Multimedia Tools and Applications (2020) 79:9523–9542
https://doi.org/10.1007/s11042-019-07974-7

View-collaborative fuzzy soft subspace clustering for automatic medical image segmentation

Kaifa Zhao, et al. *[full author details at the end of the article]*

Received: 16 January 2019 / Revised: 17 June 2019 / Accepted: 10 July 2019 /
Published online: 24 July 2019

Abstract
With the rapid development of medical imaging methodologies, such as magnetic resonance (MR) and positron emission tomography (PET)/MR, various types of MR images, which are acquired using inconsistent MR pulse sequences on the same patient, have been applied in medical-image-based diagnoses. A feature map extracted from an MR image describes the patient's condition from one perspective. By effectively using all feature maps from various MR images, it is possible to completely describe the intrinsic characteristics of the patient's condition to facilitate a diagnosis. Facing such a scenario, classic machine learning algorithms typically stack these feature maps for unified processing and do not explore the importance of each feature within a single feature map or the relationships among feature maps. To address these challenges, both multiview and subspace learning scenarios are considered in this study, and the *multiview collaboration-based fuzzy soft subspace clustering* (MVC-FSSC) algorithm is proposed. The MVC-FSSC algorithm not only strives to exploit the agreement of decisions across all views via collaborative learning but also strives to utilize the soft subspace-based weighting mechanism to automatically evaluate the contribution of each dimensional feature to each estimated cluster within a single view. Our experimental results indicate that the proposed MVC-FSSC algorithm can effectively explore the collaborative relations among all views and the importance of features in their respective views. Additionally, our MVC-FSSC method has substantial advantages over traditional clustering algorithms in MR image segmentation. Applying the MVC-FSSC algorithm to five patients' MR images, the average mean absolute prediction deviation (MAPD) is 98.62 ± 8.34, which is significantly better than the score of 131.90 ± 16.03 that was obtained using the collaborative fuzzy k-means (CO-FKM) algorithm and the score of 128.87 ± 11.32 that was obtained using the quadratic weights and Gini-Simpson diversity-based fuzzy clustering (QWGSD-FC) algorithm.

Keywords Multiview clustering · Soft subspace clustering · Medical image segmentation · Collaborative learning

1 Introduction

Advanced state-of-the-art clinical diagnostic technologies are dynamic, precise, automatic and noninvasive and involve microquantization. Medical imaging technologies, such as computed tomography (CT), magnetic resonance (MR), positron emission tomography (PET), and their combinations that can effectively display the anatomical construction or molecular function of human body, have revolutionized and transformed clinical diagnostics and have become significant methodologies in modern medicine. During medical imaging, artificial intelligence-leveraged medical image processing technologies, which qualitatively and quantitatively analyze lesions and specified areas, have been playing important roles as well as facilitating the medical diagnosis progress. For instance, Wang, G. et al. [29] fine-tuned a convolutional neural network (CNN) to adapt it to medical image segmentation; Roth, H. R. et al. [25] applied the 3-dimensional fully convolutional network (3D U-Net) in CT image segmentation tasks and the experiments demonstrated a reasonably desirable performance in regions of interest. Among these artificial intelligence technologies, clustering [24], which belongs to unsupervised learning and aims to separate the data samples into several usually disjoint subsets, each of which is called a cluster, has been particularly applied to the medical image segmentation and recognition of regions of interest in recent years. For example, Su, K. H. et al. [27] used the classic fuzzy c-means (FCM) to segment the tissue types in MR brain images and generated the pseudo CTs, whereas Zhao, K. et al. [34] utilized the transfer fuzzy clustering together with neural network to distinguish the tissue types in abdomen-pelvis and thus facilitate the attenuation correction of PET/MR.

Although many clustering algorithms have been proposed in the past few decades [4, 7, 15, 16, 20, 23, 31], research on clustering is far from complete. A key challenge is that data are becoming increasingly complex. A prominent example is exactly medical images. During medical imaging, with different scan setup, we often obtain multiple imaging signals and thereby are able to reconstruct several medical images via one-time scanning. Taking the MR scan as an example, using several MR scan sequences, e.g., the free induction decay (FID), Dixon, and ultrashort echo time (UTE), we can obtain many MR images concurrently for one subject (patient). These images essentially, separately describe the same subject's body characteristics from inconsistent views. In such context, however, traditional clustering algorithms are usually not eligible to comprehensively utilize these features existing in different medical images to segment the tissue types well in target medical images. Therefore, multiview clustering technologies, which will be detailed in Section 2.1, have been applied with increasing frequency in the past few years. Even so, there is still the room of performance improvement with respect to multiview clustering methods. Most current multiview clustering methods cannot exploit information across views and are implemented on the whole feature sets, each of which could be constituted by primitively juxtaposing all of the characteristics from all available views. In addition, most existing multiview clustering methods regard each cluster within a view as coming from the same feature space. However, clusters in some views are usually associated with different subsets of characteristics and clusters may differ among subspaces in terms of their subsets of features [5, 14], especially in some high-dimensional multiview scenarios [22, 26]. Subspace clustering, which has been extensively studied, has recently been applied as an effective data mining method for high-dimensional datasets. The objective of subspace clustering is to partition different clusters into one or more subspaces of the original data dimensions. According to the methods used to identify subspaces, existing subspace clustering can be divided into two main categories: hard subspace clustering, which utilizes a binary value to represent the membership of a feature and subspaces are assigned to different clusters [1, 2, 5], and soft subspace clustering, which uses weights to evaluate the importance of different features in different clusters [9].

Motivated by these challenges, for the segmentation of the tissue types in the same subject's multiple medical images, a novel *multiview collaboration-based fuzzy soft subspace clustering*

(*MVC-FSSC* for short) algorithm is proposed in this paper to extensively mine and leverage the knowledge existing in the multiple, given medical images and thus to ultimately benefit the segmentation of tissue types.

By incorporating two dedicated weighting metrics regarding fuzzy weights, i.e., the *quadratic-weighted collaborative learning* (*QW-CoL*) and the *soft subspace weighted attributes* (*SSWA*), into the objective function, in MVC-FSSC the concepts of interview collaboration and fuzzy soft subspace clustering are fully embodied.

The remainder of the paper is organized as follows: Multiview learning strategies, related fuzzy clustering methods and soft subspace clustering are reviewed in Section 2. In Section 3, the framework as well as detailed algorithm of MVC-FSSC are presented. In Section 4, we present, analyze and discuss the experimental results. In Section 5, we conclude this paper.

2 Related work

2.1 Multiview clustering

In recent years, research on clustering multiple-view data has become prevalent in the field of artificial intelligence. When studying multiview clustering, we must consider how to utilize multiple data points and features in each view. The current multiview clustering algorithms are mainly divided into three strategies. The first strategy involves cascading (also called a priori fusion or feature fusion) [12, 19, 30], in which the views are cascaded into a unitary view, either by directly juxtaposing the sets of features or by indirectly combining the proximity metrics that are derived from each view. For instance, Wang, X. et al. [31] studied an improved, multifeature-fusion-based fuzzy c-means algorithm for color image segmentation in which the initial cluster centers were identified by incorporating the mutation operator of the genetic algorithm into particle swarm optimization. The second strategy is the distributed strategy (also called a posteriori fusion) [6, 11]. This strategy handles data from all views separately and uses result fusion techniques to combine clustering results across all perspectives. For example, Xue, Z. et al. [32] investigated the group-aware multiview fusion approach, in which images were divided into groups and assigned weights for evaluating the pairwise resemblance between groups. Iterative optimization is applied to improve the clustering results and to optimize the fusion weights. The third strategy is the collaborative strategy. During the clustering process, this strategy can exploit the connections across views. Jiang, Y. et al. [15] proposed multiple-weighted-view fuzzy clustering algorithms, which can automatically distinguish the importance of different views by the weight of each view. As discussed above, the collaborative strategy exploits mutual links between multiple views and satisfies the requirements of data confidentiality. Thus, we will study the third strategy in this paper.

2.2 Quadratic weights and Gini-Simpson diversity-based fuzzy clustering (QWGSD-FC)

Qian, P. et al. [23] proposed the QWGSD-FC algorithm, which inherits most of the advantages of classic fuzzy clustering models, namely, FCM [4], maximum-entropy clustering (MEC) [16] and fuzzy clustering by quadratic regularization (FC-QR) [20].

Let $\mathbf{X} = \{\boldsymbol{x}_j | \boldsymbol{x}_j \in R^D, j = 1, \ldots, N\}$ be a specified dataset in which $\boldsymbol{x}_j (j = 1, \ldots, N)$ represents one data sample, and let N and D denote the data dimension and the size of dataset, respectively. The formulation of the QWGSD-FC algorithm can be written as

$$\min_{U,V}\left(\psi(U,V)=\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}^{2}\left\|x_j-v_i\right\|^2+\beta\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}^{2}\right),\quad s.t.\ 0\le u_{ij}\ \ and\ \ \sum_{i=1}^{C}u_{ij}=1. \tag{1}$$

where $C(1 < C < N)$ represents the number of clusters, $\boldsymbol{V} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_C]^T \in R^{C\times D}$ signifies the cluster centroid matrix that consists of the cluster prototype centers, and $\boldsymbol{v}_i \in R^D$, $i = 1, \ldots, C$; $\boldsymbol{U} \in R^{C\times N}$ represents the membership matrix, and each element u_{ij} denotes the fuzzy membership of data instance $\boldsymbol{x}_j$ with regard to cluster prototype $\boldsymbol{v}_i$; $d_{ij} = \|\boldsymbol{x}_j - \boldsymbol{v}_i\|^2$ signifies the distance between cluster centroid $\boldsymbol{v}_i$ and data instance $\boldsymbol{x}_j$; and $\beta > 0$ is the regularization parameter.

Utilizing the Lagrange optimization, the iteration equations of cluster centroid $\boldsymbol{v}_i$ and membership u_{ij} in Eq. (4) are respectively formulated as

$$v_i=\frac{\sum_{j=1}^{N}u_{ij}^{2}x_j}{\sum_{j=1}^{N}u_{ij}^{2}}, \tag{2}$$

$$u_{ij}=\frac{1}{\left(2\left\|x_j-v_i\right\|^2+2\beta\right)\sum_{k=1}^{C}\frac{1}{2\left\|x_j-v_i\right\|^2+2\beta}}. \tag{3}$$

2.3 Soft subspace clustering

Soft subspace clustering (SSC) algorithms utilize the weight coefficient matrix to partition high-dimensional data into different clusters. The weight coefficient matrix describes the contribution of different dimensions to the specified cluster. SSC is generally regarded as an improvement on traditional feature-weighted clustering. There are three types of SSC: conventional soft subspace clustering (CSSC) [8], independent soft subspace clustering (ISSC) and extended soft subspace clustering (XSSC) [9]. Specifically, the traditional feature-weighted clustering algorithm, in which the coordinate subspace and a mutual weight vector are applied, belongs to the CSSC category. In contrast, different clusters in ISSC possess different weight vectors; i.e., each cluster has a separate subspace. Therefore, ISSC can also be called multifeature weighted clustering. As an extension of CSSC or ISSC, XSSC uses other mechanisms to improve the performance of the clustering algorithm. Table 1 describes the three types of SSC algorithms.

Table 1 Three types of soft subspace clustering algorithms

SSC algorithms	Description
Conventional SSC (CSSC) algorithms	Traditional feature-weighted clustering algorithm in which all clusters share the same subspace and weight vector.
Independent SSC (ISSC) algorithms	Multidimensional feature-weighted clustering algorithm in which each cluster has independent weighting factors; i.e., each cluster has its own subspace.
Extended SSC (XSSC) algorithms	Extension of the CSSC or ISSC algorithms, which uses new techniques and mechanisms to improve the performance of the algorithm.

3 Multiview collaboration-based fuzzy soft subspace clustering (MVC-FSSC)

Let us summarize some concepts related to the fuzzy clustering model — QWGSD-FC — before we introduce our work.

According to [23], QWGSD-FC integrates most of the advantages of MEC, FCM and FC-QR, and its formulation function is composed of two main terms, as expressed in Eq. (1). The first term, $\sum_{i=1}^{C}\sum_{j=1}^{N}u_{ij}^2\left\|\boldsymbol{x}_j-\boldsymbol{v}_i\right\|^2$, calculates the total differences of all data samples $\boldsymbol{x}_j$, $j = 1, \ldots, N$ with respect to all cluster centroids $\boldsymbol{v}_i$, $i = 1, \ldots, C$, where u_{ij}^2 are the weight elements. The second term, $\beta\sum_{i=1}^{N}\sum_{j=1}^{C}u_{ij}^2$, which is derived from the Gini-Simpson index and equals the quadratic function in FC-QR, utilizes the maximum-entropy inference (MEI) principle to achieve an unbiased probability distribution in the clustering process [11, 25]. The quadratic weight u_{ij}^2 calculates the total intracluster deviation. This soft clustering model has three properties. First, the quadratic weights can effectively convey the desired individual impact of each $d_{ij} = \|\boldsymbol{x}_j - \boldsymbol{v}_i\|^2$in the deviation calculation. Second, the quadratic weight form in QWGSD-FC is easier to handle than the form in FC-QR, thus the iteration formulations of bothu_{ij} and $\boldsymbol{v}_i$ can be derived conveniently. Third, and most importantly, as demonstrated in [23], this model outperforms many existing models in practical applications.

Based on the above summary, we will introduce the two mechanisms that form the foundation of our work.

3.1 Two mechanisms for collaborative multiview soft subspace fuzzy clustering

3.1.1 Mechanism of quadratic-weighted collaborative learning (QW-CoL)

The collaborative learning mechanism designed in this paper aims at utilizing multiple views simultaneously, and the clustering information in each view is designed to be learned from other views. Let k ($k\in[1,K]$) be the view indicatrix, K signify the total view number, and $u_{ij,k}$ represent the membership matrix of data instance j ($j\in[1,N]$) to cluster i ($i\in[1,C]$) in the kth view. Then, the formulation for our QW-CoL mechanism is expressed as

$$\tilde{u}_{ij,k\sigma} = \sigma u_{ij,k}^2 + \frac{1-\sigma}{K-1}\sum_{\substack{k' \neq k \\ k' = 1}}^{K} u_{ij,k'}^2 \tag{4}$$

where $\sum_{k'\neq k}u_{ij,k'}^2$ denotes the sum of the membership of a data sample j ($j\in[1,N]$) to a cluster centroid i ($i\in[1,C]$) in all views excluding the kth view, $\tilde{u}_{ij,k\sigma}$ denotes the integrated membership degree matrix concerning data instance j ($j\in[1,N]$) to cluster i ($i\in[1,C]$) in view k, and$\sigma\in(0,1]$ denotes a tradeoff parameter.

During the process of clustering, the membership matrix $\tilde{u}_{ij,k\sigma}$, which incorporates clustering knowledge from all other views into the current view and in which parameter σ controls the importance of each view and implements the information exchange between views. Drawing on the theory behind QWGSD-FC, quadratic weights are used as fuzzy degrees of the membership matrix.

3.1.2 Mechanism of soft subspace weighted attributes (SSWA)

The SSWA mechanism aims to measure the contributions of a single dimension to the formula of a specified cluster and to generate the soft subspace attribute for every cluster. For this purpose, inspired by the original formulation of QWGSD-FC (Eq. (1)), SSWA is defined by

$$\begin{aligned} &\min\left(\psi_{ij,k} = \sum_{l=1}^{d_k} w_{il,k}^2 \left\|x_{jl,k} - v_{il,k}\right\|^2 + \lambda_2 \sum_{i=1}^{C} \sum_{l=1}^{d_k} w_{il,k}^2\right) \\ &s.t. \quad \sum_{l=1}^{d_k} w_{il,k} = 1 \end{aligned} \tag{5}$$

where d_k signifies the number of data dimensions in view k, $w_{il,\,k}$ denotes the weight of attribute $l\,(l \in [1, d_k])$ for cluster i in view k, $x_{jl,k} \in \boldsymbol{x}_{j,k} = \left[x_{j1,k}, x_{j2,k}, \ldots, x_{jd_k,k}\right]^T$ represents the lth feature value of the jth instance in the kth view, $v_{il,k} \in \boldsymbol{v}_{i,k} = \left[v_{i1,k}, v_{i2,k}, \ldots, v_{id_k,k}\right]^T$ denotes the lth feature value of the ith cluster prototype in the kth view, and $\lambda_2 > 0$ is a regularization factor.

The SSWA formula consists of two terms. The first term, $\sum_{l=1}^{d_k} w_{il,k}^2 \left\|x_{jl,k} - v_{il,k}\right\|^2$, calculates the feature-weighted matrix; in this term, the larger $w_{il,\,k}$ is, the more the lth-dimensional feature in view k contributes to the ith cluster. The second term, $\sum_i \sum_l w_{il,k}^2$, which is similar to $\sum_i \sum_j u_{ij}^2$ in the formula of QWGSD-FC (see Eq. (1)), denotes the Gini-Simpson index item. The principle of statistical MEI is used to achieve an unbiased probability distribution in the clustering process.

3.2 Novel framework for the MVC-FSSC algorithm

Our novel MVC-FSSC algorithm is proposed to solve the dimensional problems in multiview clustering tasks by organically incorporating the QW-CoL and SSWA mechanisms into classic QWGSD-FC. With the same notations as in Eqs. (1), (4) and (5), the framework of MVC-FSSC can be formulated as

$$\begin{aligned} &\min\left(\begin{aligned} &J_{MVC-FSSC}(U_1, V_1, w_1, \ldots, U_K, V_K, w_K) = \\ &\quad \sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{j=1}^{N} \left(\tilde{u}_{ij,k\sigma} \sum_{l=1}^{d_k} w_{il,k}^2 \left\|x_{jl,k} - v_{il,k}\right\|^2\right) + \lambda_1 \sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij,k}^2 + \lambda_2 \sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{d=1}^{d_k} w_{il,k}^2, \\ &\tilde{u}_{ij,k\sigma} = \sigma u_{ij,k}^2 + \frac{1-\sigma}{K-1} \sum_{\substack{k' \neq k \\ k' = 1}}^{K} u_{ij,k'}^2 \end{aligned}\right) \\ &s.t. \quad u_{ij,k} \in [0,1], w_{il,k} \in [0,1], \sum_{i=1}^{C} u_{ij,k} = 1, \sum_{i=1}^{C} w_{il,k} = 1, 1 \le i \le C, 1 \le j \le N, 1 \le k \le K \end{aligned} \tag{6}$$

where $x_{jl,\ k} \in \boldsymbol{x}_{j,\ k}$, $v_{il,\ k} \in \boldsymbol{v}_{i,\ k}$, $U_k = \{u_{ij,\ k} | u_{ij,\ k} \in R^{C \times N},\ k = 1, \ldots, K\}$, $w_k = \left\{w_{il,k} | w_{il,k} \in R^{C \times d_k},\ k = 1, \ldots, K\right\}$, and $\lambda_1 > 0$ and $\lambda_2 > 0$ are two regularization factors.

In Eq. (6), $\sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij,k\sigma} \sum_{l=1}^{d_l} w_{il,k}^2 \left\|x_{jl,k} - v_{il,k}\right\|^2$ calculates the total deviation between the cluster centroids and the data instances in all views, whereas the remainder,

$\lambda_1 \sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij,k}^2 + \lambda_1 \sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{l=1}^{d_k} w_{il,k}^2$, is composed of two Gini-Simpson index terms that take advantage of the MEI principle to achieve an unbiased probability distribution in the clustering process. The membership, $u_{ij,\ k}$ and the weight matrix, $w_{il,\ k}$, are represent two types of probabilities; the former represents the probability that the jth data sample in the kth view belongs to the ith cluster, and the latter represents the contribution of the lth feature in the kth view to the ith cluster. The larger the value of $u_{ij,\ k}$ or $w_{il,\ k}$ is, the larger the corresponding probability.

Theorem 1 The iteration formulation of the cluster centroids, fuzzy memberships and weighted factors in Eq. (6) are derived as follows:

$$v_{il,k} = \frac{\sum_{j=1}^{N} \tilde{u}_{ij,k\sigma} w_{il,k}^2 x_{jl,k}}{\sum_{j=1}^{N} \tilde{u}_{ij,k\sigma} w_{il,k}^2} \tag{7}$$

$$u_{ij,k} = \frac{1}{\left(\sigma \sum_{l=1}^{d_k} w_{il,k}^2 \left\| x_{jl,k} - v_{il,k} \right\|^2 + \frac{1-\sigma}{K-1} \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \sum_{l=1}^{d_k} w_{il,k'}^2 \left\| x_{jl,k'} - v_{il,k'} \right\|^2 + \lambda_1 \right) \sum_{p=1}^{C} \frac{1}{\sigma \sum_{l=1}^{d_k} w_{pl,k}^2 \left\| x_{jl,k} - v_{pl,k} \right\|^2 + \frac{1-\sigma}{K-1} \sum_{\substack{k'=1 \\ k' \neq k}}^{K} \sum_{l=1}^{d_k} w_{pl,k'}^2 \left\| x_{jl,k'} - v_{pl,k'} \right\|^2 + \lambda_1}} \tag{8}$$

$$w_{il,k} = \frac{1}{\left(\sum_{j=1}^{N} \tilde{u}_{ij,k\sigma} \left\| x_{jl,k} - v_{il,k} \right\|^2 + \lambda_2 \right) \sum_{q=1}^{d_k} \frac{1}{\sum_{j=1}^{N} \tilde{u}_{ij,k\sigma} \left\| x_{jq,k} - v_{iq,k} \right\|^2 + \lambda_2}} \tag{9}$$

where $\tilde{u}_{ij,k\sigma} = \sigma u_{ij,k}^2 + \frac{1-\sigma}{K-1} \sum_{\substack{k'=1 \\ k' \neq k}}^{K} u_{ij,k'}^2$.

Proof Utilizing Lagrange optimization, the target minimization formula of $J_{MVC-FSSC}$ can be converted into the following form:

$$\begin{aligned} L = & \sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{j=1}^{N} \left(\tilde{u}_{ij,k\sigma} \sum_{l=1}^{d_k} w_{il,k}^2 \left\| x_{jl,k} - v_{il,k} \right\|^2 \right) + \lambda_1 \sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij,k}^2 + \lambda_2 \sum_{k=1}^{K} \sum_{i=1}^{C} \sum_{d=1}^{d_k} w_{il,k}^2 \\ & + \sum_{j=1}^{N} \alpha_j \left(1 - \sum_{i=1}^{C} u_{ij,k} \right) + \sum_{j=1}^{N} \beta_j \left(1 - \sum_{i=1}^{C} w_{il,k} \right) \end{aligned} \tag{10}$$

where $\alpha_j (j \in [1, N])$ and $\beta_j (j \in [1, N])$ are Lagrange multipliers.

Then, set the derivatives with respect to $v_{il,\,k}$, $u_{ij,\,k}$ and $w_{il,\,k}$ to zero:

$$\frac{\partial L}{\partial v_{il,k}}=-2\sum_{j=1}^{N}\tilde{u}_{ij,k\sigma}w_{il,k}^{2}\left(x_{jl,k}-v_{il,k}\right)=0 \tag{11}$$

Eq. (7) can be acquired by rearranging Eq. (11).

$$\begin{aligned}
&\frac{\partial L}{\partial u_{ij,k}}=2u_{uj,k}\left(\sigma\sum_{l=1}^{d_k}w_{il,k}^{2}\left\|x_{jl,k}-v_{il,k}\right\|^{2}+\frac{1-\sigma}{K-1}\sum_{\substack{k'=1\\k'\neq k}}^{K}w_{il,k'}^{2}\left\|x_{jl,k'}-v_{il,k'}\right\|^{2}\right)+2\lambda_1 u_{ij,k}-\alpha_j=0\\
&\Leftrightarrow 2\left(\sigma\sum_{l=1}^{d_k}w_{il,k}^{2}\left\|x_{jl,k}-v_{il,k}\right\|^{2}+\frac{1-\sigma}{K-1}\sum_{\substack{k'=1\\k'\neq k}}^{K}w_{il,k'}^{2}\left\|x_{jl,k'}-v_{il,k'}\right\|^{2}+\lambda_1\right)u_{uj,k}=\alpha_j\\
&\Leftrightarrow u_{ij,k}=\frac{\alpha_j}{2\left(\sigma\sum_{l=1}^{d_k}w_{il,k}^{2}\left\|x_{jl,k}-v_{il,k}\right\|^{2}+\frac{1-\sigma}{K-1}\sum_{\substack{k'=1\\k'\neq k}}^{K}w_{il,k'}^{2}\left\|x_{jl,k'}-v_{il,k'}\right\|^{2}+\lambda_1\right)}
\end{aligned} \tag{12}$$

$\sum_{p=1}^{C}u_{pj,k}=1$ and through Eq. (12), we obtain

$$\begin{aligned}
&\frac{\alpha_j}{2}\sum_{p=1}^{C}\frac{1}{\sigma\sum_{l=1}^{d_k}w_{pl,k}^{2}\left\|x_{jl,k}-v_{pl,k}\right\|^{2}+\frac{1-\sigma}{K-1}\sum_{\substack{k'=1\\k'\neq k}}^{K}w_{pl,k'}^{2}\left\|x_{jl,k'}-v_{pl,k'}\right\|^{2}+\lambda_1}=1\\
&\Leftrightarrow\frac{\alpha_j}{2}=\frac{1}{\sum_{p=1}^{C}\frac{1}{\sigma\sum_{l=1}^{d_k}w_{pl,k}^{2}\left\|x_{jl,k}-v_{pl,k}\right\|^{2}+\frac{1-\sigma}{K-1}\sum_{\substack{k'=1\\k'\neq k}}^{K}w_{pl,k'}^{2}\left\|x_{jl,k'}-v_{pl,k'}\right\|^{2}+\lambda_1}}
\end{aligned} \tag{13}$$

By substituting Eq. (13) into Eq. (12), we immediately obtain Eq. (8).

Likewise,

$$\begin{aligned}
&\frac{\partial L}{\partial w_{il,k}}=2w_{il,k}\sum_{j=1}^{N}\tilde{u}_{ij,k\sigma}\left\|x_{jl,k}-v_{il,k}\right\|^{2}+2\lambda_2 w_{il,k}-\beta_j=0\Leftrightarrow\\
&w_{il,k}=\frac{\beta_j}{2\left(\sum_{j=1}^{N}\tilde{u}_{ij,k\sigma}\left\|x_{jl,k}-v_{il,k}\right\|^{2}+\lambda_2\right)}
\end{aligned} \tag{14}$$

Since $\sum_{q=1}^{C} w_{ql,k} = 1$, via Eq. (14), we obtain

$$\frac{\beta_j}{2}\sum_{q=1}^{C}\frac{1}{\sum_{j=1}^{N}\tilde{u}_{qj,k\sigma}\left\|x_{jl,k}-v_{ql,k}\right\|^2+\lambda_2}=1\Leftrightarrow\frac{\beta_j}{2}=\frac{1}{\sum_{q=1}^{C}\frac{1}{\sum_{j=1}^{N}\tilde{u}_{qj,k\sigma}\left\|x_{jl,k}-v_{ql,k}\right\|^2+\lambda_2}} \tag{15}$$

Substituting Eq. (15) into Eq. (14) yields Eq. (9).

From the method proposed above, we obtain an optimum soft partition for individual views. To obtain global clustering, as in other multiview clustering approaches, we use the geometric mean [3, 7] of the membership matrix of all views as the overall clustering decision, which is denoted as $\tilde{U}$:

$$\tilde{U} = \sqrt[K]{\prod_{k=1}^{K} U_k} \tag{16}$$

Now, we detail the MVC-FSSC algorithm as follows.

Input: The multiview datasets that contain K views $\boldsymbol{X}_K = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_K\}$, where $\boldsymbol{x}_i$ denotes the data in ith view; the number of clusters C; the maximum number of iterations max_iter; the convergence threshold ε; the regularization factors λ_1 and λ_2; and the interview weighted parameter σ.

Output: The global membership degree matrix $\tilde{\boldsymbol{U}}$ and the cluster to which each object belongs.

Step 1: Set the iteration counter to $t = 0$ and in the kth view, initialize the membership matrix $\boldsymbol{U}_k(\mathrm{t}) = [u_{ij,k}(t)]_{C\times N}$ under the conditions of $\sum_{i=1}^{C} u_{ij,k}(t) = 1$, and $u_{ij,k}(t) \in [0, 1]$, and initialize the feature-weighted matrix $\boldsymbol{w}_k(t) = \left[w_{il,k}(t)\right]_{C\times d_k}$ under the conditions of, $\sum_{i=1}^{C} w_{il,k}(t) = 1$, and $w_{il,k}(t) \in [0, 1]$;

Step 2: Calculate the centroids $\boldsymbol{V}_k(t) = [\boldsymbol{v}_{1,k}(t), \boldsymbol{v}_{2,k}(t), \ldots, \boldsymbol{v}_{C,k}(t)]^T$ with $\boldsymbol{v}_{i,k}(t) = \left[v_{i1,k}(t), v_{i2,k}(t), \ldots, v_{id_k,k}(t)\right]^T, i = 1, \ldots, C$, using Eq. (7), $\boldsymbol{U}_k(\mathrm{t}) = [u_{ij,k}(t)]_{C\times N}$, and $\boldsymbol{w}_k(t) = \left[w_{il,k}(t)\right]_{C\times d_k}$.

Step 3: Compute $J_{VC-SSFC-QWGSD}(t)$ in Eq. (6) using $\boldsymbol{U}_k(\mathrm{t})$, $\boldsymbol{w}_k(t)$ and $\boldsymbol{V}_k(t)$, where $k = 1, \ldots, K$;

Step 4: In the kth ($k = 1,\ldots,K$,) view, calculate the membership $\boldsymbol{U}_k(\mathrm{t}+1) = [u_{ij,k}(t+1)]_{C\times N}$ using Eq. (8), $\boldsymbol{V}_k(t) = [\boldsymbol{v}_{1,k}(t), \boldsymbol{v}_{2,k}(t), \ldots, \boldsymbol{v}_{C,k}(t)]^T$, and $\boldsymbol{w}_k(t) = \left[w_{il,k}(t)\right]_{C\times d_k}$;

Step 5: In the kth view ($k = 1,\ldots,K$,), compute the feature-weighted matrix $\boldsymbol{w}_k(t+1) = \left[w_{il,k}(t+1)\right]_{C\times d_k}$ using Eq. (9), where $\boldsymbol{U}_k(\mathrm{t}+1) = [u_{ij,k}(t+1)]_{C\times N}$ and $\boldsymbol{V}_k(t) = [\boldsymbol{v}_{1,k}(t), \boldsymbol{v}_{2,k}(t), \ldots, \boldsymbol{v}_{C,k}(t)]^T$;

Step 6: In the kth ($k = 1,\ldots,K$,) view, computer the new cluster centroids $\boldsymbol{V}_k(t+1) = [\boldsymbol{v}_{1,k}(t+1), \boldsymbol{v}_{2,k}(t+1), \ldots, \boldsymbol{v}_{C,k}(t+1)]^T$ with $\boldsymbol{v}_{i,k}(t+1) = \left[v_{i1,k}(t+1),\ v_{i2,k}(t+1),\ \ldots,\ v_{id_k,k}(t+1)\right]^T,\ i = 1, \ldots, C$, using Eq. (7), where $\boldsymbol{U}_k(\mathrm{t}+1) = [u_{ij,k}(t+1)]_{C\times N}$ and $\boldsymbol{w}_k(t+1) = \left[w_{il,k}(t+1)\right]_{C\times d_k}$;

Step 7: Calculate $J_{VC-SSFC-QWGSD}(t+1)$ in Eq. (6) using $\boldsymbol{U}_k(\mathrm{t}+1)$, $\boldsymbol{w}_k(t+1)$, and $\boldsymbol{V}_k(t+1)$, $k = 1, \ldots, K$.

Step 8: If $|J_{VC-SSFC-QWGSD}(t+1) - J_{VC-SSFC-QWGSD}(t)| < \varepsilon$ or $t = max\_iter$, proceed to Step 9; otherwise, set $t = t + 1$ and return to Step 4;

Step 9: Obtain the global membership matrix $\tilde{U}$ via Eq. (16), and calculate the group of each individual according to the assignment of each data object in its group:
$group_j = \mathrm{argmax}_i\left(\left[\tilde{u}_{ij}|\tilde{u}_{ij}\in\tilde{\boldsymbol{U}}; i = 1, .., C\right]^T\right), j = 1, \ldots, .N.$

---

4 Experiments and results

4.1 Setup

In this section, we demonstrate the performance of our novel MVC-FSSC algorithm. Two other correlative state-of-the-art algorithms are used as comparison algorithms, namely, the collaborative fuzzy k-means (CO-FKM) [7] and QWGSD-FC [23] algorithms. The CO-FKM algorithm is a collaborative multiview clustering algorithm, whereas QWGSD-FC is the base of our work in this paper. For QWGSD-FC, we take advantage of the feature fusion strategy prior to the clustering process to handle multiview data scenes. The categories of these methods are listed in Table 2.

The experiments used two traditional validity indicators, namely, the normalized mutual information (NMI) [18, 21] and the Rand index (RI) [10, 18], to evaluate the performance of the algorithms on artificial datasets. To evaluate the efficacy of these algorithms on medical image datasets, the mean absolute prediction (MAPD), the root mean square error (RMSE) and the R value [27] are employed.

The detailed parameters range of the competing algorithms listed in Table 2, which were selected primarily according to the recommendations in the relevant literature [7, 23, 28, 32]. All experiments were performed with MATLAB 2014a on a Windows 7 64-bit computer with an Intel i3-3240 CPU with 12 GB RAM. The optimal clustering performance of each algorithm is demonstrated by the mean and standard deviation of the NMI and RI after 10 runs per dataset.

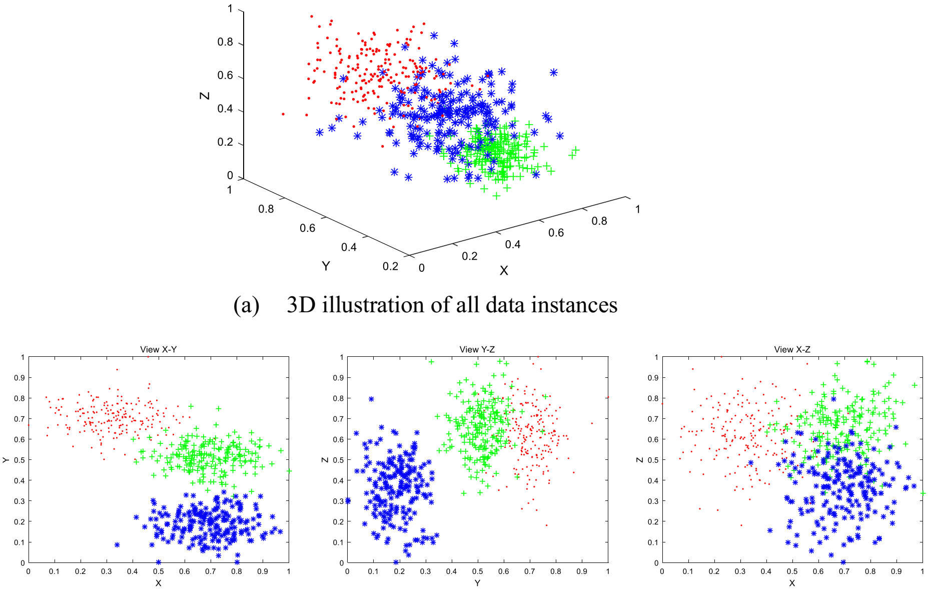
4.2 Artificial multiview datasets

To imitate datasets for multiview clustering, we generated a 3D scene with 600 data instances with X, Y and Z coordinate values, as shown in Fig. 1(a). These data are divided into 3 classes: red, green and blue (Fig. 1(a)). We examined three views of this dataset: X-Y (Fig. 1(b)), Y-Z (Fig. 1(c)), and X-Z (Fig. 1(d)). In the X-Y view, these 600 examples can be divided into 3 groups, whereas in the Y-Z and X-Z views, it is difficult to group these data points because many instances overlap at the cluster junctions.

To measure the performance of our approach, we compared our algorithm (MVC-FSSC) against two related clustering methods: CO-FKM and QWGSD-FC. Table 3 lists the metrics used to evaluate the clustering performances of these algorithms.

Table 2 Categories and parameter settings of the considered algorithms

Algorithm	Category	Parameter values or trial ranges
CO-FKM	Collaborative multiview clustering	Fuzzifier $m \in [1.05 : 0.05 : 2.5]$ Parameter $\eta \in \left[0 : 0.01 : \frac{K-1}{K}\right]$
QWGSD-FC	Soft partition clustering	Diversity measure coefficient $\beta \in [1e-4, 1e-3, 1e-2, 1e-1, 1e0, 1e1, 1e2, 1e3, 1e4]$
MVC-FSSC	Collaborative multiview clustering; Soft subspace clustering	Tradeoff factor $\sigma \in [0.01 : 0.05 : 1]$ Parameters λ_1, $\lambda_2 \in [1e-4, 1e-3, 1e-2, 1e-1, 1e0, 1e1, 1e2, 1e3, 1e4]$



(a) 3D illustration of all data instances

(b) X-Y view (c) Y-Z view (d) X-Z view

Fig. 1 Artificial 3D data scene

According to Table 3, the proposed clustering algorithm, MVC-FSSC, which benefits from a quadratic-weighted collaborative learning mechanism and a soft subspace weighted attribute mechanism, achieved the best NMI and RI scores. As shown in Fig. 1, in the X-Y view, the data instances described by the Y feature (V2 in the X-Y view) are relatively separate, which means that it is easy to determine the cluster to which those data instances belong. Correspondingly, as shown in Table 4, the attribute weights of V2 in the X-Y view are apparently larger than the attribute weights of V1. A similar conclusion can be drawn in the Y-Z view. In the X-Z view, it is difficult to distinguish data instances by either attribute X or attribute Z; hence, the attribute weights of those two features in the X-Z view approach are equally 0.5. In summary, our proposed approach can extract both inter-view and intra-view information to improve performance.

4.3 Application in medical image segmentation

Based on the results presented above, our proposed MVC-FSSC algorithm yields more accurate results in multiview scenarios than do the other methods. In this subsection, further

Table 3 Performances of the compared approaches in a synthetic 3D dataset

Dataset	Validity Metric	Algorithms		
		CO-FKM	QWGSD-FC	MVC-FSSC
Artificial 3D dataset	NMI_mean	0.6268	0.921	**0.9544**
	NMI_std	7.40E-17	0	1.17E-16
	RI_mean	0.8664	0.9781	**0.9890**
	RI_std	1.17E-16	0	1.17E-16

Table 4 Feature-weighted matrix on the artificial, 3D dataset

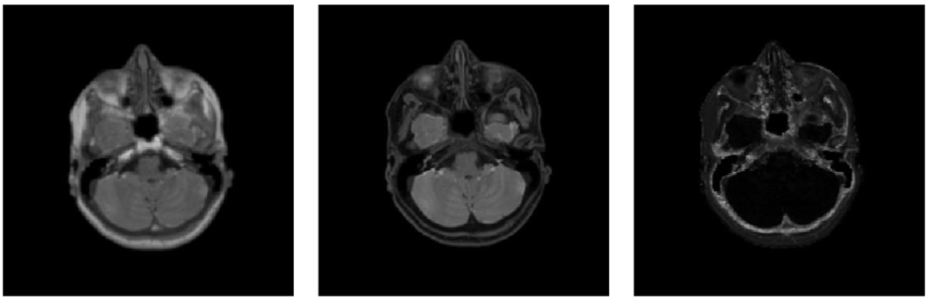
View	w_{il}			
		C_1	C_2	C_3
X-Y	V_1	0.3181	0.3419	0.3767
	V_2	0.6819	0.6581	0.6233
Y-Z	V_1	0.6841	0.6121	0.6284
	V_2	0.3159	0.3879	0.3716
X-Z	V_1	0.5027	0.4505	0.5055
	V_2	0.4973	0.5495	0.4945

*V_l denotes the lth attribute, and C_i denotes the ith cluster centroid

research and discussions of the proposed algorithm will be conducted in brain image segmentation tasks. Brain MR imaging segmentation can reflect more tissue separation information to assist physicians in the diagnosis of brain diseases. In the medical segmentation scenario, MR images are divided into five categories, i.e., Air, Bone, Fluid, Fat, and Soft Tissue, by the abovementioned algorithms. To evaluate the performance on medical image segmentation tasks, we applied our algorithm to UTE-mDixon PET/MR brain images [13, 27, 33]. Three MR sequence images ($\text{Dixon}_{\text{fat}}$, $\text{Dixon}_{\text{water}}$ and R2*) from five patients are used in this paper (Fig. 2). Each MR image contains 256 image slices and 256 × 256 pixels. To better exploit the information in the images, location features [17], which are 3D spatial coordinates that represent the anatomical location of a single point, are used in our research. Therefore, there are three views for each subject, and each view consists of four dimensions.

Table 5 compares the means and standard deviations of the RMSE, MAPD and R values from the MVC-FSSC algorithm and the other two algorithms on the five-patient dataset. To demonstrate the clustering performance, we assign a corresponding linear attenuation coefficient (LAC) to each tissue based on the clustering results and generate pseudo-CT images (Fig. 3).

According to the results in Table 1, the proposed MVC-FSSC algorithm achieves a relatively high validity index with inter-view collaboration and intraview feature weights. According to the pseudo-CT images in Fig. 3, many soft tissues are wrongly classified as bone tissue by the CO-FKM and QWGSD-FC algorithms. The pseudo-CT images generated by the MVC-FSSC segmentation result is more similar to the reference CT image than the results



(a) $\text{Dixon}_{\text{fat}}$ (b) $\text{Dixon}_{\text{water}}$ (c) $R2*$

Fig. 2 UTE-mDixon MR images

Table 5 MR brain segmentation performance comparisons

Dataset	RMSE			MAPD			R		
	CO-FKM	QWGSD-FC	MVC-FSSC	CO-FKM	QWGSD-FC	MVC-FSSC	CO-FKM	QWGSD-FC	MVC-FSSC
Patient 1	240.7	260.95	**198.46**	145.24	141.39	**114.42**	0.8804	0.7293	**0.9013**
Patient 2	**224.0**	232.90	230.95	128.79	137.75	**90.45**	**0.9090**	0.8357	0.8905
Patient 3	261.9	238.97	**228.69**	154.33	123.97	**93.58**	0.8559	0.7974	**0.8912**
Patient 4	211.0	220.55	**190.43**	121.08	131.71	**98.65**	**0.9013**	0.7766	0.9012
Patient 5	189.6	194.47	**188.36**	110.04	109.56	**96.02**	0.9014	0.8498	**0.9035**
mean	225.4	229.57	207.37	131.90	128.87	98.62	0.8896	0.7978	0.8975
std	24.71	21.89	18.64	16.03	11.31	8.34	0.02	0.0431	0.0055

from the other two competition methods. Our current work remains limited to a fixed part of the body, namely, the brain. Additional challenging anatomic segments, such as the abdomen and neck, should be examined via this technique. In addition, the use of more MR sequences and more feature maps that have been extracted from MR images could improve the performance of our algorithm.

Reference CT image

Pseudo-CT image by CO-FKM

Pseudo-CT image by QWGSD-FC

Pseudo-CT image by MVC-FSSC

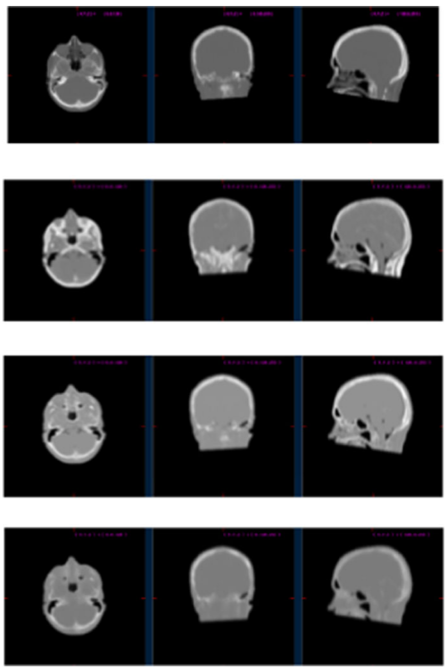


Fig. 3 Pseudo-CT images for Patient 1 by algorithm

4.4 Parameter analyses

Finally, to illustrate the reliability of our approach, we assessed the robustness of the core parameters of the MVC-FSSC algorithm: the view tradeoff parameter (σ), the Gini-Simpson diversity measure factor (λ_1), and the feature-weighted matrix measure parameter (λ_2). Due to space limitations, we separately report only the experimental results on the artificial multiview dataset and the brain MR image segmentation of Patient 4.

The performance curves for MVC-FSSC on these three datasets are plotted in Fig. 4, where Fig. 4(a), (b) and (c) correspond to the artificial dataset, and Fig. 4(d), (e) and (f) correspond to the brain MR image segmentation of Patient 4. As shown in Fig. 4, the clustering performance of the MVC-FSSC algorithm is comparatively stable. When the core factors are fixed, the evaluation index is stable in a small range, indicating that the MVC-FSSC algorithm proposed in this paper illustrates excellent robustness to parameter settings.

Table 6 illustrates the feature weight matrix on Patient 2's MR medical image segmentation dataset. In tissue segmentation tasks, the MR sequence features generally have better distinguishing performance than the 3D spatial coordinate features (Table 4). Since $\text{Dixon}_{\text{fat}}$ reflects fat tissue, the weight matrix of the second feature (Fat), contributes more than the others. Analogously, since $\text{Dixon}_{\text{water}}$ reflects water, the weight matrix of the fourth feature (Fluid), contributes more than the other features. As shown in Fig. 2, R2* is a special MR sequence that can provide more information about bones than the other sequences, which makes the value of the weight matrix on the third feature, i.e., Bone, of R2* much higher. In summary, our algorithm can automatically assign the contribution weight of each attribute in the view to each class and can analyze the feature distribution in regard to each subspace in the target dataset.

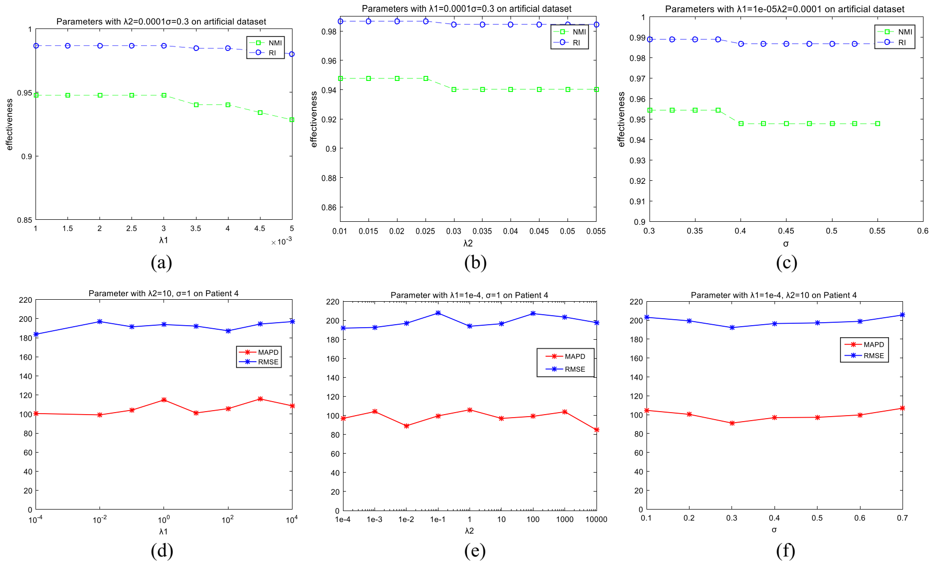


Fig. 4 Performance curves with respect to parameters λ_1, λ_2 and σ on an artificial dataset and the brain MR image segmentation of Patient 4. (a) Algorithm on the artificial dataset (λ_2 and σ are fixed, and λ_1 varies), (b) algorithm on the artificial dataset (λ_1 and σ are fixed, and $\lambda_2$varies), (c) algorithm on the artificial dataset (λ_1 and λ_2 are fixed, and σ varies), (d) algorithm on the Patient 4 dataset (λ_2 and σ are fixed, and λ_1 varies), (e) algorithm on the Patient 4 dataset (λ_1 and σ are fixed, and λ_2 varies), (f) algorithm on the Patient 4 dataset (λ_1 and $\lambda_2$are fixed, and σ varies)

Table 6 Feature-weighted matrix on the Patient 2

View		w_{il}				
		C_1 (Air)	C_2 (Fat)	C_3 (Bone)	C_4 (Fluid)	C_5 (Soft)
V1	$\text{Dixon}_{\text{fat}}$	0.8580	0.9703	0.8648	0.8440	0.4947
	X-axis	0.0446	0.0099	0.0383	0.0473	0.1194
	Y-axis	0.0613	0.0114	0.0433	0.0706	0.1528
	Z-axis	0.0361	0.0084	0.0537	0.0381	0.2332
V2	$\text{Dixon}_{\text{water}}$	0.8122	0.9846	0.7718	0.9106	0.8634
	X-axis	0.0588	0.0052	0.0520	0.0269	0.0411
	Y-axis	0.0732	0.0059	0.0634	0.0427	0.0515
	Z-axis	0.0558	0.0043	0.1128	0.0198	0.0439
V3	R2*	0.8844	0.9798	0.9505	0.3944	0.7921
	X-axis	0.0367	0.0068	0.0129	0.1582	0.0666
	Y-axis	0.0437	0.0077	0.0163	0.3143	0.0831
	Z-axis	0.0351	0.0057	0.0204	0.1331	0.0582

5 Conclusions

In this paper, the novel multiview soft subspace clustering algorithm, MVC-FSSC, is proposed. In contrast to existing multiview algorithms, the proposed algorithm can take full advantage of the interaction between views via the quadratic-weighted collaborative learning mechanism. The soft subspace weighted attribute mechanism is extended to automatically extract the contribution of each attribute to each cluster within each view. The experimental performance for both the artificial and multiview MR image datasets indicates that in each view, the MVC-FSSC algorithm performs well in terms of multiview SSC effectiveness, parameter robustness, and subspace distributions. Both the metrics and the pseudo-CT images show that the proposed multiview subspace algorithm achieves a better brain MR image segmentation performance than the other methods, which fully demonstrates the clustering performance and practical value of the proposed algorithm. However, when few features can be obtained to describe the target object, the performance of SSWA will be limited. In the future, we will extend our research on feature extraction methods to improve the MVC-FSSC algorithm, and we will utilize our algorithm to other practical applications, such as abdominal medical image analyses.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China under grants 61772241 and 61702225, the Natural Science Foundation of Jiangsu Province under grant BK20160187, the Fundamental Research Funds for the Central Universities under grant JUSRP51614A, the 2016 Qinglan Project of Jiangsu Province, the 2016 Six Talent Peaks Project of Jiangsu Province, and the Science and Technology Demonstration Project of Social Development of Wuxi under grant WX18IVJN002.

References

1. Aggarwal CC, Wolf JL, Yu PS et al (1999) Fast algorithms for projected clustering. ACM SIGMOD Rec 28(2):61–72
2. Aggarwal CC, Yu PS (2000) Finding generalized projected clusters in high dimensional spaces. ACM 29(2):70–81
3. Batuwita R, Palade V (2012) Adjusted geometric-mean: a novel performance measure for imbalanced bioinformatics datasets learning. J Bioinforma Comput Biol 10(04):1250003

4. Bezdek JC, Ehrlich R, Full W (1984) FCM: the fuzzy c -means clustering algorithm. Comput Geosci 10(2): 191–203
5. Cao Y, Wu J (2002) Projective ART for clustering data sets in high dimensional spaces. Neural Netw 15(1): 105–120
6. Chitsaz E, Jahromi MZ (2016) A novel soft subspace clustering algorithm with noise detection for high dimensional datasets. Soft Comput 20(11):4463–4472
7. Cleuziou G, Exbrayat M, Martin L, Sublemontier JH (2009) CoFKM: A centralized method for multiple-view clustering. IEEE Int Conf Data Min 752–757
8. De Soete G (1986) Optimal variable weighting for ultrametric and additive tree clustering. Qual Quant 20(2-3):169–180
9. Deng Z, Choi KS, Chung FL, Wang S (2010) Enhanced soft subspace clustering integrating within-cluster and between-cluster information. Pattern Recogn 43(3):767–781
10. Desgraupes B (2013) Clustering indices. University of Paris Ouest-Lab Modal'X 1:34
11. Domeniconi C, Gunopulos D, Ma S et al (2007) Locally adaptive metrics for clustering high dimensional data. Data Min Knowl Disc 14(1):63–97
12. Gao Y, Maggs M (2005) Feature-level fusion in personal identification. Comput Soc Conf Comput Vis Pattern Recognit 1:468–473
13. Hooijmans MT, Dzyubachyk O, Nehrke K et al (2015) Fast multistation water/fat imaging at 3T using DREAM-based RF shimming. J Magn Reson Imaging 42(1):217–223
14. Hotho A, Maedche A, Staab S (2002) Ontology-based text document clustering. KI 16(4):48–54
15. Jiang Y, Chung FL, Wang S et al (2015) Collaborative fuzzy clustering from multiple weighted views. IEEE Trans Cybern 45(4):688–701
16. Li RP, Mukaidono M (1995) A maximum-entropy approach to fuzzy clustering. IEEE Int Conf Fuzzy Syst 4:2227–2232
17. Liang F, Qian P, Su KH et al (2018) Abdominal, multi-organ, auto-contouring method for online adaptive magnetic resonance guided radiotherapy: an intelligent, multi-level fusion approach. Artif Intell Med 90:34–41
18. Liu J, Mohammed J, Carter J et al (2006) Distance-based clustering of CGH data. Bioinformatics 22(16): 1971–1978
19. Loeff N, Alm CO, Forsyth DA (2006) Discriminating image senses by clustering with multimodal features. ACL Main Conf Poster Sess 547–554
20. Miyamoto S, Umayahara K (1998) Fuzzy clustering by quadratic regularization. IEEE World Congress Comput Intell 2:1394–1399
21. Nie F, Xu D, Li X (2012) Initialization independent clustering with actively self-training method. IEEE Trans Syst Man Cybern Part B (Cybernetics) 42(1):17–27
22. Parsons L, Haque E, Liu H (2004) Subspace clustering for high dimensional data: a review. ACM Sigkdd Explor Newsl 6(1):90–105
23. Qian P, Sun S, Jiang Y et al (2016) Cross-domain, soft-partition clustering with diversity measure and knowledge reference. Pattern Recogn 50:155–177
24. Rokach L (2009) A survey of clustering algorithms. Data Min Knowl Disc Handb 269–298
25. Roth HR, Shen C, Oda H et al (2018) Deep learning and its application to medical image segmentation. Med Imaging Technol 36(2):63–71
26. Sim K, Gopalkrishnan V, Zimek A, Cong G (2013) A survey on enhanced subspace clustering. Data Min Knowl Disc 26(2):332–397
27. Su KH, Hu L, Stehning C et al (2015) Generation of brain pseudo-CTs using an undersampled, single-acquisition UTE-mDixon pulse sequence and unsupervised clustering. Med Phys 42(8):4974–4986
28. Tzortzis G, Lika A (2012) Kernel-based weighted multi-view clustering. IEEE Int Conf Data Min 675–684
29. Wang G, Li W, Zuluaga MA et al (2018) Interactive medical image segmentation using deep learning with image-specific fine tuning. IEEE Trans Med Imaging 37(7):1562–1573
30. Wang G, Liu Y, Xiong C (2015) An optimization clustering algorithm based on texture feature fusion for color image segmentation. Algorithms 8(2):234–247
31. Xiaopeng W, Shihe H, Hui Y, Wen Z (2014) The design of medical image transfer function using multi-feature fusion and improved k-means clustering. J Chem Pharm Res 6(7):2008–2014
32. Xue Z, Li G, Wang S, Zhang C, Zhang W, Huang Q (2015) GOMES: A group-aware multi-view fusion approach towards real-world image clustering. IEEE Int Conf Multimed Expo 1–6
33. Zaidi H, Ojha N, Morich M et al (2011) Design and performance evaluation of a whole-body ingenuity TF PET–MRI system. Phys Med Biol 56(10):3091
34. Zhao K, Zhou L, Qian P et al (2019) A transfer fuzzy clustering and neural network based tissue segmentation method during PET/MR attenuation correction. J Med Imaging Health Inf. Accepted

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Kaifa Zhao is a M.S. candidate at the School of Digital Media, Jiangnan University, Wuxi, Jiangsu, China. His research interests include pattern recognition as well as bioinformatics and their applications.

Yizhang Jiang received his Ph.D. degree from Jiangnan University in 2016. Now he is an Associate Professor at the School of Digital Media, Jiangnan University, Wuxi, Jiangsu, China. He has published more than 30 papers in international journals including IEEE TFS, IEEE TNNLS, IEEE T. Cyber., PR, KBS, and InS. His research interests include pattern recognition, intelligent computation, and their applications.

Kaijian Xia is now studying at China University of Mining Technology for the Ph.D. degree. He received the M.S. degree of computer science and technology from Jiangnan University in 2009. From 2009 to 2010, he worked as a lecturer at the School of Computer Science and Engineering, Changshu Institute of Technology. Since 2010, He has been working at the Changshu NO.1 People's Hospital. His research interests are in Biological image processing and Computational medicine, etc. He has published near 20 papers in international/national journals.

Leyuan Zhou now is the deputy director of the Department of Radiotherapy, Affiliated Hospital of Jiangnan University. He received the Ph.D. degree of Oncology from Fudan University in 2012. Now he is the Member of Radiotoxicology Committee of Jiangsu Toxicology Society, the Youth Member of Jiangsu Immunological Society, the Member of Cancer Radiotherapy Committee of Jiangsu Society of Integrated Traditional Chinese and Western Medicine, and the Key Talents of Science, Education and Health Engineering Medicine of Wuxi. From November 2014 to May 2015, he worked as a visiting scholar in the Department of Radiotherapy, University of Michigan. He has published near 20 papers in international journals, such as Int J Radiat Oncol Biol Phys, the most authoritative journal in the field of cancer radiotherapy.

Yangyang Chen is a M.S. candidate at the School of Digital Media, Jiangnan University, Wuxi, Jiangsu, China. His research interests include artificial intelligence and its application in medical image processing.

Ke Xu is a M.S. candidate at the School of Digital Media, Jiangnan University, Wuxi, Jiangsu, China. His research interests include deep learning and its application in multimedia data.

Pengjiang Qian received his Ph.D. degree from Jiangnan University in March, 2011. Now he is a Full Professor at the School of Digital Media, Jiangnan University, Wuxi, Jiangsu, China. He has authored or co-authored near 60 papers published in international/national journals and conferences, e.g., IEEE TNNLS, IEEE TSMC-B, IEEE T. Cyber., IEEE TFS, PR, InS, and KBS. His research interests include data mining, pattern recognition, bioinformatics and their applications, such as analysis and processing for medical imaging, intelligent traffic dispatching, and advanced business intelligence in logistics.

Affiliations

Kaifa Zhao[1] · Yizhang Jiang[1] · Kaijian Xia[2] · Leyuan Zhou[3] · Yangyang Chen[1] · Ke Xu[1] · Pengjiang Qian[1]

✉ Pengjiang Qian
qianpjiang@jiangnan.edu.cn

Kaifa Zhao
zhaokaifa@live.com

Yizhang Jiang
jyz0512@163.com

Kaijian Xia
xiakaijian@163.com

Leyuan Zhou
zhouleyuan99@126.com

Yangyang Chen
chenyangyang0620@163.com

Ke Xu
2428915905@qq.com

1 School of Digital Media, Jiangnan University, Wuxi 214122, China

2 Changshu No.1 People's Hospital, Suzhou 215500, China

3 Affiliated Hospital, Jiangnan University, Wuxi 214000, China