


# Low-Rank Subspace Learning of Multikernel Based on Weighted Truncated Nuclear Norm for Image Segmentation

LI LI<sup>1</sup>, XIAO WANG<sup>2</sup>, LEI PU<sup>2</sup>, JING WANG<sup>2</sup>, AND XIAOQIAN ZHANG<sup>1</sup><sup>2</sup>, (Member, IEEE)

<sup>1</sup>School of Computer Science and Technology, Southwest University of Science and Technology, Mianyang 621010, China

<sup>2</sup>School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China

Corresponding author: Xiaoqian Zhang (zhxq0528@163.com)


This work was supported in part by the National Natural Science Foundation of China under Grant 62102331, in part by the Natural Science Foundation of Sichuan Province under Grant 2022NSFSC0839, and in part by the Southwest University of Science and Technology Doctoral Fund Project 22zx7110.

**ABSTRACT** Previous natural image segmentation algorithms through subspace learning method have over-segmentation issues in the pre-segmentation process, which will compromise the edge information, and the subspace learning model cannot effectively utilize the nonlinear structure in the image data and has weak resistance to multiple noises. To address these problems, a multi-kernel subspace learning method based on weight truncated Schatten-p norm for image segmentation is designed in this paper. First, the original natural image pre-processing operation, which is conducting adaptive morphological reconstruction watershed transformation on the image, then the original pixels are aggregated to form a superpixel image, of which the obtained superpixel block would retain more comprehensive local information; Secondly, perform feature extraction for each superpixel block, and stack the obtained feature vectors into the desired feature matrix; Then, it is input into the weighted truncated Schatten-p low-rank multi-kernel subspace learning model to obtain a similarity matrix with cluster structure on the diagonal; Finally, the similarity matrix is used as the adjacency matrix in the spectral clustering model, and the final feature data clustering and image segmentation results are obtained by the optimization solution. The final experimental results demonstrate that contrasts to existing clustering models, the proposed method accomplishes the best clustering property on two public datasets; Compared with seven segmentation algorithms on the BSDS500 dataset, and achieved the best segmentation effect on two evaluation metrics.

**INDEX TERMS** Subspace learning, image segmentation, multi-kernel, superpixel, spectral clustering.

## I. INTRODUCTION

The segmentation [1]–[3] of natural images is a current difficult research area in the direction of machine vision and image recognition analysis. It refers to grouping the single pixels in the original input image to form a series of regions according to their feature similarity, and to separate foreground objects and image backgrounds. Segmentation methods for natural images are usually considered as five types, i.e., using specific threshold, pixel region-based, contour detection-based, clustering-based, and neural network-based segmentation methods. Subspace learning is often used in popular fields such as data analysis and face recognition,

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

and can effectively cluster high-dimensional data in images [4]–[6]. The theoretical assumption [7] can be understand as the known high-dimensional image feature data is contained in the coalition of many lower-dimensional feature subspaces. The high-dimensional image feature data is represented by a lower-dimensional feature subspace, and the representation coefficients are used to gain spatial properties of original image feature data.

The natural image segmentation method based on subspace learning [8]–[11] regards the segmentation of natural images as a data clustering issue of regional pixel feature. As shown in Fig.1, the feature information of the superpixel block is extracted from the preprocessed image, then the subspace learning algorithm is used to gather the feature data into clusters to achieve segmentation. Two more classical

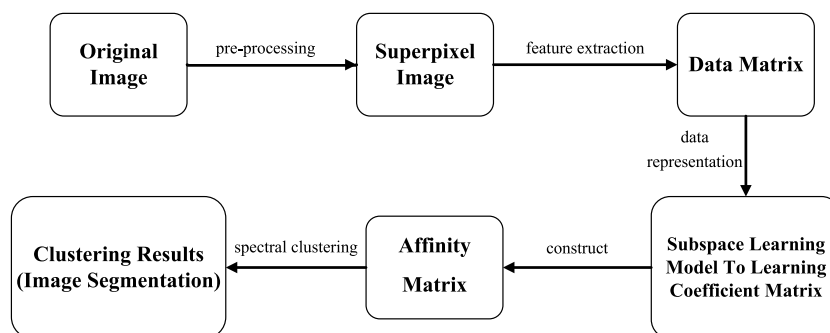


FIGURE 1. The Process of Image Segmentation Based on Subspace Learning.

subspace learning methods are sparse subspace clustering by Vidal *et al.* [12] and low-rank expression model of subspace clustering designed by Liu *et al.* [13], they impose a simple  $l_1$  norm or nuclear norm to the coefficient matrix. However, this does not help the diagonal of coefficient matrix to form a better cluster structure, which will cause confusions in the classification of image feature data. Subsequent research scholars have expanded and derived many linear subspace learning algorithms. By fusing the sparse representation with the low-rank representation model, making non-negative judgments, Zhuang *et al.* [14] designed a method to joint low-rank model and sparse representation in non-negative(NNLRSC), Hu *et al.* [15] designed a clustering model using smoothed representation (SMRC), Wang *et al.* [16] proposed adding low-rank constraints on the basis of SSC (LRSSC). The previous algorithms are mostly using linear self-representation of the features data, which can only work in linear subspace, but the feature data generated by actual images are often nonlinear structures.

Some researchers use kernel techniques to map nonlinear structural data to higher-dimensional Reproducing Kernel Hilbert Space (RKHS) [17], for conversion into linear structural data for processing, which can make up for the limitations of linear subspace learning methods in processing image feature data. Among the single-kernel subspace learning algorithms, such as Spectral Clustering (SC) [18], RKKM (Robust K-Means Using Kernel) [19] and model by Simplex Sparse Representation (SLSR) [20] are relatively popular. Tao *et al.* [21] designed a new sparse expression clustering algorithm for natural image segmentation, which takes weighted sparse constraints to better represent feature data within clusters. Li *et al.* [22] designed an improved sparse expression subspace clustering model to natural image segmentation, introducing the  $l_{p,1}$  norm to expand, so that the coefficient components tend to be uniform within the class and sparse among the classes. But, the data clustering and image segmentation capabilities of these single kernel models depends mainly on the goodness of the mapping function. so many researchers have made further enhancements to this base. The Multi-Kernel Learning (MKL) [23] frame of subspace clustering emerged. In the framework of

multi-kernel fusion learning, the representation of image data samples in the feature space mainly depends on the sparse allocation of basic kernels and weights. Current mainstream multi-kernel models, such as Robust k-means using Multi-Kernel (RMKKM) [24], Multi-kernel learning model using Self-weighted (SMKL) [25], Multi-kernel Spectral clustering (SCMK) [26], Graph rank by Low-rank Kernel representation (LKGr) [27], but the low rank property of the data structure obtained by these methods is weaker in natural image feature space. The low-rank constraint imposed by the relatively simple kernel norm on the natural image data, which is not sufficient to make the data possess good low-rank properties. Therefore, this paper incorporates weighted truncated Schatten-p norm on the basis of multi-kernel subspace learning to impose low-rank constraints. The major dedications of this work are as follows:

- A method for natural image segmentation based on weighted truncated Schatten-p multi-kernel subspace learning (AMR-WT-WTSPSC) is proposed. The adopted multi-kernel subspace framework not only enables processing of nonlinear data, but also avoids the limitations of using a single predefined kernel.
- The weighting of the correntropy metric is incorporated into the multi-kernel weighting model, which improves the previous issue of fixed weights or insufficient weight distribution, and helps to study the good consensus kernel. The optimal kernel learned with a weighted truncation Schatten-p constraint promotes it to maintain a good low-rank structure.
- Clustering experiments with nine algorithms are performed on two public face image datasets, and segmentation experiments are performed on the BSDS500 dataset. The results demonstrate that the designed method receives better clustering performance and segmentation results.

The remainder of the paper is organized as follows. We slightly comment on some developments of previous methods in Section II. In Section III, we design a WTSPSC clustering method, then list the optimal solution of the novel model in detail. In subsequent Section IV, the complete experimental and theoretical analysis is provided. The conclusion are presented in Section V.

**II. RELATED WORKS**

Next section, a brief description of some developments from previous methods, i.e., the adaptive morphological reconstruction watershed transformation method, and the kernel trick.

**A. WATERSHED TRANSFORM FOR ADAPTIVE MORPHOLOGICAL RECONSTRUCTION FOR IMAGE PRE-SEGMENTATION**

Seed region segmentation algorithms [28], such as Random Walk(RW) [29], Graph Cut(GC) [30] and Watershed Transform (WT) [31], etc., have been widely used in image pre-segmentation process. Among them WT pre-segment an image by combining topological ideas with mathematical morphology, in which takes the similarity of two adjacent pixels as an criterion, so that the pixels that are spatially adjacent and have similar characteristics are merged to form independent contour regions. The basic steps of WT are to convert the original natural image to grayscale natural image, then calculate the gradient image, and then watershed transformation is applied to the gradient image to gain pre-segmented boundary image. However, when the WT algorithm is executed, the redundant seeds in the gradient image will cause the image to be over-segmented. To this end, Lei *et al.* [32] proposed a watershed transformation method for adaptive morphological reconstruction to avoid over-segmentation of images. AMR-WT employ multi-scope architecture elements to rehabilitate the obtained extracted images, and then performs point-by-point extremum operations on multiple gradient images, finally the self-use reconstruction results are obtained. Specifically set as follows:

*Definition 1:* Assume that  $k_a \subseteq \dots \subseteq k_j \subseteq k_{j+1} \subseteq \dots \subseteq k_b$  be a spectrum of nested architecture elements, where  $j$  is the scope parameter of a architecture element,  $1 \leq a \leq j \leq b$ ,  $a, j, b \in N^+$ . Relative to the gradient natural image  $d$  has that  $I = \varepsilon_{k_j}(d)$  and  $I \leq d$ , the adaptive morphological reconstruction represent by  $\phi$  of  $d$  from  $I$  is defined as:

$$\phi(d, a, b) = \vee_{a \leq j \leq b} \{R_d^\phi(I)_{k_j}\} \quad (1)$$

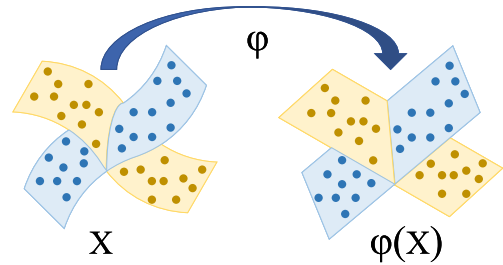
**B. SUBSPACE LEARNING BASED ON SELF-REPRESENTING LOW-RANK KERNELS**

The LRR algorithm [33] adopts the “self-expression” of the feature data, that is, the feature data points attaching to the same subspace can be linearly combined with each other to express the rest of the samples. Unluckily, this type of method cannot be applied to data with nonlinear structure, therefore, the kernel trick [34], [35] is needed to perform kernel mapping on the original nonlinear structure data to keep it in linear space for analysis (as shown in the fig. 2). The non-linear transformation model of LRR is:

$$\min_{H, C} \|\varphi(Y)\|_* + \lambda \|C\|_1 \quad \text{s.t. } \varphi(Y) = \varphi(Y)C, \quad C_{ii} = 0 \quad (2)$$

where  $Y$  is a data sample feature matrix,  $C$  is a self-expression coefficient matrix,  $H = \varphi(Y)^T \varphi(Y)$  is Gram matrix,

$\varphi(Y) = [\varphi(y_1), \varphi(y_2), \dots, \varphi(y_N)]$  represents the mapped linear space, and  $\lambda$  is a balance parameter. The best alternative to rank  $(\varphi(Y))$  is  $\|\varphi(Y)\|_*$ . This guarantees that the transformed feature sample data can be distributed in multiple linear feature subspaces.



**FIGURE 2.** The mapping of nonlinear structure data.

**III. THE MODEL AND OPTIMIZATION OF MULTI-KERNEL SUBSPACE LEARNING BASED ON WEIGHT TRUNCATE SCHATTEN-P**

**A. THE MODEL OF WTSPSC**

Since the image feature data contains non-linear structure, it is mapped into a higher-dimensional RKHS to perform a linear space analysis. The learning of the similarity matrix is carried out in a self-expressive graph learning framework, mainly for spectral clustering. At the same time, the kernel mapping and self-expression characteristics are fused, and the kernel self-representation majorization problem is reformulated as formula (3):

$$\begin{aligned} \min_C & \frac{1}{2} \|\varphi(Y) - \varphi(Y)C\|_F^2 + \alpha R(C) \\ \text{s.t. } & C \geq 0, \quad \text{diag}(C) = 0 \end{aligned} \quad (3)$$

where  $C$  in the penalty term  $R(C)$  represents the kernel self-expression coefficient matrix, and  $\alpha$  is a non-negative compromise parameter.

With the kernel strategy, the kernel matrix  $E = \varphi(Y)^T \varphi(Y)$  in the kernel space is mainly considered, instead of depending on the unknown kernel function, where  $E_{ij} = \langle \varphi(Y_i), \varphi(Y_j) \rangle$  refers to the  $(i, j)$ -th element of  $E$ . In this regard, formula (3) can be formulated as:

$$\|\varphi(Y) - \varphi(Y)C\|_F^2 = \text{Tr}(E - 2EC + C^T E C) \quad (4)$$

In theoretical research, when a feature data matrix  $Y$  with  $k$  clusters or categories is input, the similarity matrix constructed by the kernel self-representation coefficient matrix  $C$  should also have  $k$  appropriately arranged block diagonals. According to Lu [36] *et al.*, in order to keep the diagonal chunk property of  $C$  in the independent underlying subspace, a Block Diagonal Regularizer (BDR) is incorporated into the model. The Laplacian matrix  $L$  homologous to the similarity matrix  $H \in R^{N \times N}$  is defined as  $L = G - H$ , where  $G_i = \sum_j H_{ij}$ . Assume that the  $i$ -th smallest eigenvalue of  $L$  is  $\lambda_i(L)$ ,  $i \in (1, N)$  in descending sequence, and the sum of the

$k$  smallest eigenvalues is BDR:

$$\|H\|_{\boxed{k}} = \sum_{i=N-k-1}^N \alpha_i(L) \quad (5)$$

The construction of a weighting strategy is the core content of MKL, which can effectively utilize complementary information, especially in practical applications with noisy situations. In the multi-kernel learning frame, assuming that every single small disturbances of the prospective best consistent kernel matrix  $E$  is the underlying kernel  $E^i$ , denoted as  $\min \|E^i - E\|_F^2$ , and the weights of individual kernels are robust to partial noise, thus considering the application of the correntropy metric induced multi-kernel learning model (CMMKL) [37] to deal with the non-linear structure of image feature data and more study the consistent kernel. Eventually, a weighted truncated Schatten-p norm ( $\|\cdot\|_{w,r}^p$ ) is used to the model to constrain the learned optimal kernels, for induce a low-rank feature structure.

Finally, the objective function model of this paper as a whole is:

$$\begin{aligned} \min_{C,E,\omega} & \frac{1}{2} \text{Tr} (E - 2EC + C^T EC) + \alpha \|C\|_{\boxed{k}} \\ & + \beta \sum_{h=1}^r \omega_i \|E^i - E\|_F^2 + \lambda \|E\|_{w,r}^p \\ \text{s.t. } & C \geq 0, \quad \text{diag}(C) = 0, \quad C = C^T \end{aligned} \quad (6)$$

where  $\alpha$  and  $\beta$  are the non-negative penalty parameters of the model, associated with BDR and CMMKL.  $\lambda > 0$ , the  $w_i$  represents the  $i$ -th predefined base kernel.

### B. THE OPTIMIZATION OF WTSPSC

Since the BDR is adopted in the model, the constraint of  $C$  is non-negative and symmetric, namely,  $C \geq 0$ ,  $\text{diag}(C) = 0$ ,  $C = C^T$ , which affects the self-representation ability of  $C$  to some extent. In addition, since both  $C$  and  $E$  in the optimization objective involve two terms, which are not conducive to solving, we consider introducing two auxiliary variables  $D$  and  $F$  to dissociate variables and optimize independent sub-problems. Set  $C = D$ ,  $E = F$ , the optimization issue of Eq. 6 is changed into:

$$\begin{aligned} \min_{C,E,\omega,D,F} & \frac{1}{2} \text{Tr} (E - 2EC + C^T EC) + \alpha \|D\|_{\boxed{k}} \\ & + \beta \sum_{h=1}^r \omega_i \|E^i - E\|_F^2 + \lambda \|F\|_{w,r}^p \\ \text{s.t. } & C \geq 0, \quad \text{diag}(C) = 0, \quad C = C^T, \\ & C = D, \quad E = F \end{aligned} \quad (7)$$

The introduction of two auxiliary matrices makes the sub-problems of  $C$  and  $E$  extremely convex, so that a unique and stable solution is obtained for all relevant variables. Then, the augmented Lagrangian model of Eq.7 can be presented

as formula 8:

$$\begin{aligned} L(C, E, \omega, D, F) & = \frac{1}{2} \text{Tr} (E - 2EC + C^T EC) \\ & + \alpha \|D\|_{\boxed{k}} + \beta \sum_{h=1}^r \omega_i \|E^i - E\|_F^2 + \lambda \|F\|_{w,r}^p \\ & + \frac{\mu}{2} (\|D - C + L_1/\mu\|_F^2 + \|F - E + L_2/\mu\|_F^2) \\ \text{s.t. } & C \geq 0, \quad \text{diag}(C) = 0, \quad C = C^T \end{aligned} \quad (8)$$

where  $\mu > 0$  is a balance parameter,  $L_1$  and  $L_2$  are the Lagrangian multipliers in the model, respectively.

The alternating direction method for multipliers (ADMM) is used to iteratively optimize the above optimization model to obtain the final closed-form solution. The next step is to update in turn and fixing the other variables.

#### (1) Update $D$ :

By fixing  $F = F^t$ ,  $\omega = \omega^t$ ,  $C = C^t$ ,  $E = E^t$ , can be optimized by the following formula:

$$\begin{aligned} D^{t+1} & = \arg \min_D \alpha \text{Tr}(LS) + \frac{\mu}{2} \|D - C + L_1/\mu\|_F^2 \\ \text{s.t. } & D \geq 0, \quad \text{diga}(D) = 0, \quad D = D^T \end{aligned} \quad (9)$$

Set  $M = C - L_1/\mu - \frac{\alpha}{\mu} (\text{diag}(S)\mathbf{1}^T - S)$ ,  $\tilde{M} = M - \text{Diag}(\text{diag}(M))$ , then the optimal solution of  $D$  can be found as:

$$D^{t+1} = \max \left( 0, \frac{\tilde{M} + \tilde{M}^T}{2} \right) \quad (10)$$

#### (2) Update $F$ :

By fixing  $D = D^t$ ,  $\omega = \omega^t$ ,  $C = C^t$ ,  $E = E^t$ , for matrix  $F$ , there is the following relationship:

$$\min_F \lambda \|F\|_{w,r}^p + \frac{\mu}{2} \|F - E + L_2/\mu\|_F^2 \quad (11)$$

Among them, define  $Q = E - L_2/\mu$ , the singular value decomposition of  $Q$  is  $Q = U\Sigma V^T$ . Its optimal solution is:

$$F^{t+1} = U \max\{\Sigma - \eta \text{diag}(g), 0\} V^T \quad (12)$$

#### (3) Update $w$ :

Correlation entropy is highly robust to impulse and non-Gaussian noise. Extends CIM (Correlation Inducement Metrics) to general similarity measures between arbitrary kernel matrices or square matrices. The fresh strong metric is named Matrix Correlation Induction Metrics (MCIM), and it can be expressed as equation 13:

$$\text{MCIM}(U, V) = \sqrt{\left( k_\delta(0) - \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M k_\delta(\Delta_{ij}) \right)} \quad (13)$$

Among them  $\Delta = U - V, k_\delta(\Delta_{ij}) = \exp(-(\Delta_{ij})^2/2\delta^2)$ ,  $\delta^2 = (2M^2\|\Delta\|_F^2)^{-1}$ , the closed-form solution of  $w$



can be obtained as:

$$\omega_i^{t+1} = 1 - \text{MCIM}(\mathbf{E}^i, \mathbf{E}) \quad (14)$$

**(4) Update C:**

By fixing  $\mathbf{D} = \mathbf{D}^t, \mathbf{F} = \mathbf{F}^t, \omega = \omega^t, \mathbf{E} = \mathbf{E}^t$ , then the following sub-problems can be optimized to update the variables:

$$\min_C \frac{1}{2} \text{Tr}(\mathbf{E} - 2\mathbf{E}\mathbf{C} + \mathbf{C}^T\mathbf{E}\mathbf{C}) + \frac{\mu}{2} \left\| \mathbf{D} - \mathbf{C} + \frac{L_1}{\mu} \right\|_F^2 \quad (15)$$

Let  $\frac{\partial L(\mathbf{C})}{\partial \mathbf{C}} = 0$ , the optimal solution of  $\mathbf{C}$  can be obtained by further optimizing the solution as follows:

$$\mathbf{C}^{t+1} = (\mu\mathbf{I} + \mathbf{E}) \setminus (\mathbf{E} + \mu\mathbf{D} + L_1) \quad (16)$$

**(5) Update E:**

By fixing  $\mathbf{D} = \mathbf{D}^t, \mathbf{F} = \mathbf{F}^t, \omega = \omega^t, \mathbf{C} = \mathbf{C}^t$ , the optimal solution of  $\mathbf{E}$  can be obtained by further optimizing the solution as follows:

$$\min_E \frac{1}{2} \text{Tr}(\mathbf{E} - 2\mathbf{E}\mathbf{C} + \mathbf{C}^T\mathbf{E}\mathbf{C}) + \beta \sum_{h=1}^r \omega_i \left\| \mathbf{E}^i - \mathbf{E} \right\|_F^2 + \frac{\mu}{2} \left\| \mathbf{F} - \mathbf{E} + L_2/\mu \right\|_F^2 \quad (17)$$

The optimal solution can be made  $\frac{\partial L(\mathbf{E})}{\partial \mathbf{E}} = 0$  to obtain:

$$\mathbf{E}^{t+1} = \frac{2\beta \sum_{k=1}^h \omega_i \mathbf{E}^i + \mu\mathbf{F} + L_2 - \frac{1}{2} + \mathbf{C}^T - \frac{\mathbf{C}\mathbf{C}^T}{2}}{\mu + 2\beta \sum_{k=1}^h \omega_i} \quad (18)$$

The pseudo-code of WTSPSC is shown in Algorithm 1 to update the corresponding value of the relevant variable, and stop the update when the count of iterations run surpass the maximum or the convergence condition is reached. For each iterative calculation, use Eq 19 to determine whether the stopping criterion is met:

$$\begin{aligned} \text{diff}_D &= \left\| \mathbf{D}^{t+1} - \mathbf{D}^t \right\|_F, \quad \text{diff}_F = \left\| \mathbf{F}^{t+1} - \mathbf{F}^t \right\|_F \\ \text{diff}_C &= \left\| \mathbf{C}^{t+1} - \mathbf{C}^t \right\|_F, \quad \text{diff}_E = \left\| \mathbf{E}^{t+1} - \mathbf{E}^t \right\|_F \\ \text{diff}_L &= L^{t+1} - L^t, \\ \max(\text{diff}_C, \text{diff}_D, \text{diff}_F, \text{diff}_E) &\leq \varepsilon \end{aligned} \quad (19)$$

where the value of the optimization model at iteration  $t$  is  $L^t$ , and  $\varepsilon$  refers to the stopping threshold.

After solving to the auxiliary matrix  $\mathbf{D}$ , construct  $\mathbf{A} = (|\mathbf{D}| + |\mathbf{D}^T|)/2$  using the uniform construction form, and then perform spectral clustering on  $\mathbf{A}$  to obtain the final data clustering results.

**IV. EXPERIMENTAL DATA VERIFICATION AND THEORETICAL ANALYSIS**

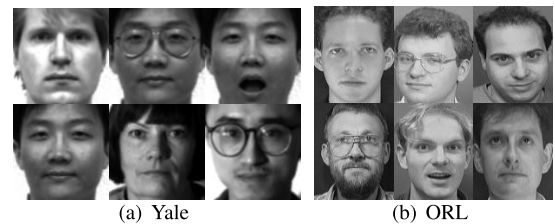
Next section, the proposed WTSPSC is experimentally validated. The face image clustering experiments were conducted on two public benchmark datasets, and segmentation experiments are performed on the BSDS500 natural image dataset. The experiments were conducted with MATLAB R2020b on

**Algorithm 1** Solving WTSPSC via ADMM

**Input:**  $\mathbf{Y}$ , kernel matrix  $\{\mathbf{E}^i\}_{i=1}^r$ , parameter  $\alpha, \beta, \lambda, \mu$ , numbers of clusters  $k$ .

- 1: Initialize:  $\mathbf{S}^1 = 0, \mathbf{C}^1 = 0, \mathbf{E}^1 = \frac{1}{r} \sum_{i=1}^r \mathbf{E}^i, \{\omega_i^1\}_{i=1}^r = \frac{1}{r}, L^1 = 0, L^2 = 0, \varepsilon = 10^{-4}, t = 1, t > \max \text{Iter} = 100$ .
- 2: **while** the algorithm does not converged and  $t < \max \text{Iter}$ , perform the following steps **do**
- 3:     **for**  $t = 1$  to  $\max \text{Iter}$  **do**
- 4:         Assuming  $\mathbf{F}^t, \omega^t, \mathbf{C}^t, \mathbf{E}^t$  is known, solve  $\mathbf{D}^{t+1}$
- 5:         according to (10);
- 6:         Assuming  $\mathbf{D}^{t+1}, \omega^t, \mathbf{C}^t, \mathbf{E}^t$  is known, solve  $\mathbf{F}^{t+1}$  according to (12);
- 7:         Assuming  $\mathbf{D}^{t+1}, \mathbf{F}^{t+1}, \mathbf{C}^t, \mathbf{E}^t$  is known, solve  $\omega^{t+1}$  according to (14);
- 8:         Assuming  $\mathbf{D}^{t+1}, \mathbf{F}^{t+1}, \omega^{t+1}, \mathbf{E}^t$  is known, solve  $\mathbf{C}^{t+1}$  according to (16);
- 9:         Assuming  $\mathbf{D}^{t+1}, \mathbf{F}^{t+1}, \omega^{t+1}, \mathbf{C}^{t+1}$  is known, solve  $\mathbf{E}^{t+1}$  according to (18);
- 10:     **end for**
- 11:     Update Lagrangian parameters  $L_1$  and  $L_2$ ;
- 12:     Update  $t = t + 1$ ;
- 13:     Check the convergence conditions according to (19);
- 14: **end while**
- 15: End the algorithm until the algorithm converges or  $t > \max \text{Iter}$ .

**Output:** The learning results.



**FIGURE 3.** Some samples of face dataset.

a computer with Intel Core i5-11400 CPU (2.70GHz) and 16GB memory.

**A. CLUSTERING EXPERIMENT OF WTSPSC**

To verify the performance of the WTSPSC algorithm, multi-kernel clustering experiments will be performed on the public datasets Yale and ORL. Nine clustering algorithms are compared, namely LRR [33], MKKM [38], AASC [39], RMKKM [24], SCMK [26], LKGr [27], SMKl [25], JMKSC [40] and LRMKC [41]. The parameter settings of the comparison method refer to the original paper.

- **Yale dataset<sup>1</sup>:** The dataset contains 15 test persons with a total of 165 grayscale face images. Grayscale images of 11 different facial expressions or external environments, such as blinking, happy, normal, left glare,

<sup>1</sup><http://www.vision.caltech.edu/archive.html>

TABLE 1. Details of Kernel Functions.

| Kernel     | Expressions  | #n | Parameter setting  |
|------------|--|----|--|
| Gaussian   | $k(x_1, x_2) = \exp\left(-\ x_1 - x_2\ _2^2 / (2t\sigma^2)\right)$ | 7  | $\sigma$ is the maximum distance between $x_1$ and $x_2$ |
| Linear     | $k(x_1, x_2) = x_1^T x_2$  | 1  | -  |
| Polynomial | $k(x_1, x_2) = (x_1^T x_2 + a)^b$                                  | 4  | $a \in \{0, 1\}, b \in \{2, 4\}$                         |



FIGURE 4. Face image clustering.

without glasses, etc. Fig. 3(a) is a partial sample of the Yale dataset.

- **ORL dataset<sup>2</sup>**: The dataset was expanded to reach 40 test subjects, each with 10 different pictures. These images were taken at a certain time of day, with different states of expression and under different intensity lighting conditions. Fig. 3(b) is a partial sample of the ORL dataset.

The above two face datasets are mainly used to assess the robustness of multiple methods and image clustering property of different models. In all comparison algorithms in the experiment, the count of clustering is based on the true count of standard categories, the clustering experiments were performed 20 times without difference and the average of the run results was taken as the final clustering result. This paper shares 12 kernels, and the minutiae of the kernel functions are shown in Table 1.

In this experiments, the image clustering property of all algorithms is evaluated by three manifold evaluation metrics [42], including Purity, Normalized Mutual Information (NMI) and Clustering Accuracy (ACC). The three evaluation values after the run are sorted in descending order, and the algorithm with the highest value has the best properties, where ACC is the calculation of the optimal matching arrangement between the clustered labels and the standard labels; NMI is mainly a weigh of the resemblance among two clustering results; For purity, it is measured by calculating percentage of the count of rightly assigned clusters to the overall count of data samples. Face image clustering is to perform orderly clustering of multiple disordered images through the feature information of faces, as shown in Fig. 4.

<sup>2</sup><http://mlg.ucd.ie/datasets/3sources.html>

By setting a series of parameters and combining the evaluation metrics, the experimental results can be obtained on the Yale and ORL face image datasets as shown in Table 2.

It can be seen that a proper kernel weighting strategy in multi-kernel learning can improve the clustering performance, as expected, the proposed WTSPSC method in this paper obtains the best clustering property among all the contrasting methods, which illustrates the effectiveness of the multi-kernel weighting strategy in subspace clustering. Compared with the previously proposed methods such as SMKL, JMKSC, LRMKC, etc., the WTSPSC method produces dramatically different results. Theoretically, the largest difference is that WTSPSC utilizes a low-rank constraint based on weighted truncated Schatten-p norm, and secondly it uses an entropy-based weighting strategy, which improves the previous simple measure of euclidean distance. In the multi-kernel weighting strategy, kernels with high correlation are weighted more (with an upper limit of 1), while insignificant kernels are approached to 0. In addition, BDR is incorporated into the WTSPSC model to promote the similarity matrix to have a block-diagonal architecture. In the MKL framework, WTSPSC can obtain the lowest weight standard deviation, resulting in good robustness. Finally, the experimental effects on the Yale and ORL face image datasets display that compared with existing clustering methods, WTSPSC has better clustering performance on three evaluation metrics.

**B. SEGMENTATION EXPERIMENT OF AMR-WT-WTSPSC**

In this subsection, Fig. 5 clearly shows the block diagram of the image segmentation process based on AMR-WT-WTSPSC. Firstly, the original image is pre-segmented, and the superpixel image is obtained by adaptive morphological watershed transformation. Secondly, the feature information is extraceted from the independent superpixel

TABLE 2. Experimental results of Yale and ORL dataset.

| Dataset | Metrics | LRR   | MKKM  | AASC  | RMKKM | SCMK  | LKGr  | SMKL  | JMKSC | LRMKC | WTSPSC       |
|---------|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------------|
| Yale    | ACC     | 0.470 | 0.456 | 0.406 | 0.523 | 0.527 | 0.542 | 0.583 | 0.627 | 0.670 | <b>0.685</b> |
|         | NMI     | 0.526 | 0.501 | 0.468 | 0.554 | 0.573 | 0.563 | 0.616 | 0.630 | 0.658 | <b>0.680</b> |
|         | Purity  | 0.543 | 0.473 | 0.423 | 0.537 | 0.613 | 0.554 | 0.664 | 0.681 | 0.701 | <b>0.715</b> |
| ORL     | ACC     | 0.663 | 0.474 | 0.274 | 0.553 | 0.654 | 0.617 | 0.575 | 0.723 | 0.719 | <b>0.728</b> |
|         | NMI     | 0.812 | 0.687 | 0.437 | 0.747 | 0.808 | 0.792 | 0.733 | 0.850 | 0.854 | <b>0.858</b> |
|         | Purity  | 0.698 | 0.513 | 0.314 | 0.601 | 0.699 | 0.659 | 0.648 | 0.751 | 0.763 | <b>0.795</b> |

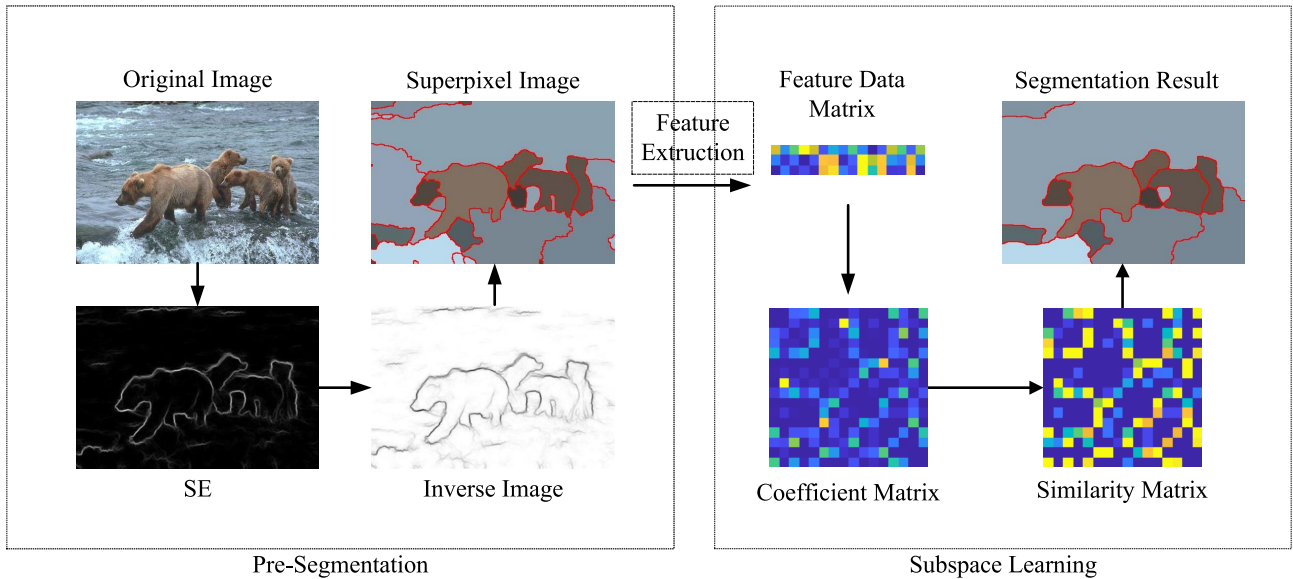


FIGURE 5. The flowchart of the AMR-WT-WTSPSC.

block. The feature data matrix is formed and input into WTSPSC to learning coefficient matrix and similarity matrix. Then perform spectral clustering operation on the obtained similarity matrix. Finally, receive the clustering and segmentation result of the image. So as to confirm the segmentation capability of the AMR-WT-WTSPSC algorithm, we use two superpixel segmentation methods [43], [44] and the other three novel subspace learning methods to form seven comparison algorithms, namely SLIC-LRR, AMR-WT-LRR, SLIC-JMKSC, AMR-WT-JMKSC, SLIC-LRMKC, AMR-WT-LRMKC, and SLIC-WTSPSC. Then, perform image segmentation experiments with the AMR-WT-WTSPSC method on the BSDS500 dataset [45]. The dataset contains a total of 500 natural images, such as: people, animals, plants, buildings and landscapes, etc. The size of each image is  $321 \times 481$ , and they are all color images. Each image has a corresponding segmentation standard map (Ground Truth), which is calibrated by multiple experts according to corresponding rules, and the segmentation effects of all algorithms are shown in Fig.6.

By observing the experimental performances, it can be found that when using the same superpixel segmentation algorithm, the impact on the natural image segmentation

capability mainly relies on the clustering ability of the subspace learning method on the feature data matrix. After the feature data matrix is obtained from the superpixel block, different ways of subspace learning will produce different subspace representations, which can lead to significantly varying performances, while the block diagonal regularizer introduced by WTSPSC can obtain a better block diagonal structure, and the group effect of image feature data is better. For the first line of segmentation display, in contrast, ours divides the foreground trees into two regions, and the contrast algorithm divides the foreground into three regions. The mid-range area comparison algorithm does not adequately segment the contours of the mountains, and divides the two mountains into a whole area, while the AMR-WT-WTSPSC method clearly segments the contour edges of the mountains. When using the same subspace learning method, it can be found that the SLIC superpixel segmentation method has poor segmentation performance, because SLIC is always segmented into rectangular areas of similar shape or size, while AMR-WT more focus on the edge detail information of objects in the image, not just focus on the number of superpixel blocks, which makes up for the defect that the superpixel blocks contain both foreground objects and background patterns, making the obtained image feature data more



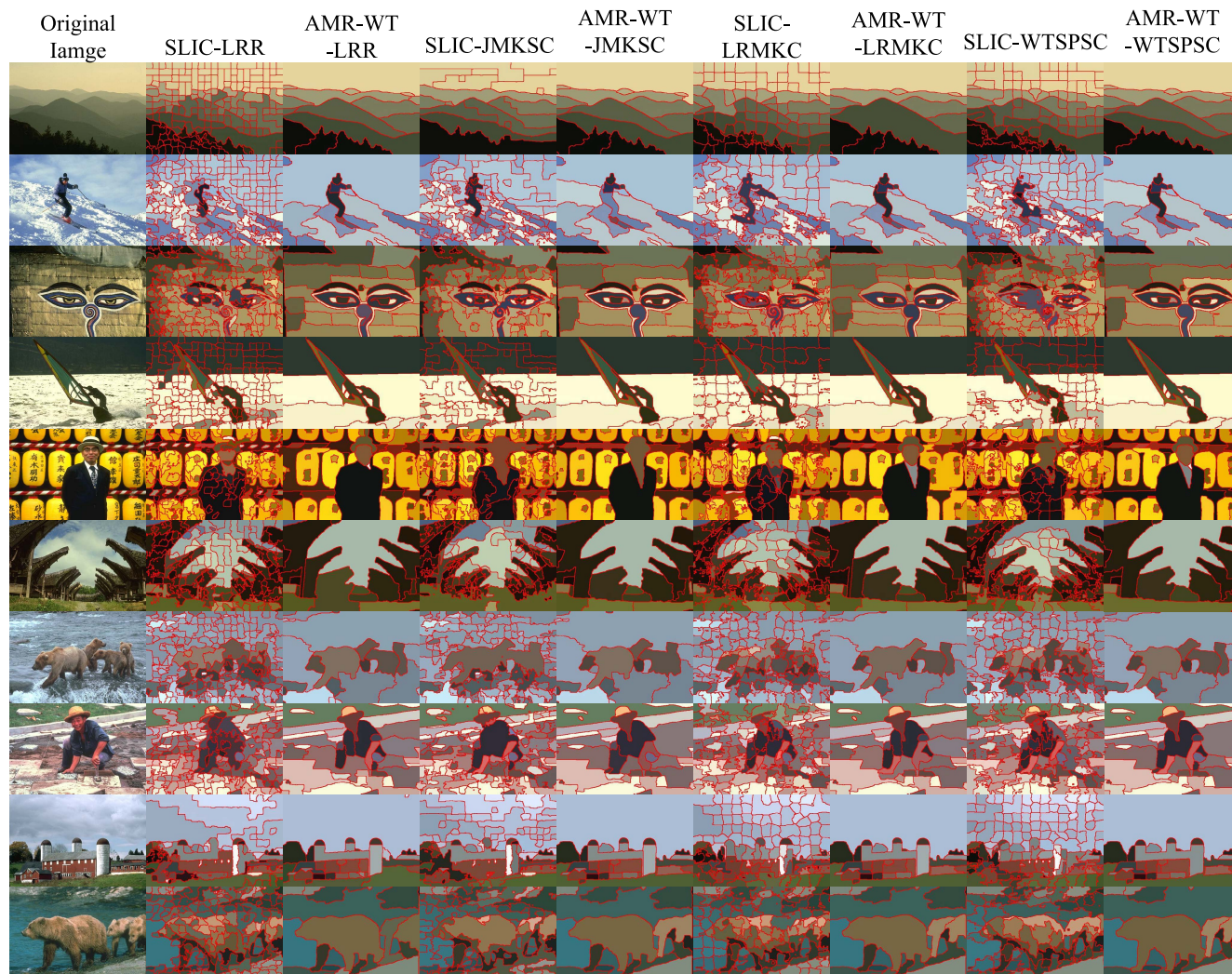


FIGURE 6. Experimental results of each method.

TABLE 3. Experimental results of Yale and ORL dataset.

| Metrics | SLIC-LRR | AMR-WT-LRR | SLIC-JMKSC | AMR-WT-JMKSC | SLIC-LRMKC | AMR-WT-LRMKC | SLIC-WTSPC | AMR-WT-WTSPSC |
|---------|----------|------------|------------|--------------|------------|--------------|------------|---------------|
| PRI     | 0.651    | 0.759      | 0.691      | 0.776        | 0.667      | 0.770        | 0.677      | <b>0.810</b>  |
| VOI     | 4.364    | 2.480      | 3.591      | 2.210        | 4.805      | 2.306        | 4.295      | <b>1.970</b>  |

accurate. It can accurately reflect the essential information of the image, which is beneficial for subsequent segmentation tasks.

In order to evaluate the image segmentation performance of each segmentation method from a quantitative perspective [46], two popular evaluation metrics: Probabilistic Rand Index (PRI) [47] and Rate of Change of Information (VOI) [48] are used to reflect the experimental performances, as shown in Table 3. When the PRI value is the highest and the VOI value is the lowest, the segmentation effect is the best at this time. By observing the results in the table, it is not difficult to find that for the same pre-segment method, the

segmentation performance depends on the subspace learning model. When clustering and segmenting image feature data, the AMR-WT-WTSPSC method uses the Schatten-p norm with weighted truncated to constrain the data with low rank, and receives the best result compared with other algorithms. At the same time, the superpixel segmentation method has an important impact on the segmentation task of the subsequent subspace learning. The more the superpixel block conforms to the local information of the image, the more conducive to the subsequent segmentation. It can be seen from the evaluation indicators that ours method obtains the best segmentation performance.



## V. CONCLUSION

In this work, a novel image segmentation algorithm with weighted truncated Schatten-p multi-kernel subspace learning is proposed. First, the adaptive morphological watershed transform is used as a superpixel segmentation method to pre-segment the original image. Second, a weighted truncated Schatten-p subspace learning model with low-rank multi-kernel is designed, and the correlation entropy measure is introduced into the weighting strategy. The optimal consensus kernel can be learned when processing data. Besides, applying low-rank constraints to the learned optimal kernel to induce it to fit the image feature data effectively, which can handle nonlinear structure in image feature data with good robustness to impulse noise and non-Gaussian noise. At the same time, applying block diagonal constraints on the coefficient matrix, so that a tighter block-diagonal structure can be obtained, which is beneficial to the correct classification of data. The clustering and segmentation experiments demonstrate that the proposed algorithm receives the best data clustering and image segmentation performances. In future research works, we will study the multi-dimensional feature information of images and consider image segmentation with multi-view subspace clustering.

## REFERENCES

- [1] Y. Cheng and B. Li, "Image segmentation technology and its application in digital image processing," in *Proc. IEEE Asia-Pacific Conf. Image Process., Electron. Comput. (IPEC)*, Apr. 2021, pp. 1174–1177.
- [2] M. Zhang, L. Jiao, R. Shang, X. Zhang, and L. Li, "Unsupervised EA-based fuzzy clustering for image segmentation," *IEEE Access*, vol. 8, pp. 8627–8647, 2020.
- [3] X. Jia, T. Lei, P. Liu, D. Xue, H. Meng, and A. K. Nandi, "Fast and automatic image segmentation using superpixel-based graph clustering," *IEEE Access*, vol. 8, pp. 211526–211539, 2020.
- [4] H. Zhu, R. Vial, S. Lu, X. Peng, H. Fu, Y. Tian, and X. Cao, "YouTube: Searching action proposal via recurrent and static regression networks," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2609–2622, Jun. 2018.
- [5] Z. Wang and J. Liu, "Multiple kernel subspace clustering based on consensus Hilbert space and second-order neighbors," *IEEE Access*, vol. 8, pp. 124633–124645, 2020.
- [6] Y. Zheng, X. Zhang, Y. Xu, M. Qin, Z. Ren, and X. Xue, "Robust multi-view subspace clustering via weighted multi-kernel learning and co-regularization," *IEEE Access*, vol. 8, pp. 113030–113041, 2020.
- [7] Z. Xie and L. Wang, "Active block diagonal subspace clustering," *IEEE Access*, vol. 9, pp. 83976–83992, 2021.
- [8] J. Francis, A. Johnson, B. Madathil, and S. N. George, "A joint sparse and correlation induced subspace clustering method for segmentation of natural images," in *Proc. IEEE 17th India Council Int. Conf. (INDICON)*, Dec. 2020, pp. 1–7.
- [9] J. Francis, M. Baburaj, and S. N. George, "An  $l_{1/2}$  and graph regularized subspace clustering method for robust image segmentation," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 18, no. 2, pp. 1–24, May 2022.
- [10] X. Xue, X. Wang, X. Zhang, J. Wang, and Z. Liu, "Image segmentation based on non-convex low rank multi kernel clustering," in *Proc. CAAI Int. Conf. Artif. Intell.*, Cham, Switzerland: Springer, 2021, pp. 420–431.
- [11] W. Zhu, J. Lu, and J. Zhou, "Nonlinear subspace clustering for image clustering," *Pattern Recognit. Lett.*, vol. 107, pp. 131–136, May 2018.
- [12] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, Nov. 2013.
- [13] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *Proc. ICML*, vol. 1. Princeton, NJ, USA: Citeseer, 2010, p. 8.
- [14] L. Zhuang, H. Gao, Z. Lin, Y. Ma, X. Zhang, and N. Yu, "Non-negative low rank and sparse graph for semi-supervised learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2328–2335.
- [15] H. Hu, Z. Lin, J. Feng, and J. Zhou, "Smooth representation clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 3834–3841.
- [16] Y.-X. Wang, H. Xu, and C. Leng, "Provable subspace clustering: When LRR meets SSC," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 26, 2013, pp. 1–9.
- [17] M. Z. Nashed and G. G. Walter, "General sampling theorems for functions in reproducing kernel Hilbert spaces," *Math. Control, Signals, Syst.*, vol. 4, no. 4, pp. 363–390, Dec. 1991.
- [18] A. Ng, M. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 14, 2001, pp. 849–856.
- [19] I. S. Dhillon, Y. Guan, and B. Kulis, "Kernel k-means: Spectral clustering and normalized cuts," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2004, pp. 551–556.
- [20] J. Huang, F. Nie, and H. Huang, "A new simplex sparse learning model to measure data similarity for clustering," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, 2015, pp. 3569–3575.
- [21] L. Tao, W. Wei-Wei, Z. Dong, and J. Xi-Xi, "Weighted-sparse subspace clustering method for image segmentation," *Syst. Eng. Electron.*, vol. 36, no. 3, pp. 580–585, 2014.
- [22] X. P. Li, W. W. Wang, L. Luo, and S. Q. Wang, "Improved sparse subspace clustering method for image segmentation," *Syst. Eng. Electron.*, vol. 37, no. 10, pp. 2418–2424, 2010.
- [23] Z. Xu, R. Jin, I. King, and M. Lyu, "An extended level method for efficient multiple kernel learning," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 21, 2008, pp. 1825–1832.
- [24] L. Du, P. Zhou, L. Shi, H. Wang, M. Fan, W. Wang, and Y.-D. Shen, "Robust multiple kernel k-means using  $\ell_{2,1}$ -norm," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, 2015, pp. 1–7.
- [25] Z. Kang, X. Lu, J. Yi, and Z. Xu, "Self-weighted multiple kernel learning for graph-based clustering and semi-supervised classification," 2018, *arXiv:1806.07697*.
- [26] Z. Kang, C. Peng, Q. Cheng, and Z. Xu, "Unified spectral clustering with optimal graph," in *Proc. AAAI Conf. Artif. Intell.*, vol. 32, no. 1, 2018, pp. 3366–3373.
- [27] Z. Kang, L. Wen, W. Chen, and Z. Xu, "Low-rank kernel learning for graph-based clustering," *Knowl.-Based Syst.*, vol. 163, pp. 510–517, Jan. 2019.
- [28] H. Fan, F. Meng, Y. Liu, F. Kong, J. Ma, and Z. Lv, "A novel breast ultrasound image automated segmentation algorithm based on seeded region growing integrating gradual equipartition threshold," *Multimedia Tools Appl.*, vol. 78, no. 19, pp. 27915–27932, Oct. 2019.
- [29] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, Nov. 2006.
- [30] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient ND image segmentation," *Int. J. Comput. Vis.*, vol. 70, no. 2, pp. 109–131, Nov. 2006.
- [31] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 6, pp. 583–598, Jun. 1991.
- [32] T. Lei, X. Jia, T. Liu, S. Liu, H. Meng, and A. K. Nandi, "Adaptive morphological reconstruction for seeded image segmentation," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5510–5523, Nov. 2019.
- [33] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.
- [34] V. M. Patel and R. Vidal, "Kernel sparse subspace clustering," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 2849–2853.
- [35] P. Ji, I. Reid, R. Garg, H. Li, and M. Salzmann, "Adaptive low-rank kernel subspace clustering," 2017, *arXiv:1707.04974*.
- [36] C. Lu, J. Feng, Z. Lin, T. Mei, and S. Yan, "Subspace clustering by block diagonal representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 2, pp. 487–501, Feb. 2019.
- [37] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5286–5298, Nov. 2007.
- [38] H. C. Huang, Y. Y. Chuang, and C. S. Chen, "Multiple kernel fuzzy clustering," *IEEE Trans. Fuzzy Syst.*, vol. 20, no. 1, pp. 120–134, Feb. 2012.
- [39] H.-C. Huang, Y.-Y. Chuang, and C.-S. Chen, "Affinity aggregation for spectral clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 773–780.
- [40] C. Yang, Z. Ren, Q. Sun, M. Wu, M. Yin, and Y. Sun, "Joint correntropy metric weighting and block diagonal regularizer for robust multiple kernel subspace clustering," *Inf. Sci.*, vol. 500, pp. 48–66, Oct. 2019.

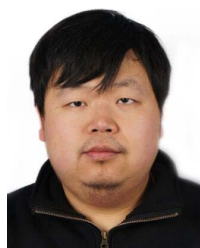
- [41] X. Zhang, X. Xue, H. Sun, Z. Liu, L. Guo, and X. Guo, "Robust multiple kernel subspace clustering with block diagonal representation and low-rank consensus kernel," *Knowl.-Based Syst.*, vol. 227, Sep. 2021, Art. no. 107243.
- [42] W. Xu, X. Liu, and Y. Gong, "Document clustering based on non-negative matrix factorization," in *Proc. 26th Annu. Int. ACM SIGIR Conf. Res. Develop. Informaion Retr. (SIGIR)*, 2003, pp. 267–273.
- [43] R. J. Al-Azawi, Q. S. Al-Jubouri, and Y. A. Mohammed, "Enhanced algorithm of superpixel segmentation using simple linear iterative clustering," in *Proc. 12th Int. Conf. Develop. eSyst. Eng. (DeSE)*, Oct. 2019, pp. 160–163.
- [44] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012, doi: 10.1109/TPAMI.2012.120.
- [45] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2, Feb. 2001, pp. 416–423.
- [46] X. Li, W. Wang, A. Razi, and T. Li, "Nonconvex low-rank sparse factorization for image segmentation," in *Proc. 11th Int. Conf. Comput. Intell. Secur. (CIS)*, Dec. 2015, pp. 227–230.
- [47] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "Toward objective evaluation of image segmentation algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 929–944, Jun. 2007.
- [48] M. Meilă, "Comparing clusterings—An information based distance," *J. Multivariate Anal.*, vol. 98, no. 5, pp. 873–895, May 2007.



**LEI PU** received the B.S. degree from the Southwest University of Science and Technology, Mianyang, China, in 2020. He is currently pursuing the M.S. degree in electronic information engineering. His research interests include deep learning, pattern recognition, and medical image processing.



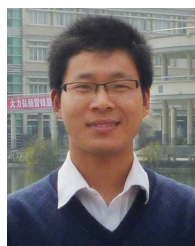
**JING WANG** received the B.A. degree in process equipment and control engineering from the Tianjin University of Technology (TUT), Tianjin, China, in 2019. She is currently pursuing the M.S. degree in control science and engineering with the Southwest University of Science and Technology (SWUST), Mianyang, China. Her research interests include subspace clustering, sparse representation, and multi-view clustering and their applications in image processing.



**LI LI** was born in 1981. He received the Ph.D. degree in mechanical and electronic engineering from Sichuan University (SCU), China, in 2013. He is currently an Assistant Professor with the School of Computer Science and Technology, Southwest University of Science and Technology, and is leading a research group in doing some edge computing and CPS designing. His research interests include edge computing, CNN, droid navigation, and machine vision.



**XIAO WANG** received the B.S. degree in communication engineering from the North China Institute of Aerospace Engineering (NCIAE), Langfang, China, in 2020. He is currently pursuing the M.S. degree in electrical and information engineering with the Southwest University of Science and Technology (SWUST), Mianyang, China. His research interests include subspace clustering, image segmentation, and image multi-feature.



**XIAOQIAN ZHANG** (Member, IEEE) received the M.S. degree in circuits and systems from the Southwest University of Science and Technology (SWUST), Mianyang, China, in 2013, and the Ph.D. degree in control science and engineering from the Nanjing University of Science and Technology, Nanjing, China. He is also working as a Teacher at SWUST. His research interests include subspace clustering, sparse representation, and deep learning and their applications in image processing.

...