containing tumor layers is randomly selected, and all tumor points and the same number of background points are selected as training samples. Due to the different tumor positions of different patients, the neighborhood information around the tumor is not the same, so for different patients, the optimal original input layer neighborhood size is determined adaptively through the grid optimization method; after determining the neighborhood size, through multiple learning of training samples, obtain the final multimodal 3D-CNNs of each layer convolution weight and bias parameters [18].

## 5. Experimental Results and Analysis

In this section of the experiment, we first determine the value range of the original input layer neighborhood of multimodal 3D-CNNs through experiments; then, we use the method of adding multimodal 3D-CNNs features to realize the clustering of brain tumor MRI images and analyze the differences shown by different patients compared with the method that does not add the features of multimodal 3D CNNs one by one; finally, we compare the method based on the features of multimodal 3D-CNNs and the method based on the features of multimodal 2D-CNNs. In order to verify the effectiveness and necessity of the method in this paper, the dice coefficient (dice similarity coefficient), sensitivity and false positive rate (FP), and other technical indicators are used to evaluate the clustering results, where the dice coefficient represents the experimental clustering. The result is similar to the result of manual clustering by experts. Sensitivity indicates the proportion of tumor points with correct clustering.

*5.1. Parameter Range Determination.* The original input layer of multimodal 3D-CNNs needs two convolutions and two downsamplings to get the initial features, so the original input layer size must be $(10 + 4n) \times (10 + 4n) \times 4$, where $n$ is a natural number. Figure 4 shows the average clustering results of the same patient training layer with different neighborhood sizes. It can be seen from the figure that the optimal neighborhood size appears between 14 and 26, and considering the clustering time and the clustering accuracy of small tumors, the neighborhood value should not be too large, so the neighborhood optimization range is 10–30 [17].

*5.2. Analysis of the Results of Different Patients.* Figure 5 shows the clustering results of each training layer of 10 patients. It can be seen from the figure that the clustering accuracy of patients 2, 3, 4, 5, and 6 plus multimodal 3D-CNNs features have no significant change [18]. This is since the neighborhood grayscale and Haar wavelet low-frequency coefficients can already describe the characteristics of each pixel well, for patients 1, 7, 8, 9, 10; after adding the 3D-CNNs features, the clustering accuracy of the training layer is obtained. Significant improvement is due to modal 3D-CNNs adaptively extracting features that are conducive to classification. Experiments show that even in the training layer, the applicability of subjective feature extraction is not high.

Figure 6 shows the different coefficients of the clustering results of 7 patients. Table 1 shows the average clustering results of the 7 patients.

From Figure 6 and Table 1, we can see that for these 7 patients, after applying the features of multimodal 3D-CNNs, the dice coefficients of 7 patients have been improved to varying degrees, which is mainly reflected in the sensitivity coefficient, that is, the false negative rate has been significantly improved, and the average dice coefficient has

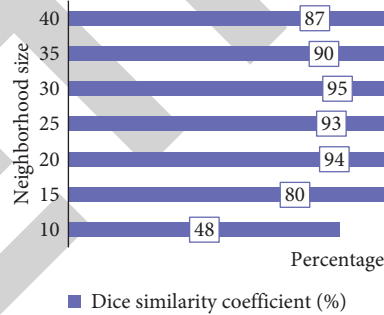Figure 3: 3D-CNNs feature extraction method of MRI brain tumor subspace clustering algorithm.



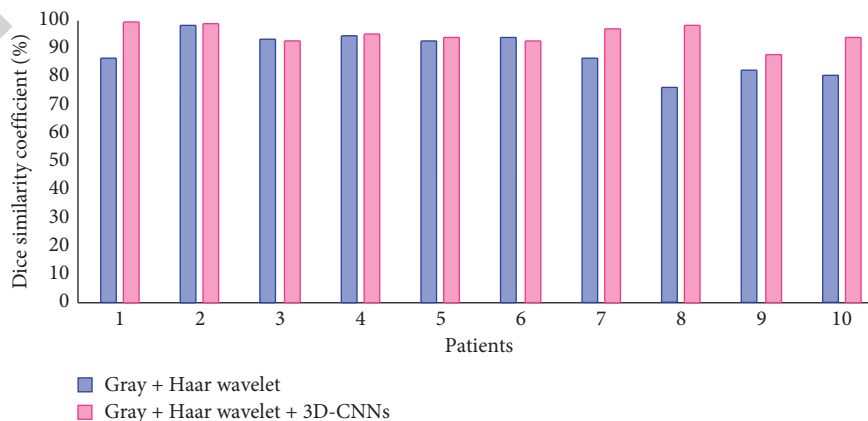Figure 4: Multimodal 3D-CNNs and input neighborhood size.



Figure 5: Comparison of clustering results of different feature training layers.
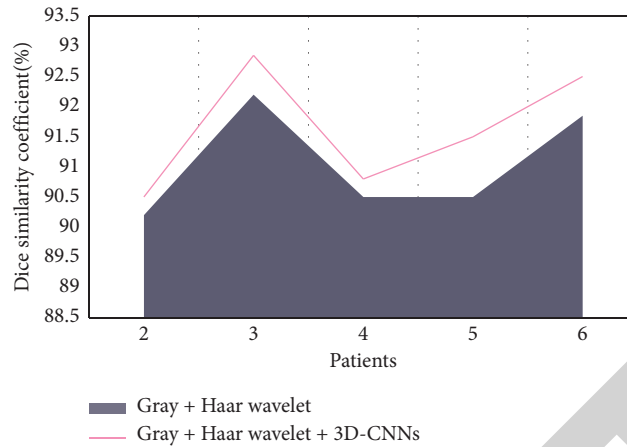
Figure 6: Comparison of clustering results of different feature test layers.

Table 1: Average clustering results of 7 patients.

| Feature type | Dice coefficient (%) | Sensitivity (%) | False negative rate (%) |
|---|---|---|---|
| Gray + Haar wavelet | 83.11 ± 1. 64 | 86. 37 ± 2. 10 | 18. 55 ± 10. 32 |
| Gray + Haar wavelet + 3D-CNNs | 88. 52 ± 2. 21 | 92. 58 ± 3. 74 | 18.3 9 ± 6. 21 |

increased from 83.11 to 88.52. The experimental results prove that the supervised feature extraction method of multimodal 3D-CNNs can obtain more boundary information from the original features of the large neighborhood, and the convolution template obtained through machine learning can obtain more boundary information [19].

The feature information is more conducive to classification, and at the same time, downsampling removes part of the redundant information so that the number of features is not too large, and the final clustering accuracy is greatly improved. However, not all patients will increase the clustering accuracy after adding the multimodal 3D-CNNs features. As shown in Figure 7, for patients 7, 8, and 9, after adding new features, the dice coefficient ignorant significant changes, this is because, for these three patients, the edema area around the tumor is small, the gray and texture characteristics of the tumor area and the tumor area are obvious, and the neighborhood gray plus Haar wavelet coefficient can already distinguish tumor spots from nontumor spots [20].

5.3. Comparison of Superpixel Generation Methods with Local k-Means Clustering. Figure 8 shows the multimodal MRI diagram of patient 5 and patient 6. It can be seen from the figure that for patient 5, the tumor is very poorly distinguished from peripheral edema, and $T_1$ mode basically cannot provide any grayscale, and texture information is used for classification. T1C mode has rich texture information in the center of the tumor, but it is difficult for $T_1C$ and $T_2$ modes to distinguish the boundaries of tumor and nontumor regions. From the training layer to the test layer, the clustering accuracy of patient 5 is not ideal. For patient 6, the gray level information of FLAIR and $T_1$ modalities, the texture information of $T_1C$ and $T_2$ modalities can distinguish tumor spots from nontumor spots well, and there is
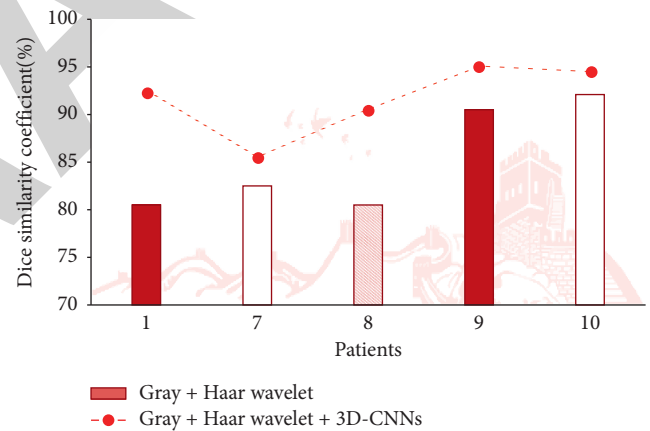


Figure 7: Clustering results of different feature test layers.

less edema around the tumor, so only adjacent domain grayscale and Haar wavelet low-frequency coefficients can achieve high clustering accuracy, but the clustering progress of patient 6 has not improved. Figure 8 shows the clustering results of the training layer of patient 5. It can be seen from the true value that the boundary between tumor and nontumor is very fuzzy. For the neighborhood grayscale and Haar wavelet features, the clustering results effectively divided the surrounding large areas of edema. After removing the tumor and adding multimodal 3D-CNNs features, this situation has been significantly improved.

Table 2 shows the average clustering results of 10 patients using neighborhood grayscale and Haar wavelet low-frequency coefficients (basic features), basic features plus multimodal 3D-CNNs, and basic feature classic 2D-CNNs. Among them, classic 2D-CNNs use 4 CNNs with different neighborhood sizes for feature learning for 4 modalities, and the neighborhood size is obtained within 20–50 through grid
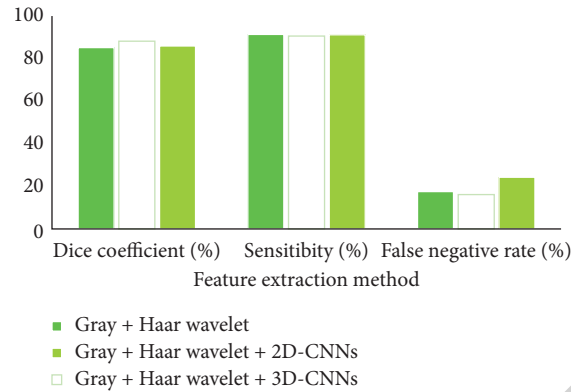
FIGURE 8: Clustering results of patient 5.

TABLE 2: Average clustering results of 10 patients.

| Feature extraction method | Dice coefficient (%) | Sensitivity (%) | False negative rate (%) |
| --- | --- | --- | --- |
| Gray + Haar wavelet | 85.17 ± 1. 64 | 91. 36 ± 1. 81 | 17. 55 ± 8. 17 |
| Gray + Haar wavelet + 3D-CNNs | 88. 17 ± 1. 21 | 90. 58 ± 2. 77 | 16.14 ± 4. 28 |
| Gray + Haar wavelet + 2D-CNNs | 85. 93 ± 3. 58 | 91. 18 ± 3. 48 | 24. 28 ± 1. 84 |

optimization. The results in Table 2 show that the feature extraction method of multimodal 3D-CNNs is significantly better than classic 2D-CNNs. With the features of multimodal 3D-CNNs, the dice coefficient becomes 88.17%, which is significantly better than classic 2D-CNNs. And after adding the features of classic 2D-CNNs, the clustering results are worse than using only the neighborhood grayscale and Haar wavelet low-frequency coefficients. This is because, first of all, brain tumors are generally close to spherical, and the size of tumors in different layers of the same patient is different. The 2D-CNNs model trained on the training layer is difficult to adapt to the entire patient's tumor layer image; secondly, the 4 models that are separately performing 2D-CNNs feature extraction on the state can theoretically obtain richer difference information between different modalities, but too much feature information increases the degree of linear indivisibility of each pixel, making the clustering result worse; modal 3D-CNNs not only overcome the shortcomings of the above classic 2D-CNNs, the three-dimensional combination of the four modalities is more conducive to the combination of the difference information between the modalities while removing redundant information to promote the realization of effective classification.

## 6. Conclusion

A clustering method of MRI brain tumors based on the feature extraction of multimodal 3D-CNNs is proposed. There are many feature extraction methods in image clustering, but because all feature extraction methods are preset based on subjective experience, they are not suitable for the specificity of brain tumor size, shape, and grayscale. CNNs is a supervised feature extraction method for specific learning of classified objects. It has been successfully applied in many fields. However, for image clustering, especially multimodal

MRI brain tumor clustering, conventional 2D-CNNs, it cannot achieve the purpose of feature extraction and high-precision clustering. In response to the above, combined with the characteristics of multimodal MRI images, this paper proposes a multimodal 3D-CNNs feature extraction method, which can make full use of the different information of each modal while taking into account the difference in tumor size and extract more abundant neighbors. Domain information and boundary information can better distinguish tumor points with fuzzy boundaries from nontumor points. Experimental results show that this method can adapt to multivariable and multimodal MRI brain tumor images and accurately cluster brain tumors. In the following research, we will further analyze how to better combine multimodal and comodal 3D neighborhood CNNs to make full use of 3D MRI image information upon feature extraction and how to more effectively improve the clustering speed in order to further improve the clustering strategy proposed in this article.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no known conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

## References

[1] J. Wu, V. S. Sheng, J. Zhang et al., "Multi-label active learning algorithms for image classification: overview and future promise," *ACM Computing Surveys (CSUR)*, vol. 53, no. 2, pp. 1–35, 2020.

[2] C. Q. Xia, K. Han, Y. Qi, Y. Zhang, and D. J. Yu, "A self-training subspace clustering algorithm under low-rank representation for cancer classification on gene expression data," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 4, pp. 1315–1324, 2017.

[3] M. G. Parsa, H. Zare, and M. Ghatee, "Unsupervised feature selection based on adaptive similarity learning and subspace clustering," *Engineering Applications of Artificial Intelligence*, vol. 9, no. 5, pp. 103855-103856, 2020.

[4] C. Dupont, N. Betrouni, N. Reyns, and M. Vermandel, "On image segmentation methods applied to glioblastoma: state of art and new trends," *IRBM*, vol. 37, no. 3, pp. 131–143, 2016.

[5] Y. Yang, L. F. Yan, X. Zhang et al., "Optimizing texture retrieving model for multimodal MR image-based support vector machine for classifying glioma," *Journal of Magnetic Resonance Imaging*, vol. 49, no. 5, pp. 1263–1274, 2019.

[6] Y. Guo, H. Li, M. Cai, and L. Li, "Integrative subspace clustering by common and specific decomposition for applications on cancer subtype identification," *BMC Medical Genomics*, vol. 12, no. 9, p. 191, 2019.

[7] H. Luo, Q. Zhuang, Y. Wang et al., "A novel image signature-based radiomics method to achieve precise diagnosis and prognostic stratification of gliomas," *Laboratory Investigation*, vol. 2, no. 2, pp. 1–3, 2020.

[8] T. Wu, Y. Xiao, M. Guo, and F. Nie, "A general framework for dimensionality reduction of K-means clustering," *Journal of Classification*, vol. 37, no. 3, pp. 616–631, 2020.

[9] J. Liu, Y. Cheng, X. Wang, X. Cui, Y. Kong, and J. Du, "Low rank subspace clustering via discrete constraint and hypergraph regularization for tumor molecular pattern discovery," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 15, no. 5, pp. 1500–1512, 2018.

[10] A. Saha, S. Banerjee, S. Kurtek et al., "DEMARCATE: density-based magnetic resonance image clustering for assessing tumor heterogeneity in cancer," *NeuroImage: Clinical*, vol. 12, no. 1, pp. 132–143, 2016.

[11] M. Latha and R. Surya, "Brain tumour detection using neural network classifier and k-means clustering algorithm for classification and segmentation," *European Journal of Applied Sciences*, vol. 9, no. 2, pp. 66–71, 2017.

[12] T. Zhan, F. Shen, X. Hong et al., "A glioma segmentation method using cotraining and superpixel-based spatial and clinical constraints," *IEEE Access*, vol. 6, no. 1, pp. 57113–57122, 2018.

[13] G. Mirzaei and H. Adeli, "Segmentation and clustering in brain MRI imaging," *Reviews in the Neurosciences*, vol. 30, no. 1, pp. 31–44, 2018.

[14] A. Bi, W. Ying, and Z. Qian, "Spatial fuzzy clustering and its application for MRI and CT image segmentation," *Journal of Medical Imaging and Health Informatics*, vol. 11, no. 2, pp. 409–412, 2020.

[15] D. Komura and S. Ishikawa, "Machine learning methods for histopathological image analysis," *Computational and Structural Biotechnology Journal*, vol. 1, no. 6, pp. 34–42, 2018.

[16] S. K. Rao and B. Lingappa, "Image analysis for MRI based brain tumour detection using hybrid segmentation and deep learning classification technique," *International Journal of Intelligent Engineering and Systems*, vol. 1, no. 2, pp. 53–62, 2019.

[17] Z. A. Al-Saffar and T. Yildirim, "A novel approach to improving brain image classification using mutual information-accelerated singular value decomposition," *IEEE Access*, vol. 1, no. 8, pp. 52575–52587, 2020.

[18] G. Litjens, T. Kooi, B. E. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 4, no. 2, pp. 60–88, 2017.

[19] X. Zeng, A. Chen, and M. Zhou, "Color perception algorithm of medical images using density peak based hierarchical clustering," *Biomedical Signal Processing and Control*, vol. 4, no. 8, pp. 69–79, 2019.

[20] Y. Zhu, K. M. Ting, and M. J. Carman, "Grouping points by shared subspaces for effective subspace clustering," *Pattern Recognition*, vol. 8, no. 3, pp. 230–244, 2018.