IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

SPECIAL SECTION ON DATA-ENABLED INTELLIGENCE FOR DIGITAL HEALTH

Received October 4, 2019, accepted October 12, 2019, date of publication October 16, 2019, date of current version October 31, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2947689*

# Epileptic EEG Detection Using a Multi-View Fuzzy Clustering Algorithm with Multi-Medoid

**QIANYI ZHAN[1,2], YIZHANG JIANG[1,2,3], (Member, IEEE), KAIJIAN XIA[2,4], JING XUE[5], WEI HU[5], HUANGXING LIN[6], AND YUAN LIU[1,2]**

[1]School of Digital Media, Jiangnan University, Wuxi 214122, China
[2]Jiangsu Key Laboratory of Media Design and Software Technology, Wuxi 214122, China
[3]Jiangsu Jiangda Smart Science and Technology Ltd., Jingjiang 214500, China
[4]Changshu No.1 People's Hospital, Changshu 215500, China
[5]Affiliated Wuxi People's Hospital of Nanjing Medical University, Wuxi 214023, China
[6]Fujian Key Laboratory of Sensing and Computing for Smart City, School of Informatics, Xiamen University, Xiamen 361005, China

Corresponding authors: Yizhang Jiang (yzjiang@jiangnan.edu.cn) and Yuan Liu (lyuan1800@sina.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61702225 and Grant 61902152, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20180600, in part by the 2018 Six Talent Peaks Project of Jiangsu Province under Grant XYDXX-127, and in part by the Science and Technology Demonstration Project of Social Development of Wuxi under Grant WX18IVJN002.

**ABSTRACT** Using clustering algorithms to automatically analyze EEGs of patients and to identify the characteristic waves of epilepsy is of high clinical value. Traditional clustering algorithms mostly use a calculated virtual single representative medoid point to describe the cluster structure, but this single representative medoid point has insufficient information. To accurately capture more accurate intracluster structural information, a representative multi-medoid points strategy is adopted, which describes the cluster structure by assigning representative weights to each sample in the cluster. Considering that the multi-view learning mechanism combines information from each view to improve the algorithm's clustering performance, a multi-view fuzzy clustering algorithm with multi-medoid (MvFMMdd) is proposed. This algorithm discards the approach of the traditional fuzzy clustering algorithm, which uses a single virtual representative point to characterize the cluster structure, and uses several real representative points to describe the cluster structure. Experiments verify the medical significance of the proposed algorithm.

**INDEX TERMS** Epileptic EEG, multi-view, multi-medoid, fuzzy clustering.

## I. INTRODUCTION

Epilepsy, a disease of the brain, causes dysfunction in consciousness, sensation, movement and mentality, causing substantial pain and serious physical and mental damage to the patient [1]. Seizures are the result of sudden abnormal synchronous discharges in neuron groups in the brain [2]. EEG is an important method used in the characterization of the electrophysiological activity of brain tissue and the functional state of the brain. Therefore, the analysis of EEGs of patients with epilepsy has high value for the diagnosis of epilepsy and the judgment of the type of seizure. However, understanding EEG signals is still based on the subjective judgment of the doctor through visual inspection. When relying on the artificial interpretation of EEG signals, problems such as subjectivity and inconsistent inspection standards consume a high level of manpower and introduce a high error rate. Therefore, it is necessary to use current computer technology to automatically analyze patients' EEGs.

The associate editor coordinating the review of this manuscript and approving it for publication was Yongtao Hao.

At present, the algorithms used for the detection of epilepsy mainly include artificial neural networks [3]–[5]; time-frequency analysis algorithms [6], [7]; and fuzzy clustering [8], [9], migration clustering [10]–[13], multi-view clustering [14]–[16], multitasking clustering [17] and other types of clustering algorithms [18], [19]. Liu *et al.* [20] proposed an integrated radial basis neural network to analyze epilepsy EEG signals and improve the stability of the model. Asha *et al.* [21] fused neural networks and support vector machines and applied this combined method to multichannel EEG signals to automatically detect epilepsy. Juárez-Guerra *et al.* [22] used wavelet transforms and neural networks to detect epilepsy. Srinivasan *et al.* [23] introduced the concept of entropy into artificial neural networks to classify epilepsy EEG signals. Subasi [24] used a discrete wavelet transform for preprocessing and, then, neural networks for classification, which improved the performance of the classifier. Khan and Gotman [25] mainly study the law of fluctuations between different frequencies and backgrounds. This method can detect epilepsy automatically with higher sensitivity. Bhuiyan and Das [26] preprocessed the EEG signals and then used the neural network to classify them. The advantage of this algorithm is an increase in the speed of calculation. Hassanpour and Boashash [27] proposed a new time-frequency-based spike detection technique and evaluated it in neonatal EEG data with good results. The fuzzy clustering algorithm automatically diagnoses epilepsy by classifying normal and seizure EEG signals. Sucharitha *et al.* [28] applied the improved fuzzy c-means (FCM) algorithm to epilepsy detection. Kuwata *et al.* [29] proposed a self-organized additive fuzzy clustering method and applied the algorithm to brain EEG recognition. The segmentation results obtained by the algorithm were more adaptive. Harikumar and Vijayakumar [30] compared the classification performance of epilepsy EEG signals by k-means clustering and several other classic algorithms. All of the abovementioned methods use a calculated virtual single cluster center to describe the cluster structure, which is not a real sample. Therefore, drift occurs under the influence of noise and outliers, making this method unable to accurately capture the internal structure of the cluster. In addition, for medical images, the virtual center does not express organ tissue very well.

To address the abovementioned challenges, an MvFMMdd algorithm is given, and the algorithm is applied to epilepsy EEG signals. Some contributions made in the process of analyzing EEG signals in epilepsy mainly include the following.

(1) The use of a single representative medoid point in describing the cluster structure is prone to insufficient information. To accurately capture more accurate intra-cluster structural information, a cluster structure expression strategy based on multiple representative medoid points is introduced.
(2) A new multi-view clustering model comprehensively considers the quality of each view and assigns corresponding weights according to the quality of each view, which effectively improves the clustering performance.
(3) Experiments verify the effectiveness of the clustering performance of the MvFMMdd algorithm; the strategy of expressing the cluster structure with real representative multi-medoid points is more relevant for medical data.

## II. RELATED WORK

### A. MULTI-MEDOID REPRESENTATIVE POINT EXPRESSION STRATEGY

The multi-medoid representative point expression strategy treats all samples in the sample space as potential cluster centers and assigns a weighting factor to each sample to indicate the extent to which it represents the cluster. Therefore, the cluster structure in the sample space can be exemplified by the included samples, and more cluster structure information can be captured than by the cluster center being represented by a single representative medoid point. To make the strategy more general, it can be considered that each sample in the data set may become a representative point of the cluster, and this possibility can be quantized into a weight coefficient within the range [0,1]. The larger the weight value of the sample is, the more likely the sample is to represent its cluster. The schematic diagram of the strategy is as follows:

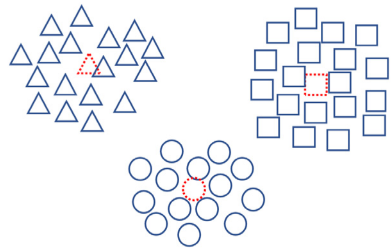


**FIGURE 1.** Sample graph of the virtual single-medoid representative point expression cluster structure.

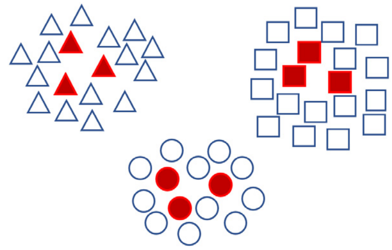


**FIGURE 2.** Sample graph of the real multi-medoid representative point expression cluster structure.

**FIGURE 1** shows a sample graph of the virtual single-medoid representative point expression cluster structure. **FIGURE 2** shows a sample graph of the real multi-medoid representative point expression cluster structure. The cluster center expressions shown in Figure 2 are all real sample points, which can better express the cluster structure.

### B. MULTI-VIEW LEARNING MECHANISM BASED ON VIEW FUZZY THEORY

For some data sets with complex structures or large numbers, the multi-view learning mechanism is introduced. Due to the increased complexity of the data, when people observe complex data sets, they can often interpret data through multiple views to obtain multi-view data. Multi-view data refers to the fact that some objects have multiple feature sets from different attribute spaces that are the result of synthesizing multiple types of features from different views. The main feature of the multi-view learning mechanism is that it analyzes data samples constructed from different features of the same object and uses the synergy between the various perspectives to find

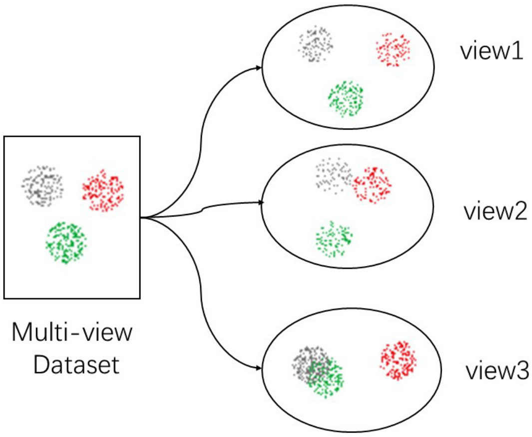



FIGURE 3. Schematic diagram of multi-view data.

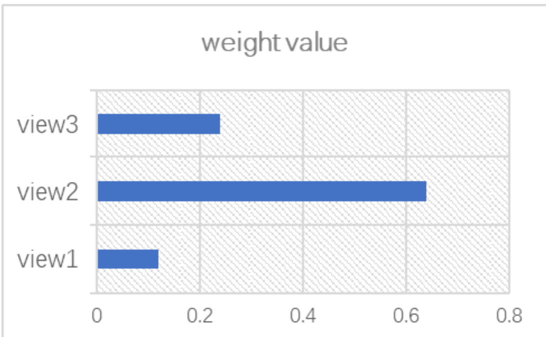



FIGURE 4. Schematic diagram of the view fuzzy mechanism.

the similar components between the perspectives to obtain a global decision result that tends to be consistent. Since the learning mechanism comprehensively considers the characteristic information of the research object from various views, the decision result that is obtained is more comprehensive and reliable than the traditional decision result based on the single view feature space. In the multi-view learning mechanism, the quality of each view is different; thus, each view makes a different contribution to the segmentation result. The problem description is as follows.

In **FIGURE 3**, the view 1 division quality is obviously better than that of views 2 and 3. The division quality of view 2 is better than that of view 1, and that of view 3 is the worst. To get a more satisfactory result, this paper introduces the view fuzzy weighting mechanism, which gives different weights to different views according to the quality of the division of each view. The principle being followed is that the data with a good-quality of view will have a large weight, and the data with a poor-quality view will have a small weight. For the multi-view data in Figure 3, the schematic diagram of the view fuzzy mechanism is as follows.

The process of the traditional clustering algorithm for processing multi-view data is as follows. First, traditional algorithms are used to process each view separately, and then, the integration strategy [31] is used to obtain the global result. Directly decomposing multi-view data into multiple single-view data sets for processing will have a bad influence on the final global segmentation result due to the obvious difference in the clustering results of the various views, resulting in the deterioration of the final global segmentation result and poor or unstable algorithm performance. The multi-view learning mechanism solves this problem very well. Many scholars have begun to study related algorithms based on the multi-view learning mechanism. Yamanishi *et al.* [32] used a cooperative clustering algorithm to handle a multi-view scene. Zhou and Burges [33] extended traditional spectral clustering and proposed a multi-view spectrum clustering algorithm. Pedrycz [34] introduced a collaborative learning idea based on FCM and proposed a cooperative method. Cleuziou *et al.* [35] proposed a multi-view fuzzy c-means (MvFCM) algorithm based on the classic FCM. Chen *et al.* [36] upgraded the traditional k-means algorithm to a multi-view k-means algorithm.

## III. MULTI-VIEW FUZZY CLUSTERING ALGORITHM WITH MULTI-MEDOID

Considering the multi-view learning mechanism mentioned above and the multi-medoid representative point expression strategy in related work, a multi-view fuzzy clustering algorithm with multi-medoid (MvFMMdd) is given. Given a sample set $X = \{x, i = 1, 2, \ldots, N\}$, $U$ represents the $X'$ fuzzy partition, and the goal of $V$ is to describe the internal structure of each cluster. The objective function of MvFMMdd is

$$J_{MvFMMdd}(U, V) = \sum_{r=1}^{R} w_r^{m_1} \sum_{c=1}^{K} \sum_{i=1}^{N} \sum_{j=1}^{N} u_{ci,r}^{m_2} v_{cj,r}^{n} d_{ij,r} \tag{1}$$

$$\sum_{c=1}^{K} u_{ci,r} = 1, u_{ci,r} \geq 0 \tag{2}$$

$$\sum_{j=1}^{N} v_{cj,r} = 1, v_{cj,r} \geq 0 \tag{3}$$

$$\sum_{r=1}^{R} w_r = 1, w_r \geq 0 \tag{4}$$

where $R$ is the number of views, $w_r$ is the weight of view $r$, $m_1$ and $m_2$ are fuzzier, $K$ is the cluster number, $N$ is the total number of samples, $u_{ci,r}$ represents the sample $x_i$'s membership to cluster c in view $r$, $v_{cj,r}$ is the sample $x_j$'s prototype weight in cluster $c$ in view$r$, the parameter $n$ is used to adjust the smoothness of the distribution of the prototype weights between all the objects in each cluster, $D$ is the dissimilarity matrix, and $d_{ij,r}$ is the dissimilarity between samples $x_i$ and $x_j$. $U_{K\times N}$ is a membership matrix belonging to $K$ clusters for $N$ samples, $V_{K\times N}$ represents the prototype weights of $N$ samples for $K$ clusters. The Lagrangian multiplier method is used to solve Eq. (1), and the variable expressions are as follows.

$$u_{ci,r}=\frac{\left(\sum_{j=1}^{N} v_{cj,r}^{n} d_{ij,r}\right)^{\frac{-1}{m-1}}}{\sum_{h=1}^{K}\left(\sum_{j=1}^{N} v_{hj,r}^{n} d_{ij,r}\right)^{\frac{-1}{m-1}}} \tag{5}$$

$$v_{cj,r}=\frac{\left(\sum_{i=1}^{N} u_{ci,r}^{m} d_{ij,r}\right)^{\frac{-1}{n-1}}}{\sum_{l=1}^{N}\left(\sum_{i=1}^{N} u_{ci,r}^{m} d_{il,r}\right)^{\frac{-1}{n-1}}} \tag{6}$$

$$w_{r}=\frac{\left(\sum_{c=1}^{K}\sum_{i=1}^{N}\sum_{j=}^{N} u_{ci,r}^{m_2} v_{cj,r}^{n} d_{ij,r}\right)^{\frac{-1}{m_1-1}}}{\sum_{g=1}^{R}\left(\sum_{c=1}^{K}\sum_{i=1}^{N}\sum_{j=}^{N} u_{ci,g}^{m_2} v_{cj,g}^{n} d_{ij,g}\right)^{\frac{-1}{m_1-1}}} \tag{7}$$

For easy understanding, set

$$\begin{aligned} v_{c,r}^{n} &= \left(v_{c1,r}^{n}, v_{c2,r}^{n}, \ldots, v_{cN,r}^{n}\right)^{T},\\ d_{i,r} &= \left(d_{i1,r}, d_{i2,r}, \ldots . d_{iN,r}\right)^{T},\\ z_{c,r} &= v_{c,r}^{n} \end{aligned} \tag{8}$$

$$a\left(x_{i,r}, z_{c,r}\right)=\sum_{j=1}^{N} z_{cj,r} d_{ij,r}=z_{c,r}^{T} d_{i,r} \tag{9}$$

Substituting Eq. (8) and (9) into Eq. (5), the expression of $u_{ci,r}$ is as follows.

$$u_{ci,r}=\frac{a\left(x_{i,r}, z_{c,r}\right)^{-1/(m_2-1)}}{\sum_{f=1}^{K} a\left(x_{i,r}, z_{f,r}\right)^{-1/(m_2-1)}} \tag{10}$$

The integration strategy for the global division result is as follows

$$U=\sum_{r=1}^{R} w_r u_{ci,r} \tag{11}$$

where $z_{c,r}$ is used to characterize the representative vector of each sample in the $c^{th}$ cluster in the view $r$. This vector can be seen as the "center vector" of the $c^{th}$ cluster, and each dimension in the vector represents the weight of each representative sample. $z_{c,r}$ is used to characterize the $c^{th}$ cluster, and $\left(x_{i,r}, z_{c,r}\right)$ represents the dissimilarity between $x_{i,r}$ and cluster $c$. In Eq.(10), the membership degree of sample $x_{i,r}$ to the $c^{th}$ cluster is related to the dissimilarity between the sample and the cluster. $u_{ci,r}$ is proportional to the dissimilarity between $x_{i,r}$ and cluster $c$ and is inversely proportional to the sum of the dissimilarities between sample $x_{i,r}$and other clusters. When a "center vector" of all the clusters is obtained, each object can be characterized by the distance between the object and all clusters.

Set

$$\begin{aligned} u_{c,r}^{m_2} &= \left(u_{c1,r}^{m_2}, u_{c2,r}^{m_2}, \ldots, u_{c,r}^{m_2}\right)^{T}\\ y_{c,r} &= u_{c,r}^{m_2} \end{aligned} \tag{12}$$

$$b\left(x_{j,r}, y_{c,r}\right)=\sum_{i=1}^{N} y_{ci,r} d_{ij,r}=y_{c,r}^{T} d_{j,r} \tag{13}$$

Substituting (12) and (13) into (8), the transformed expression of $v_{cj,r}$ is as follows

$$v_{cj,r}=\frac{b\left(x_{j,r}, y_{c,r}\right)^{\frac{-1}{n-1}}}{\sum_{h=1}^{N} b\left(x_{h,r}, y_{c,r}\right)^{\frac{-1}{n-1}}} \tag{14}$$

where $y_{c,r}$ is used to indicate the membership of all samples in view $r$ to the $c^{th}$ cluster. $b\left(x_{j,r}, y_{c,r}\right)$ represents the dissimilarity between the sample $x_{j,r}$in view $r$ and other objects in cluster $c$. The size of $b\left(x_{j,r}, y_{c,r}\right)$ characterizes the typical extent of sample $x_{j,r}$ in the $c^{th}$ cluster. In (14), $v_{cj,r}$ is $b\left(x_{j,r}, y_{c,r}\right)$ normalized over all the objects in the same cluster in view $r$, and the value of $v_{cj,r}$ does not depend on other clusters. Therefore, we call $\boldsymbol{V}$ a representative measure inside the cluster.

From the above series of formula derivation and analysis, $x_{j,r}$ in view $r$ is ranked by comparing the centrality of $x_{j,r}$ with other samples in the same cluster, and cluster c in each view is characterized by the representative weight of all objects in that cluster. After each cluster is characterized, each object's membership to $K$ clusters is assigned by the relative distances between this object and each of the clusters. The proposed MvFMMdd algorithm is described as follows.

| **Algorithm:**MvFMMdd | |
|---|---|
| **Input** | Given view number $R$, cluster number $K$, sample number $N$, iteration number $T_{max}$, iterative precision, fuzzy coefficients $m_1$ and $m_2$, and smoothing parameters $n$, initialize the distance matrix $\boldsymbol{D}$ of each view sample. |
| **Output** | the global fuzzy partition matrix $\boldsymbol{U}$ |
| **1** | Set iteration number $T$ =0; |
| **2** | Repeat |
| **3** | $T = T+1$; |
| **4** | From $r$ =1:$R$ |
| **5** | Iteratively calculate $\boldsymbol{u}_{ci,r}$ by (5); |
| **6** | Iteratively calculate the $\boldsymbol{v}_{cj,r}$ by (6); |
| **7** | Iteratively calculate the $\boldsymbol{w}_{\mathbf{r}}$by (7); |
| **8** | End |
| **9** | Until$\left\|J^{T}-J^{T+1}\right\| \leq \varepsilon$ or $T = T_{max}$ |
| **10** | Compute the global fuzzy partition matrix based on Eq (11) |

**TABLE 1.** Parameter settings.

| Algorithm | Value of each parameter |
|---|---|
| FCM | Fuzzifier $m$ is from {0.5,1.0,…,2.5} |
| CoFKM<br>WV-Co-FCM | Fuzzifier $m$ is from {0.5,1.0,…,2.5}<br>parameter $\eta$ is from $[0, \frac{C-1}{C}]$ with steps of 0.01, where $C$ is the number of views |
| FCMdd | Fuzzifier $m$ is from {0.5,1.0,…,2.5} |
| MvFMMdd | Fuzzifiers $m_1, m_2$ are from {0.5,1.0,…,2.5}<br>Smoothing coefficient n is from {1.1,1.2,…,2.0} |

The parameters in the table are optimally set using the grid search.

## IV. SIMULATION EXPERIMENTS AND RESULTS ANALYSIS

### A. EXPERIMENTAL SETTINGS

To demonstrate the performance of the MvFMMdd algorithm in processing multi-view data, three types of contrast algorithms were selected. One is the FCM algorithm and the k-means algorithm for a single-view single-medoid representative point. The second is the CoFKM algorithm for a multi-view single-medoid representative point [35], TW-(k)-means algorithm [36] and WV-Co-FCM [16]. The third is the FCMdd algorithm with multiple views and multiple medoids representative points [16]. To objectively compare the performance of each algorithm, the evaluation indicators used here are normalized mutual information (NMI) and the rand index (RI). The value range of both indicators is [0, 1], and the value is proportional to the cluster performance. The specific calculation of the two indicators is referenced [13].

The parameter settings involved in the algorithm and the comparison algorithm are shown in Table 1.

The processor of the PC used in the experiment is an Intel Core i5-6200 CPU @ 2.30 GHz, the memory is 4.00 GB, and the system type is a 64-bit operating system. The simulation software version is MATLAB R2018a. The experimental results were obtained by running the program 15 times to obtain the mean value.

### B. EXPERIMENTAL DATA SET

The experimental data used in this article can be downloaded from http://epileplotlogiebonn,de/cms/front_content.php?idcat=193&lang=3&changelange=3. The data set contains 500 samples and is divided into 5 groups of 100 samples. Groups A and B are EEG signals collected from healthy people, and groups C-E are EEG signals collected from patients. A detailed description of the data set can be found on the database download site.

To construct the multi-view EEG data set required for the experiment, this paper first uses the feature extraction method to extract the features of the EEG signal and, then, combines the extracted features into multiple different combinations, with one combination being regarded as the data set. The view number is equal to the method number for feature extraction. The feature extraction approaches used in this paper include kernel principal component analysis (KPCA), wavelet packet decomposition (WPD) and short-time Fourier transform (STFT); thus, the total number of views is 3, and a total of 6 data sets are constructed. See Table 2 for the details of the constructed multi-view data sets.

**TABLE 2.** Constructed multi-view EEG data sets.

| Datasets | Groups | Views (KPCA, WPD, STFT) | Cluster | Number of samples | Dimension |
|---|---|---|---|---|---|
| S1 | A, C | 3 | 2 | 200 | 12 |
| S2 | C, E | 3 | 2 | 200 | 12 |
| S3 | A, D, E | 3 | 3 | 300 | 18 |
| S4 | A, B, C, D | 3 | 4 | 300 | 24 |
| S5 | B, C, D, E | 3 | 4 | 300 | 24 |
| S6 | A, B, C, D, E | 3 | 5 | 500 | 30 |

### C. EXPERIMENTAL RESULT

The results of the operation of each algorithm on the six multi-view EEG data sets are shown in Table 3. The data in bold show the best performance for the corresponding experimental data sets.

Observing Table 3, the following conclusions can be drawn.

(1) For the multi-view EEG data, it can be seen from the experimental data of S1 to S6 that the clustering performance of the single-view clustering FCM and K-means algorithms is worse than that of the multi-view clustering algorithms CoFKM, WV-Co-FCM, and TW-(k)- means, including our proposed MvFMMdd algorithm. This result fully demonstrates that the clustering performance of the multi-view clustering algorithm is indeed more suitable for multi-view data.

(2) The clustering performance of the FCMdd algorithm is not much different from that of FCM or K-means algorithms, indicating that the multi-medoid representative points strategy did not substantially improve the clustering performance. However, this strategy is very important for the expression of medical data.

(3) The MvFMMdd algorithm proposed in this article has the best clustering performance for S2 and S4, but the effect on other datasets was not optimal. However, the MvFMMdd algorithm's performance on the six data sets is better than that of the single-view FCMdd algorithm, which further demonstrates that the multi-view algorithm is more suitable for multi-view data. However, although the MvFMMdd algorithm proposed in this paper results in limited improvement in clustering performance, its performance is always better than that of the single-view clustering algorithm, and the cluster structure expression strategy based on multi-medoid representative points is more suitable for clustering medical data – not only because the cluster structure expressed by multiple medoids representative points is

**TABLE 3.** Comparison of the clustering performance of each algorithm on multi-view EEG data sets.

| Dataset | Index | FCM | *k*-means | Co-FKM | WV-Co-FCM | TW-(k)-means | FCMdd | MvFMMdd |
|---|---|---|---|---|---|---|---|---|
| S1 | NMI | 0.3212 | 0.2302 | 0.3789 | **0.3813** | 0.3762 | 0.3304 | 0.3588 |
| | RI | 0.7041 | 0.6987 | 0.7172 | **0.7177** | 0.6728 | 0.6983 | 0.7012 |
| S2 | NMI | 0.6092 | 0.6013 | 0.6128 | 0.6134 | 0.6085 | 0.6070 | **0.6142** |
| | RI | 0.5289 | 0.5208 | 0.8114 | 0.8119 | 0.8093 | 0.8024 | **0.8120** |
| S3 | NMI | 0.3288 | 0.2765 | 0.4210 | **0.4362** | 0.4116 | 0.4068 | 0.4182 |
| | RI | 0.6193 | 0.6104 | 0.6831 | **0.6987** | 0.6351 | 0.6289 | 0.6806 |
| S4 | NMI | 0.2569 | 0.2692 | 0.2702 | 0.2722 | 0.2673 | 0.2615 | **0.2726** |
| | RI | 0.6270 | 0.6710 | 0.6815 | 0.6831 | 0.6722 | 0.6648 | **0.6837** |
| S5 | NMI | 0.2583 | 0.2522 | 0.2697 | **0.2815** | 0.2559 | 0.2562 | 0.2536 |
| | RI | 0.6286 | 0.6423 | 0.6863 | **0.6870** | 0.6765 | 0.6311 | 0.6774 |
| S6 | NMI | 0.3221 | 0.2505 | 0.3811 | **0.3906** | 0.3786 | 0.3196 | 0.3764 |
| | RI | 0.7102 | 0.6998 | 0.7272 | **0.7369** | 0.6964 | 0.7079 | 0.7143 |

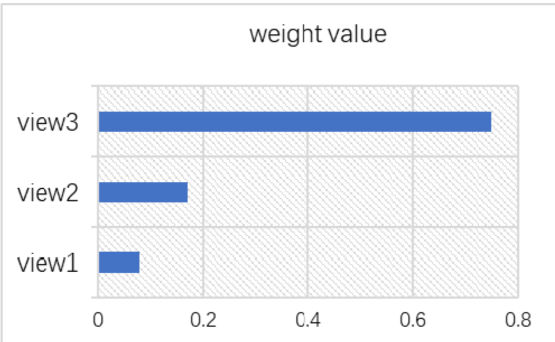



**FIGURE 5.** Three views' weights for S1 in the MvFMMdd algorithm.

richer and more accurate but also because the multiple medoids representative points are real data points, which is more medically meaningful.

Traditional multi-view clustering algorithms such as the Co-FKM algorithm, generally consider the importance of each view to be the same for clustering. However, in practical applications, there are often cases in which the clustering effect of some perspectives is not ideal. Then, the results obtained after the final decision will be subject to those views that have no obvious clustering characteristics, resulting in poor clustering performance in some cases. To solve this problem, the proposed MvFMMdd algorithm performs fuzzy weighting on each view and assigns corresponding weights according to the importance of each view. The influence of the determination of the view weight on the clustering performance has also been studied.

To verify the difference in the quality of each view, **FIGURE 5** shows the weighting of each view in the algorithm for the S1 data set. In Figure 5, each view contributes differently to the final partition. The weights of views 1 and 3 are small, indicating that the data overlap between the two views is high, and the weight of view 2 is the largest, indicating that the data in view 2 has good spatial division characteristics. The method used to assign corresponding weights according to the different degrees that each view contributes can substantially heighten the clustering performance and avoid interference from poor-quality samples in some views.

## V. CONCLUSION

In this study, a multi-view fuzzy clustering algorithm with multi-medoid is presented and used to analyze epilepsy EEG data. On the one hand, the introduction of a multi-view learning mechanism based on view fuzzy weighting considers the correlations and differences between the various view data; thus, the final result tends to be consistent. On the other hand, the view fuzzy weighting strategy that assigns corresponding weights according to the degree of contribution from each view substantially improves clustering performance and effectively avoids interference from the poor-quality data in some views. The introduction of the real multi-medoid representative points strategy can describe the cluster structure more richly and accurately. For the identification of medical data, real representative points can also be medically meaningful. The clustering performance of the MvFMMdd algorithm approaches that of the classic Co-FKM algorithm, and the clustering performance has not been substantially improved or decreased. However, the multi-medoid representative points strategy algorithm has high value for epileptic EEG data processing, since the multiple representative points are real data points; thus, the use of real points as the cluster center representative points is more medically meaningful for the processing of medical data.

Although the algorithm presented in this paper has a good clustering effect on the multi-view epilepsy EEG dataset, there is still room for further expansion in subsequent research. For example, it will be important to explore how to significantly improve clustering accuracy, avoid interference from noise points, and the like.

## REFERENCES

[1] B. S. Chang and D. H. Lowenstein, "Epilepsy," *New England J. Med.*, vol. 349, no. 13, pp. 1257–1266, Sep. 2003.

[2] R. S. Fisher, W. Van E. Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel, "Epileptic seizures and epilepsy: Definitions proposed by the international league against epilepsy (ILAE) and the international bureau for epilepsy (IBE)," *Epilepsia*, vol. 46, no. 4, pp. 470–472, Apr. 2005.

[3] J. Saini and M. Dutta, "Epilepsy classification using optimized artificial neural network," *Neurol. Res.*, vol. 40, no. 11, pp. 982–994, Aug. 2018.

[4] J. M. Haglin, G. Jimenez, and A. E. M. Eltorai, "Artificial neural networks in medicine," *Health Technol.*, vol. 9, no. 1, pp. 1–6, Jan. 2019.

[5] C. Huang, G. Tian, Y. Lan, Y. Hao, Y. Cheng, Y. Peng, and W. Che, "A new pulse coupled neural network (PCNN) for brain medical image fusion empowered by shuffled frog leaping," *Frontiers Neurosci.*, vol. 13, Mar. 2019, Art. no. 210.

[6] F. Lieb, H.-G. Stark, and C. Thielemann, "A stationary wavelet transform and a time-frequency based spike detection algorithm for extracellular recorded data," *J. Neural Eng.*, vol. 14, no. 3, Jun. 2017, Art. no. 036013.

[7] N. Richard, B. Laursen, M. Grupe, A. M. Drewes, C. Graversen, H. B. D. Sørensen, and J. F. Bastlund, "Adapted wavelet transform improves time-frequency representations: A study of auditory elicited P300-like event-related potentials in rats," *J. Neural Eng.*, vol. 14, no. 2, Feb. 2017, Art. no. 026012.

[8] P. Qian, Y. Jiang, Z. Deng, L. Hu, S. Sun, S. Wang, and R. F. Muzic, "Cluster prototypes and fuzzy memberships jointly leveraged cross-domain maximum entropy clustering," *IEEE Trans. Cybern.*, vol. 46, no. 1, pp. 181–193, Jan. 2016.

[9] P. Qian, Y. Jiang, S. Wang, K. H. Su, J. Wang, L. Hu, and R. F. Muzic, "Affinity and penalty jointly constrained spectral clustering with all-compatibility, flexibility, and robustness," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 5, pp. 1123–1138, May 2017.

[10] P. Qian, K. Zhao, Y. Jiang, K. H. Su, Z. Deng, S. Wang, and R. F. Muzic, "Knowledge-leveraged transfer fuzzy c-means for texture image segmentation with self-adaptive cluster prototype matching," *Knowl.-Based Syst.*, vol. 130, pp. 33–50, Aug. 2017.

[11] P. Qian, C. Xi, M. Xu, Y. Jiang, K. H. Su, S. Wang, and R. F. Muzic, "SSC-EKE: Semi-supervised classification with extensive knowledge exploitation," *Inf. Sci.*, vol. 422, pp. 51–76, Jan. 2018.

[12] P. Qian, S. Sun, Y. Jiang, K. H. Su, T. Ni, S. Wang, and R. F. Muzic, "Cross-domain, soft-partition clustering with diversity measure and knowledge reference," *Pattern Recognit.*, vol. 50, pp. 155–177, Feb. 2016.

[13] Y. Jiang, D. Wu, Z. Deng, P. Qian, J. Wang, G. Wang, F. L. Chung, K. S. Choi, and S. Wang, "Seizure classification from EEG signals using transfer learning, semi-supervised learning and TSK fuzzy system," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 12, pp. 2270–2284, Dec. 2017.

[14] P. Qian, J. Zhou, Y. Jiang, F. Liang, K. Zhao, S. Wang, K. Su, and R. Muzic, "Multi-view maximum entropy clustering by jointly leveraging inter-view collaborations and intra-view-weighted attributes," *IEEE Access*, vol. 6, pp. 28594–28610, 2018.

[15] Y. Jiang, Z. Deng, F.-L. Chung, G. Wang, P. Qian, K.-S. Choi, and S. Wang, "Recognition of epileptic EEG signals using a novel multiview TSK fuzzy system," *IEEE Trans. Fuzzy Syst.*, vol. 25, no. 1, pp. 3–20, Feb. 2017.

[16] Y. Jiang, F.-L. Chung, S. Wang, Z. Deng, J. Wang, and P. Qian, "Collaborative fuzzy clustering from multiple weighted views," *IEEE Trans. Cybern.*, vol. 45, no. 4, pp. 688–701, Apr. 2015.

[17] Y. Jiang, F.-L. Chung, H. Ishibuchi, Z. Deng, and S. Wang, "Multitask TSK fuzzy system modeling by mining intertask common hidden structure," *IEEE Trans. Cybern.*, vol. 45, no. 3, pp. 534–547, Mar. 2015.

[18] P. Qian, F.-L. Chung, S. Wang, and Z. Deng, "Fast graph-based relaxed clustering for large data sets using minimal enclosing ball," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 3, pp. 672–687, Jun. 2012.

[19] C. Huang, Y. Xie, Y. Lan, Y. Hao, F. Chen, Y. Cheng, and Y. Peng, "A new framework for the integrative analytics of intravascular ultrasound and optical coherence tomography images," *IEEE Access*, vol. 6, pp. 36408–36419, 2018.

[20] L. Liu, Y. Ji, Y. Gao, and L. Kuang, "Classification of epileptic electroencephalograms signal based on integrated radius-basis-function neural-network," *J. Med. Imag. Health Inform.*, vol. 8, no. 7, pp. 1462–1467, Sep. 2018.

[21] S. A. Asha, C. Sudalaimani, P. Devanand, T. E. Thomas, and S. Sudhamony, "Automated seizure detection from multichannel EEG signals using support vector machine and artificial neural networks," in *Proc. Int. Mutli-Conf. Autom., Comput., Commun., Control Compress. Sens.*, Mar. 2013, pp. 558–563.

[22] E. Juárez-Guerra, V. Alarcon-Aquino, and P. Gómez-Gil, "Epilepsy seizure detection in EEG signals using wavelet transforms and neural networks," in *New Trends in Networking, Computing, E-Learning, Systems Sciences, and Engineering*. Cham, Switzerland: Springer, 2015, pp. 261–269.

[23] V. Srinivasan, C. Eswaran, and N. Sriraam, "Approximate entropy-based epileptic EEG detection using artificial neural networks," *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 3, pp. 288–295, May 2007.

[24] A. Subasi, "Epileptic seizure detection using dynamic wavelet network," *Expert Syst. Appl.*, vol. 29, no. 2, pp. 343–355, Aug. 2005.

[25] Y. U. Khan and J. Gotman, "Wavelet based automatic seizure detection in intracerebral electroencephalogram," *Clin. Neurophysiol.*, vol. 114, no. 5, pp. 898–908, May 2003.

[26] M. I. H. Bhuiyan and A. B. Das, "A subband correlation-based method for the automatic detection of epilepsy and seizure in the dual tree complex wavelet transform domain," in *Proc. Conf. Biomed. Eng. Sci.*, Dec. 2014, pp. 811–816.

[27] H. Hassanpour and B. Boashash, "A time-frequency approach for EEG spike detection," *Iranica J. Energy Environ.*, vol. 2, no. 4, pp. 390–395, Oct. 2011.

[28] M. Sucharitha, D. Jackson, and M. D. Anand, "Brain image segmentation using adaptive mean shift based fuzzy C means clustering algorithm," *Procedia Eng.*, vol. 38, pp. 4037–4042, Dec. 2012.

[29] T. Kuwata, M. Sato-Ilic, and L. C. Jain, "A learning based self-organized additive fuzzy clustering method and its application for EEG data," *Int. J. Knowl. Based Intell. Eng. Syst.*, vol. 16, no. 1, pp. 69–78, Jan. 2012.

[30] D. R. Harikumar and T. Vijayakumar, "Comprehensive analysis of hierarchical aggregation functions decision trees, SVD, K-means clustering, PCA and rule based AI optimization in the classification of fuzzy based epilepsy risk levels from EEG signals," *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.*, vol. 5, pp. 260–268, Feb. 2013.

[31] S. Asur, D. Ucar, and S. Parthasarathy, "An ensemble framework for clustering protein–protein interaction networks," *Bioinformatics*, vol. 23, no. 13, pp. i29–i40, Jul. 2007.

[32] Y. Yamanishi, J.-P. Vert, and M. Kanehisa, "Protein network inference from multiple genomic data: A supervised approach," *Bioinformatics*, vol. 20, no. 1, pp. i363–i370, Aug. 2004.

[33] D. Zhou and C. J. C. Burges, "Spectral clustering and transductive learning with multiple views," in *Proc. 24th Int. Conf. Mach. Learn.*, Jun. 2007, pp. 1159–1166.

[34] W. Pedrycz, "Collaborative fuzzy clustering," *Pattern Recognit. Lett.*, vol. 23, no. 14, pp. 1675–1686, Dec. 2002.

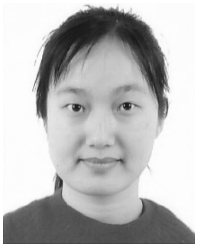
[35] G. Cleuziou, M. Exbrayat, L. Martin, and J.-H. Sublemontier, "CoFKM: A centralized method for multiple-view clustering," in *Proc. 9th IEEE Int. Conf. Data Mining*, Dec. 2009, pp. 752–757.

[36] X. Chen, X. Xu, J. Z. Huang, and Y. Ye, "TW-K-means: Automated two-level variable weighting clustering algorithm for multiview data," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 4, pp. 932–944, Apr. 2013.

**QIANYI ZHAN** received the B.S. degree in computer science from Nanjing Normal University, and the Ph.D. degree in computer science from Nanjing University. She is currently an Assistant Professor with the School of Digital Media, Jiangnan University. She has published more than ten refereed articles in journals and major conferences. Her principal research interests focus on data mining and social network analysis.

**YIZHANG JIANG** (M'12) received the Ph.D. degree from Jiangnan University, Wuxi, China, in 2015. He is currently an Associate Professor with the School of Digital Media, Jiangnan University. He has also been a Research Assistant with the Department of Computing, Hong Kong Polytechnic University, for two years. He has authored or coauthored more than 50 research articles in international/national journals, including the IEEE Transactions on Fuzzy Systems, the IEEE Transactions on Neural Networks and Learning Systems, the IEEE Transactions on Cybernetics, *ACM Transactions on Multimedia Computing, Communications, and Applications*, and *Information Sciences*. His research interests include pattern recognition, and intelligent computation and their applications. He has served as a reviewer or co-reviewer of several international conferences and journals, such as ICDM, TKDE, TFS, TNNLS, *Pattern Recognition*, *Neurocomputing*, and *Neural Computing & Applications*. He has also served as a Lead Guest Editor or Guest Editor for several international journals, such as *Journal of Ambient Intelligence and Humanized Computing* and *Journal of Cloud Computing*. He is currently an Associate Editor of IEEE Access, and is also an Editorial Review Board of the *International Journal of Healthcare Information Systems and Informatics*.

**KAIJIAN XIA** received the M.S. degree from Jiangnan University, Wuxi, China, in 2010. He is currently pursuing the Ph.D. degree with the China University of Mining and Technology. His research direction is medical information and medical image proceeding. He is currently with the Department of Computer, Changshu No. 1 People's Hospital. He has published several articles in international journals, including *Journal of medical systems* and *Journal of medical imaging and health informatics*.



**JING XUE** received the M.D. and M.M. degrees from the Department of Nephrology, Nanjing Medical University, Nanjing, China, in 2019. Her research interest includes machine learning and its applications on medicine. She is currently a Doctor with the Affiliated Wuxi People's Hospital of Nanjing Medical University, China.

**WEI HU** received the M.M. degree from Huashan Hospital affiliated to Fudan University, Shanghai, China. He is currently a Doctor with the Affiliated Wuxi People's Hospital of Nanjing Medical University, China. His principal research interest focuses on neuroimaging.

**HUANGXING LIN** received the B.S. degree in tracking technology and instrument from Beijing Jiaotong University, in 2015, and the M.S. degree in control theory and control engineering from Xiamen University, in 2018. He is currently pursuing the Ph.D. degree in signal and information processing with Xiamen University, Xiamen, China. His research interests include MRI reconstruction, medical image segmentation, and machine learning.

**YUAN LIU** is the principal investigator of many provincial research projects and has successfully completed them over the years. He has published more than 40 academic articles (including more than ten SCI and EI articles) and books in the authoritative and core journals. His main researches focus on the software development of network information systems, network security, and digital media applications. His current research interests include network traffic measurement, social networks, and digital media. He is a member of the 863 expert panel in the information security technology domain of the Ministry of Science and Technology, a senior member of the China Computer Federation (CCF), and a member of the CyberSecurity Association of China (CSAC). In 2013, he was also selected as a member of the 333 High level personnel training project and a member of the Digital-Media and Animation Education Steering Committee in the Ministry of Education, from 2013 to 2017. In 2014, he received the Grand Price in Technology Award from the China General Chamber of Commerce (CGCC). From 2009 to 2011, he has been named as the Excellent Educator in the Wuxi City and with the Jiangnan University. In 2014, due to his outstanding contribution, he was named as a Young and Middle-aged Expert in the Wuxi City. He is a Reviewer of the Journal of Chongging University of Posts and Telecommunications. He is also a member of Editorial Board of the *International Journal of Communication Networks and Information Security* (IJCNIS).