

## Research Article

# Mixed Region Covariance Discriminative Learning for Image Classification on Riemannian Manifolds

Xi Liu <sup>1</sup>, Zhengming Ma <sup>1,2</sup> and Guo Niu <sup>3</sup>

<sup>1</sup>School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China

<sup>2</sup>Nanfang College, Sun Yat-sen University, Guangzhou 510275, China

<sup>3</sup>School of Electronic Information Engineering, Foshan University, Foshan 528000, China

Correspondence should be addressed to Zhengming Ma; [issmzm@mail.sysu.edu.cn](mailto:issmzm@mail.sysu.edu.cn)

Received 25 July 2018; Revised 18 December 2018; Accepted 21 January 2019; Published 28 February 2019

Academic Editor: Łukasz Jankowski

Copyright © 2019 Xi Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Covariance matrices, known as symmetric positive definite (SPD) matrices, are usually regarded as points lying on Riemannian manifolds. We describe a new covariance descriptor, which could improve the discriminative learning ability of region covariance descriptor by taking into account the mean of feature vectors. Due to the specific geometry of Riemannian manifolds, classical learning methods cannot be directly used on it. In this paper, we propose a subspace projection framework for the classification task on Riemannian manifolds and give the mathematical derivation for it. It is different from the common technique used for Riemannian manifolds, which is to explicitly project the points from a Riemannian manifold onto Euclidean space based upon a linear hypothesis. Under the proposed framework, we define a Gaussian Radial Basis Function- (RBF-) based kernel with a Log-Euclidean Riemannian Metric (LERM) to embed a Riemannian manifold into a high-dimensional Reproducing Kernel Hilbert Space (RKHS) and then project it onto a subspace of the RKHS. Finally, a variant of Linear Discriminative Analyze (LDA) is recast onto the subspace. Experiments demonstrate the considerable effectiveness of the mixed region covariance descriptor and the proposed method.

## 1. Introduction

Image classification is an important and prevalent topic in pattern recognition and computer vision research. Particularly when the object detecting [1] and visual tracking [2, 3] are widely developed, these technologies have been effectively applied to numerous real world scenarios [4–8].

Feature representation is one of the fundamental steps for the detection, tracking, and classification task. The raw pixel value is the simplest choice. This has been used in many traditional classification methods such as Nonparametric Discriminant Multimanifold Learning (NDML) [9] and Discriminant Analysis with Local Gaussian Similarity Preserving (DA-LGSP) [10]. However, the descriptor is usually limited by the high-dimensional representation, and it is not robust to noise.

A natural extension of raw pixel values is to search for more useful information, such as histogram of image gradients and edges, colors. Using information about primary

colors (Red, Green, and Blue), Marcin et al. calculate values of Hue-Saturation-Brightness (HSB) and build models of input image regions by the HSB in some practical application scenes [11, 12]. However it is nonuniversal to describe objects features only in color space. Histogram-based representations such as Scale-Invariant Feature Transform (SIFT) descriptor [13] and Histogram of Oriented Gradients (HOG) [14] descriptor are more robust and discriminative. Nevertheless, these descriptors are difficult to handle regarding occlusion. Model the object appearance using multiple features is another choice, but it requires a large amount of data [15, 16].

Recently, covariance descriptors [17, 18] have attracted increased attention in computer vision and pattern recognition applications. Tuzel et al. [17] first propose representing image regions by a nonsingular covariance matrix of  $d$ -features such as location, intensity, higher order directional derivatives, and edge orientation. The dimensionality of the covariance descriptor is independent of the image size and has only  $d \times (d + 1)/2$  different values due to symmetry.

Furthermore, the covariance matrix does not have any information regarding the ordering and the number of points, which implies a certain scale and rotation invariance in different images. But, region covariance descriptor neglects the mean of feature vectors, while we think the mean vector could increase discriminative validity to some extent. For this reason, we propose a new region descriptor by considering the mean vector of  $d$ -features.

Covariance matrices lie on Riemannian manifolds, not on Euclidean space. It is not adequate to neglect the nonlinear geometry and directly use most of the conventional machine learning methods. This often yields undesirable effect such as the swelling of diffusion tensors [19]. One of the key issues is how to develop efficient discriminative learning algorithms on Riemannian manifolds. A popular choice is mapping points of a manifold onto the tangent space at a particular point and then designing the classification there [20]. This has been applied to pedestrians detecting [1], action classification [21] and multiclass classification problems [22]. However, flattening a manifold through tangent spaces may generate inaccurate modeling, especially for the regions far away from the tangent pole. Another framework follows the idea of the kernel method in Euclidean space [23]. Kernel-based methods embed the manifold into a high-dimensional Reproducing Kernel Hilbert Spaces (RKHS) and further project the elements in RKHS onto Euclidean space by an explicit mapping. Benefiting from the explicit mapping, many classification methods originally proposed on Euclidean space can be extended to Riemannian manifolds [24–29]. Riemannian Locality Preserving Projections (RLPP) [30] is a more recent kernel-based method that has been proposed for the visual classification task. Modeling images as region covariance matrices, RLPP proposes mapping a Riemannian manifold onto Euclidean space by a Riemannian pseudo kernel, and then exploiting Locality Preserving Projection (LPP) for discriminative learning. Considerable results on visual recognition tasks have been achieved. However, the Riemannian pseudo kernel derived from Affine-Invariance Riemannian Metric (AIRM) [31] is not guaranteed to be positive definite. Besides, RLPP is a supervise method, but we argue that RLPP fails to take full advantage of the labels. Jayasumana et al. [32] present a framework on Riemannian manifolds to identify the positive definiteness of Gaussian RBF kernel and utilize the Log-Euclidean Gaussian kernel to Kernel Principal Component Analysis (KPCA) for the recognition task.

Based on the related works, we propose a discriminative learning method, Mixed Region Covariance Discriminative Learning (MRCDL) for image classification. The main contributions within this paper are as follows:

(1) We propose a novel region descriptor, which is more discriminative than the region covariance descriptor.

(2) To address the classification task, we proposed a subspace projection framework for Riemannian manifolds. The previous framework projects the elements of Riemannian manifolds onto Euclidean space based upon a linear hypothesis, but the final feature space in our framework is a subspace of RKHS, which is isomorphic to Euclidean space. We give the well-founded derivation of the subspace mapping.

(3) We exploit KLDA on the final low-dimensional subspace to adapt to the subspace projection framework.

The rest of the paper is organized as follows: we briefly review the Riemannian geometry and the properties of RKHS in Section 2. Section 3 presents the related works of our proposed method. In Section 4, we present the proposed MRCDL in detail. In Section 5, the experimental results are depicted to demonstrate the effectiveness of MRCDL and mixed region covariance. Conclusions are made in Section 6.

## 2. Preliminaries

In this section, we briefly review the geometry of Riemannian manifolds and the properties of RKHS.

*2.1. Riemannian Geometry on SPD Matrices.* The space of  $d \times d$  real SPD matrices, denoted by  $Sym_d^+$ , is a connected Riemannian manifold. A Riemannian manifold is a differentiable manifold endowed with a symmetric positive definite and smoothly second order tensor field. So-called tensor is a multilinear functional on linear space. The tensor at a given point on a differentiable manifold is the tensor on the tangent space at the point. The tensor fields map the points on the manifold to the tensors on the tangent space at the points. Once we define a symmetric positive definite and smoothly second order tensor field on a differentiable manifold, we define a smoothly varying inner product on each tangent, vice versa. The family of symmetric positive definite and smoothly second order tensor fields is known as the Riemannian metric of the manifold. It enables us to induce the geodesic distance between two points on the manifold. The geodesic distance between two points is the length of the shortest curve connecting the two points. This is the most suitable measure of distance on a Riemannian manifold. In the following, we discuss some popular Riemannian metrics proposed on  $Sym_d^+$ .

The affine-invariant geodesic distance induced by AIRM [31] is the most natural measure of dissimilarity between two points lying on a Riemannian manifolds. The distance of AIRM between  $X_i$  and  $X_j$  is defined as

$$D_{AI}(X_i, X_j) = \left\| \log(X_i^{-1/2} X_j X_i^{-1/2}) \right\|_F \quad (1)$$

where  $\log(\cdot)$  is an ordinary matrix operator which is computed through singular value decomposition, and  $\| \cdot \|_F$  denotes the matrix Frobenius norm.

The AIRM enjoys several useful theoretical properties; however, this metric has the drawback of high time complexity in practice. The Log-Euclidean geodesic distance [33] is the distance between two SPD matrices measured on the tangent space of  $Sym_d^+$  at the identity matrix. The Log-Euclidean geodesic distance between  $X_i$  and  $X_j$  is given by

$$D_{LE}(X_i, X_j) = \left\| \log(X_i) - \log(X_j) \right\|_F \quad (2)$$

This distance is a lower bound on the actual geodesic distance. It is more efficient than the affine-invariant geodesic distance on computation.

The Jensen-Bregman LogDet Divergence (JBLD) [34] can be understood as a proxy of the geodesic distance based on matrix divergence. This divergence is also much faster to compute and turns out to be empirically very effective. The JBLD between  $X_i$  and  $X_j$  is defined as

$$D_{JK}(X_i, X_j) = \left( \log \left| \frac{X_i + X_j}{2} \right| - \frac{1}{2} \log |X_i X_j| \right)^{1/2} \quad (3)$$

where  $|\cdot|$  denotes the determinant of a matrix.

This measure does not require matrix eigenvalue computations or logarithms and at the same time enjoys many of the properties of AIRM. But JBLD is not the true geodesic distances and can lead to poor accuracy.

**2.2. RKHS.** Let  $S(\Omega) = \{f \mid f : \Omega \rightarrow R, \int_{\Omega} |f(x)|^2 < +\infty\}$  be a function space equipped with an inner product  $\langle \bullet, \bullet \rangle$  which could induce a complete inner product space, then  $H = (S(\Omega), \langle \bullet, \bullet \rangle)$  [35] is a Hilbert space of functions. If there exists a function  $k : \Omega \times \Omega \rightarrow R$ ,

- (1) for all  $x \in \Omega, k_x(\bullet) = k(\bullet, x) \in S(\Omega)$
- (2) for all  $x \in \Omega$  and  $f \in S(\Omega), f(x) = \langle f, k(\bullet, x) \rangle$

then  $H$  is called Reproducing Kernel Hilbert Space and  $k(x, y)$  is the reproducing kernel of  $H$ .

Let  $\phi : \Omega \rightarrow H$  be a mapping defined by  $k(x, y)$ , for all  $x \in \Omega, \phi(x) = k(\bullet, x) = k_x \in H$ . It can be induced that

$$\begin{aligned} \langle \phi(x), \phi(y) \rangle &= \langle k_x, k(\bullet, y) \rangle = k_x(y) = k(y, x) \\ &= k(x, y) \end{aligned} \quad (4)$$

Hilbert space can be generated from a kernel function. A function  $k(x, y) : \Omega \times \Omega \rightarrow R$  can be called a kernel function if it satisfies the following:

- (1) Symmetric: for all  $x, y \in \Omega, k(x, y) = k(y, x)$
- (2) Square integrable: for all  $x \in \Omega, k_x = k(\bullet, x)$  is square integrable
- (3) Positive definite: for all positive integers  $n \in N$  and all  $x_1, \dots, x_n \in \Omega$ , the matrix  $K = \begin{bmatrix} k(x_1, x_1) & \dots & k(x_1, x_n) \\ \vdots & \ddots & \vdots \\ k(x_n, x_1) & \dots & k(x_n, x_n) \end{bmatrix}$  is positive definite.

According to Moore–Aronszajn Theorem [35], as long as a kernel function is defined, a unique RKHS as well as its reproducing kernel is also defined. Therefore, a kernel function can be used to represent a RKHS.

The success of many state-of-the-art computer vision algorithms arises from the use of kernel methods on Euclidean spaces. One of the advantages of kernel methods is that it transforms the nonlinear space to a linear space by embedding a manifold into RKHS. In addition, we never need to compute the actual vectors since the inner products in the RKHS can be evaluated by means of the kernel function.

### 3. Related Work

**3.1. Region Covariance Descriptor.** The region covariance descriptor, as a special case of SPD matrices, proposes a natural way of fusing multiple features that might be correlated.

Let  $R$  be the input image or the rectangular region in the input image. For each pixel inside  $R$ , it can be represented by a feature vector including pixel coordinates, intensity or colors, the first and the second order derivatives of intensities in the  $x$  and  $y$  directions, etc. For example, we can extract the features from the pixel located in the  $x$ -th row and the  $y$ -th column in a given image region:

$$z = [x, y, I(x, y), |I_x|, |I_y|] \quad (5)$$

where  $I(x, y)$  is the image intensity of the pixel;  $I_x = |I(x+1, y) - I(x, y)|$  and  $I_y = |I(x, y+1) - I(x, y)|$  are the first order derivatives of intensities in the  $x$  and the  $y$  directions, respectively. Suppose there are  $h$  pixels in height and  $w$  pixels in width in  $R$ . Let  $z_k$  be the  $d$ -dimensional feature vectors of points inside  $R$ , where  $k = 1, \dots, h \times w$ . The region covariance  $\Sigma \in R^{d \times d}$  of  $R$  can be computed by

$$\Sigma = \frac{1}{n-1} \sum_{k=1}^n (z_k - \mu)(z_k - \mu)^T \quad (6)$$

where  $n = h \times w$ ;  $\mu = (1/n) \sum_{k=1}^n z_k \in R^d$  denotes the mean of the points.

**3.2. RLPP.** Next, we briefly introduce the implementation of RLPP [30]. Suppose there are  $N$  training images belonging to  $c$  classes. RLPP characterizes input images as covariance matrices by using the region covariance descriptor. Let  $\chi = \{X_1, X_2, \dots, X_N\} \subseteq M$  denote the  $N$  training samples; each  $X_i$  is a SPD matrix corresponding to an input image region (or image), and let  $l = \{l_1, l_2, \dots, l_N\}$  denote the corresponding labels. As mentioned above, covariance matrices are SPD matrices lying on a Riemannian manifold. Classical learning algorithms cannot be directly utilized on the manifolds. Thus RLPP exploits a Riemannian pseudo kernel to map the  $Sym_d^+$  from  $M$  onto a linear space  $M'$ . The Riemannian pseudo kernel defined in RLPP is formulated as

$$k_{ij} = k(X_i, X_j) = \exp \left\{ -\sigma^{-1} \left\| \log \left( X_i^{-1/2} X_j X_i^{-1/2} \right) \right\|_F^2 \right\} \quad (7)$$

RLPP implements the visual recognition task in an extrinsic manner by confining the solution to be linear. Given a SPD matrix  $X_i$  as the input, its projection  $y_i$  on the feature space is obtained by

$$y_i = AK_i^T \quad (8)$$

where  $A$  is the projection matrix and  $K_i = [k_{i1} \ \dots \ k_{iN}]$ .

The aim of RLPP is to optimize  $A$  by preserving the local geometrical structure of the original space, which can be modeled by a similarity graph  $W$ . The binary graph  $W$  incurs a heavy penalty if points from the same class are mapped far away in the mapped feature space:

$$W_{ij} = \begin{cases} 1 & \text{if } l_i = l_j \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

The optimized mapping is obtained by minimize the following objective function:

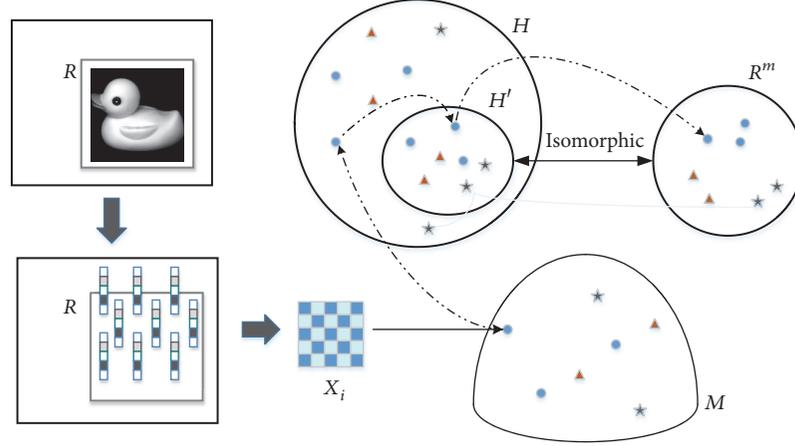


FIGURE 1: Overview of the proposed MRCDL.

$$\begin{aligned}
 \frac{1}{2} \sum_{i,j} (y_i - y_j)^2 W_{ij} &= \sum_i A^T K_i K_i^T A W_{ii} \\
 &\quad - \sum_{i,j} A^T K_i K_i^T A W_{ij} \\
 &= KDK^T A - A^T KWK^T A \\
 &= A^T K L K^T A
 \end{aligned} \tag{10}$$

Therefore, the minimization problem reduces to finding:

$$\begin{aligned}
 \arg \min_A \quad & A^T K L K^T A \\
 \text{s.t.} \quad & A^T K D K^T A = 1
 \end{aligned} \tag{11}$$

The optimal  $A$  is given by the generalized eigenvectors corresponding to the  $m$  smallest nonzero eigenvalues of  $K L K^T A = \lambda K D K^T A$ . Then, the recognition task over Riemannian manifolds is reduced to classification problem in vector space. The  $N$  training samples can be projected to  $M'$  and explicitly represented by a series of  $m$ -dimensional vectors through (7). As such, for any given testing image, the corresponding covariance matrix  $X_i$  and the kernel vector  $K_i = [k_{i1} \cdots k_{iN}]$  are firstly compute through (6) and (7) respectively. According to (8), the test image is also projected to  $M'$ . Nearest Neighbor (NN) classification or support vector machine is then conducted in this  $m$ -dimensional vector space.

## 4. The Proposed Method

In this section, we give a detail description of the proposed MRCDL.

**4.1. Framework.** Generally speaking, the learning framework of training can be viewed as a four-stage procedure. The first step is to represent the  $N$  labeled training samples with the proposed mixed region covariance descriptor, denoted as the covariance matrices as  $\chi = \{X_1, X_2, \dots, X_N\} \subseteq M$ . The second step is to embed the SPD matrices from the

Riemannian manifold  $M$  into a high-dimensional RKHS  $H$  with the Gaussian RBF kernel induced by LERM. As previously stated, RKHS is high-dimensional inner product space of functions rather than vectors, which means that the discriminant analysis cannot be utilized directly. Thus, our third step is to project the elements in the RKHS onto a subspace of the RKHS, which is isomorphic to Euclidean space  $R^m$ . Benefiting from the notion of inner space and the properties of subspace, the mapping from the Riemannian manifold to the final linear subspace can be achieved by an explicit mapping. The subspace projection framework is novel although it is similar to the method in [18, 30, 36]. Finally, discriminative learning based LDA is conducted to obtain the optimized subspace  $H'$ . Then, the samples are mapped to this low-dimensional discriminant subspace  $H'$ . The projections of  $N$  training samples can be denoted as  $y_1, \dots, y_N$ . An overall illustration of training stage is shown in Figure 1.

For the testing stage, using the mixed region covariance matrices of test images as input, we project them onto the discriminant subspace by the explicit mapping learned from the training stage. The classification task can be realized in terms of K-Nearest-Neighbor (KNN) classifier on the subspace.

**4.2. Mixed Region Covariance Descriptor.** As depicted in Section 3.1, given a rectangular image region  $R$ , we can extract the feature vectors such as intensity, color, gradients and filter responses for each pixel. Let  $\{z_k\}_{k=1, \dots, n}$  be the  $d$ -dimensional feature vectors of the  $n$  pixels inside  $R$ . Firstly, we compute the region covariance descriptor  $\Sigma_i \in R^{d \times d}$  as (6) and the mean vector  $\mu_i = (1/n) \sum_{k=1}^n z_k \in R^d$ . By employing the information geometry theory [37, 38], the mean vector and covariance matrix of  $R$  can be embedded into a SPD matrix, and then the image region  $R$  is represented as a  $D \times D$  SPD matrix:

$$X_i = |\Sigma_i|^{-1/(d+1)} \begin{vmatrix} \Sigma_i + \mu_i \mu_i^T & \mu_i \\ \mu_i^T & 1 \end{vmatrix} \tag{12}$$

where  $D = d + 1$ .

The mixed region covariance descriptor inherits the advantages of the region covariance descriptor, namely, good robustness, antinoise capability, and low dimensionality. Furthermore, the mean vectors could improve the discriminative learning ability. Increased performance can be proved by the result of experiments in Section 5.

It is worth noting that the proposed mixed region covariance descriptor totally differs from Gaussian Mixture Model (GMM), which is a parametric statistical model under the assumption of Gaussian distribution for images [39] or image sets [38]. The images are assigned to Gaussian components with respect to a posterior probability. While mixed region covariance descriptor represents images in a global manner, it makes no assumptions about the data distribution. The mean vector and region covariance are computed with feature vector of pixels rather than Gaussian components.

**4.3. Gaussian Kernel and Subspace Projection of RKHS.** It is known that Gaussian kernel is effective in Euclidean spaces. We derive a Gaussian kernel on Riemannian manifolds by replacing Euclidean distance with geodesic distance induced by LERM, which could conserve the geometric properties of manifold and avoid high time complexity. Most importantly, it yields a positive definite kernel, which is a necessary condition for reproducing kernel. The kernel is defined as

$$k_{ij} = k(X_i, X_j) = \exp \left\{ -\frac{d_{LE}^2(X_i, X_j)}{2\sigma^2} \right\} \quad (13)$$

where  $d_{LE}(X_i, X_j) = \|\log(X_i) - \log(X_j)\|_F$  is the Log-Euclidean distance.

Let  $(H, k)$  be the RKHS generated by the reproducing kernel defined in (13). We can define a mapping from  $Sym_d^+$  to RKHS  $H$ :

$$\varphi(X_i) = k(\bullet, X_i) \in H \quad (14)$$

The maps of the input covariance matrices  $X_1, \dots, X_N$  on the RKHS can be denoted as  $\varphi(X_1), \dots, \varphi(X_N)$ . Although each  $\varphi(X_i)$  in the RKHS is a function rather than a vector, the inner product of the Hilbert space can be evaluated by the property of kernel function:

$$\langle \varphi(X_i), \varphi(X_j) \rangle = k(X_i, X_j) \quad (15)$$

Next, we construct an  $m$ -dimensional subspace in RKHS  $H$ . Let  $\phi_i = \sum_{j=1}^N w_{ij} \varphi(x_j)$ ,  $i = 1, \dots, d$ ; the set  $\{\phi_1, \dots, \phi_d\}$  could be an orthogonal basis of  $H$ , only if it is orthonormal, which means that  $\langle \phi_i, \phi_j \rangle = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$ . We have

$$\begin{aligned} \langle \phi_i, \phi_j \rangle &= \left\langle \sum_{p=1}^N w_{ip} \varphi(x_p), \sum_{q=1}^N w_{jq} \varphi(x_q) \right\rangle \\ &= \sum_{p=1}^N \sum_{q=1}^N w_{ip} w_{jq} \langle \varphi(x_p), \varphi(x_q) \rangle \\ &= \sum_{p=1}^N \sum_{q=1}^N w_{ip} w_{jq} k(x_p, x_q) = W_{iRow} K W_{jRow}^T \end{aligned} \quad (16)$$

$$\text{where } W = \begin{bmatrix} w_{11} & \dots & w_{1N} \\ \vdots & \ddots & \vdots \\ w_{m1} & \dots & w_{mN} \end{bmatrix} = \begin{bmatrix} W_{1Row} \\ \vdots \\ W_{dRow} \end{bmatrix}, \quad K = \begin{bmatrix} k(X_1, X_1) & \dots & k(X_1, X_N) \\ \vdots & \ddots & \vdots \\ k(X_N, X_1) & \dots & k(X_N, X_N) \end{bmatrix} = \begin{bmatrix} K_{1Row} \\ \vdots \\ K_{NRow} \end{bmatrix}$$

Then, we can get the constraint of the projection matrix  $W$ :

$$\begin{aligned} I_m &= \begin{bmatrix} \langle \phi_1, \phi_1 \rangle & \dots & \langle \phi_1, \phi_m \rangle \\ \vdots & \ddots & \vdots \\ \langle \phi_m, \phi_1 \rangle & \dots & \langle \phi_m, \phi_m \rangle \end{bmatrix} \\ &= \begin{bmatrix} W_{1Row} K W_{1Row}^T & \dots & W_{1Row} K W_{mRow}^T \\ \vdots & \ddots & \vdots \\ W_{mRow} K W_{1Row}^T & \dots & W_{mRow} K W_{mRow}^T \end{bmatrix} \\ &= W K W^T \end{aligned} \quad (17)$$

where  $I_m$  is the  $m$ -dimensional identity matrix. Thus, we know that  $\{\phi_1, \dots, \phi_m\}$  could span a subspace  $H'$  of  $H$  only if the constraint  $W K W^T = I_m$  is satisfied.

We further project  $\varphi(x_1), \dots, \varphi(x_N)$  onto the subspace spanned by  $\{\phi_1, \dots, \phi_m\}$ . The projection of  $\varphi(x_1), \dots, \varphi(x_N)$  on  $span\{\phi_1, \dots, \phi_m\}$  is

$$\begin{aligned} y_i &= \begin{bmatrix} \langle \varphi(x_i), \phi_1 \rangle \\ \vdots \\ \langle \varphi(x_i), \phi_m \rangle \end{bmatrix} = \begin{bmatrix} \left\langle \varphi(x_i), \sum_{j=1}^N w_{1j} \varphi(x_j) \right\rangle \\ \vdots \\ \left\langle \varphi(x_i), \sum_{j=1}^N w_{mj} \varphi(x_j) \right\rangle \end{bmatrix} \\ &= \begin{bmatrix} \sum_{j=1}^N w_{1j} \langle \varphi(x_i), \varphi(x_j) \rangle \\ \vdots \\ \sum_{j=1}^N w_{mj} \langle \varphi(x_i), \varphi(x_j) \rangle \end{bmatrix} \\ &= \begin{bmatrix} \sum_{j=1}^N w_{1j} k(x_i, x_j) \\ \vdots \\ \sum_{j=1}^N w_{mj} k(x_i, x_j) \end{bmatrix} = \begin{bmatrix} W_{1Row} K_{iRow}^T \\ \vdots \\ W_{dRow} K_{iRow}^T \end{bmatrix} \\ &= W K_{iRow}^T; \quad i = 1, \dots, N \end{aligned} \quad (18)$$

Thus, we can project the points  $X_1, \dots, X_N$  from a Riemannian manifold onto a low-dimensional linear subspace under the constraint of  $W K W^T = I_m$ . Let  $Y = [y_1 \dots y_N] \in R^{m \times N}$  be the low-dimensional embedding of  $X_1, \dots, X_N$ ; we have

$$Y = [WK_{1Row}^T \cdots WK_{NRow}^T] = WK^T \quad (19)$$

The subspace projection framework is different from the mapping in [18, 30, 36], which claims that the points on a Riemannian manifold can be mapped onto Euclidean space directly or indirectly based upon a linear hypothesis, but they have not given the justification, while in the proposed subspace projection framework, the constraint of  $WKW^T = I_m$  is an important precondition.

**4.4. Discriminative Learning on Riemannian Manifolds.** A finite dimensional linear space must be isomorphic to the same dimensional Euclidean space [37]; thus the subspace  $H'$  of RKHS  $H$  is isomorphic to  $m$ -dimensional Euclidean space  $R^m$ . The proposed MRCDL recast LDA [40] from Euclidean space to the subspace  $H'$  and develop a variant of LDA to seek the optimal subspace. The LDA is well known for its effectiveness to learn discriminant subspaces. In order to make the points from the same class be more compact and the points from different classes to be further separated, LDA seeks for a projection matrix such that the ratio of the between-class scatter matrix and the within-class scatter matrix is maximized after the projection.

Suppose the input data come from  $C$  classes. We denote the data of the  $c$ -th class as  $Y_c$ , and  $Y_c$  can be represented as

$$Y_c = [y_{c_1} \cdots y_{c_{N_c}}] = YS_c = WK^T S_c; \quad c = 1, \dots, C \quad (20)$$

where  $S_c$  is the selection matrix such that  $Y_c = YS_c$ ;  $N_c$  is the number of images in the  $c$ -th class.

We denote the mean and the centralization of  $Y_c$  as  $m_c$  and  $\hat{Y}_c$ , respectively, and they can be computed by

$$m_c = \frac{1}{N_c} Y_c \Gamma_{N_c} = \frac{1}{N_c} WK^T S_c \Gamma_{N_c} \quad (21)$$

$$\begin{aligned} \hat{Y}_c &= [y_{c_1} - m_c \cdots y_{c_{N_c}} - m_c] = Y_c E_{N_c} \\ &= WK^T S_c E_{N_c} \end{aligned} \quad (22)$$

where  $\Gamma_{N_c} \in R^{N_c}$  is a  $N_c$  dimensional vector whose elements are all 1;  $E_{N_c} = I_{N_c} - (1/N_c)\Gamma_{N_c}\Gamma_{N_c}^T$  is the centralized matrix.

The between-class measure of the  $c$ -th class can be computed by

$$\begin{aligned} \sum_{i=1}^{N_c} (y_{c_i} - m_c)(y_{c_i} - m_c)^T &= \hat{Y}_c \hat{Y}_c^T \\ &= WK^T S_c E_{N_c} E_{N_c}^T S_c^T K W^T = WL_c W^T \end{aligned} \quad (23)$$

where  $L_c = K^T S_c E_{N_c} E_{N_c}^T S_c^T K$ .

Then the between-class scatter matrix can be represented as

$$S_W = \frac{1}{C} \sum_{c=1}^C WL_c W^T = W \left( \frac{1}{C} \sum_{c=1}^C L_c \right) W^T = WL_W W^T \quad (24)$$

where  $L_W = (1/C) \sum_{c=1}^C L_c$ .

Let  $M = [m_1 \cdots m_C]$ ; we have

$$\begin{aligned} M &= \left[ \frac{1}{N_1} WK^T S_1 \Gamma_{N_1} \cdots \frac{1}{N_C} WK^T S_C \Gamma_{N_C} \right] \\ &= WK^T \left[ \frac{1}{N_1} S_1 \Gamma_{N_1} \cdots \frac{1}{N_C} S_C \Gamma_{N_C} \right] = WK^T S \end{aligned} \quad (25)$$

where  $S = [(1/N_1)S_1\Gamma_{N_1} \cdots (1/N_C)S_C\Gamma_{N_C}]$ .

The between-class scatter matrix can be formulated as

$$S_B = ME_C E_C^T M^T = WK^T S E_C E_C^T S^T K W^T = WL_B W^T \quad (26)$$

where  $E_C = I_N - (1/N)\Gamma_N \Gamma_N^T$  and  $L_B = K^T S E_C E_C^T S^T K$ .

The MRCDL seeks to solve the following optimization under the constraint of  $WKW^T = I_m$ :

$$W = \underset{W}{\operatorname{argmax}} (S_B - S_W) = \underset{W}{\operatorname{argmax}} W (L_B - L_W) W^T \quad (27)$$

$$\text{s.t. } WKW^T = I_m$$

The optimal  $W$  is given by the  $m$  largest eigenvectors of the eigenproblem:  $(L_B - L_W)W = \lambda KW$ . We can get the  $m$ -dimensional embedding of the training samples by (19). Given a test example  $X_t \in \text{Sym}_d^+$ , its  $m$ -dimensional projection  $y_t$  in the discriminant subspace is obtained likewise. Then the class of test region can be predicted by KNN classification in the  $m$ -dimensional vector space.

The proposed MRCDL is presented in Algorithm 1.

**4.5. Computational Complexity.** The complexity of computing the covariance of rectangular region mainly focuses on the computation of region covariance. A fast computation of the region covariance descriptor can be realized based on the integral image [1, 6], and the computational cost for any input image region is  $O[d^2(n+1)]$ . Considering the properties of the LERM, we perform the matrix logarithm operate on each mixed region covariance matrix; it can be computed in  $O[D^3]$  time. The complexity of computing the descriptor and matrix log operation is linear in  $N$ .

Neglecting the complexity of the exponential in kernel functions and considering the symmetry of kernel matrix, the complexity of computing kernel matrix is  $O[(N(N-1)/2)D^2]$ . The optimization of the projection matrix involves the calculation of within-scatter, the calculation of between-class scatter, and the eigenvalue decomposition. The complexity of computing within-class scatter and between-class scatter is  $O[\sum_{c=1}^C mN_c^2]$  and  $O[mC^2]$ , respectively. The complexity of the eigenvalue decomposition is  $O[m^3]$ .

In summary, the entire computational complexity of the MRCDL is

$$\begin{aligned} &O \left[ N(d^2(n+1) + D^3) + \frac{N(N-1)}{2} D^2 \right. \\ &\quad \left. + m \left( \left( \sum_{c=1}^C N_c^2 \right) + C^2 + m^2 \right) \right] \end{aligned} \quad (28)$$

*Input:*

(i) The mixed region covariance descriptors  $\{X_1, X_2, \dots, X_N\}$  of the  $N$  image regions for training and the corresponding labels  $\{l_1, l_2, \dots, l_N\}$ .

(ii) The number of images in the  $c$ -th class  $N_c$ ,  $\sum_{c=1}^C N_c = N$ .

(iii) The mixed region covariance descriptor  $X_t$  of the test image region.

*Processing:*

1: Calculate the matrix logarithm  $\log(X_i)$  for all  $X_i$ ,

2: *for* each point  $X_i$  *do*

3: *for* each point  $X_j$  *do*

4: Calculate the log-Euclidean distance  $d_{LED}(X_i, X_j)$  between  $X_i$  and  $X_j$  (2),

5: *end for*

6: *end for*

7: Calculate the kernel matrix  $K$  (13),

8: *for* each class  $Y_c$  *do*

9: Calculate the mean  $m_c$  (21) and the centralized matrix  $\hat{Y}_c$  (22),

10: *end for*

11: Calculate the within-class scatter matrix  $L_W$  (24),

12: Calculate the between-class scatter matrix  $L_B$  (26),

13: Solve the eigen-problem:  $(L_B - L_W)W = \lambda KW$ . The eigenvectors corresponding to the  $m$  largest eigenvalues of  $K^{-1}(L_B - L_W)$  form  $W$ ,

14: Calculate the projections  $y_1 \dots y_N$  of  $\{X_1, X_2, \dots, X_N\}$  (19),

15: Calculate the projection  $y_t$  of the testing matrix  $X_t$  (18),

16: The label  $l_t$  of the test region is determined by KNN classifier between  $y_t$  and  $y_1 \dots y_N$ ,

*Output:*

(i) The class of the test region  $l_t$ .

## ALGORITHM 1

## 5. Experiment

Experiments are performed on the recognition and categorization tasks. Several state-of-the-art approaches for image classification are adopted for comparison. The performance of the mixed region covariance and the proposed MRCDL can be verified in this section. All of the methods were implemented on an Intel(R) Core(TM) i5-4670K (3.40GHz) PC, using Matlab R2014.

**5.1. Datasets and Experimental Settings.** We used the COIL-20 (Columbia Object Image Library) dataset [41] and the ETH80 dataset [42] for the object categorization task. The COIL-20 consists of 20 different objects; for each object there are 72 images of different viewpoints. The images are manually cropped and are normalized to  $128 \times 128$  pixels. To calculate the mixed region covariance descriptor of an image, each pixel in the image was mapped onto a 5-dimensional feature space  $F(x, y) = [I, |I_x|, |I_y|, |I_{xx}|, |I_{yy}|]$ , where  $I$  is the image intensity and  $I_x, I_y, I_{xx}, I_{yy}$  are the first and the second order derivatives of intensities in the  $x$  and the  $y$  directions, then each image was described by a  $6 \times 6$  covariance matrix. We randomly selected 10 images per class for training and the remaining 62 images for testing.

ETH80 dataset contains 8 categories including apples, pears, tomatoes, cars, cows, cups, dogs, and horses. Each category comprises pictures of 10 object instances and 41 images per object. Images in the dataset are manually cropped and are rescaled to  $128 \times 128$  pixels. Each pixel was mapped onto a 10-dimensional feature vector  $F(x, y) = [x, y, R_{x,y}, G_{x,y}, B_{x,y}, I_{x,y}, |I_x|, |I_y|, |I_{xx}|, |I_{yy}|]$  to

calculate the mixed region covariance matrix, where  $x$  and  $y$  are pixel locations,  $R_{x,y}, G_{x,y}, B_{x,y}$  represent the corresponding RGB color values,  $I_{x,y}$  is the intensity, and  $I_x, I_y, I_{xx}, I_{yy}$  are the first and the second order derivatives of intensities in the  $x$  and the  $y$  directions. The descriptor for each image therefore was a  $11 \times 11$  dimensional SPD matrix. In each experiment, this database was split into the training one and the testing one, where each category had 5 objects for training and the remaining 5 objects for testing. We randomly chose 20 images per object in the training objects and the testing objects. That is to say, there were 100 images for training and 100 images for testing in each category. The test images and the training images came from different object instances even in the same categories. The KNN classifier was adopted to predict the labels of the input testing images.

We conduct the texture classification task on the Brodatz dataset, which consists of 111 textures of size  $640 \times 640$  [43]. Each image is divided into four  $320 \times 320$  images, half of which are used for training and the remaining for testing. In each training image, the mixed region covariance descriptors of 50 randomly chosen image regions of size  $128 \times 128$  were computed from a 5-dimensional feature vector  $[I_{x,y}, |I_x|, |I_y|, |I_{xx}|, |I_{yy}|]$ . For a given test image, we computed 100 covariance matrices from randomly selected image regions of size between  $16 \times 16$  and  $128 \times 128$ . We used the KNN classifier ( $k=5$ ) for the classification of each input region; thus we obtained 100 labels for each image. The classes of the testing images were predicted using majority voting among the 100 labels.

TABLE 1: Recognition rates (%) comparisons of different methods.

Methods	COIL-20	ETH80	Brodatz
NDML	90.05	58.13	\
DA-LGSP	91.98	64.18	\
KPCA	81.05	72.61	94.14
KPCA+MRC	83.79	73.7	95.05
RLPP	85.89	74.08	95.5
RLPP+MRC	88.79	75.63	96.85
RCDL	95.21	79.81	97.75
MRCDL	96.7	81.88	99.1

TABLE 2: Computation time (seconds) of different methods on COIL-20 dataset.

Methods	NDML	DA-LGSP	KPCA	KPCA+MRC	RLPP	RLPP+MRC	RCDL	MRCDL
Time	1.121	1.448	5.035	5.274	182.172	182.411	5.669	5.908

The proposed MRCDL is compared with KPCA [32], RLPP [30], NDML [9], and DA-LGSP [10]. To validate the effectiveness of the mixed region covariance, the proposed MRCDL is also conducted under the original region covariance descriptor, denoted by RCDL. The KPCA and the RLPP are also conducted under the mixed region covariance descriptor, denoted by KPCA+MRC and RLPP+MRC, respectively. In NDML and DA-LGSP, we preserve nearly 95% image energy to select the number of principal components. The KNN ( $k = 8$ ) search is used for constructing the neighborhood graphs, as well as RLPP. The parameter  $\sigma$  in Gaussian kernel of KPCA, RLPP, and MRCDL is set to the average value of distances. Besides, the 5NN classifier is adopted to predict the labels of the testing images (or image regions).

**5.2. Performance Comparison with Other Methods.** The performance of the methods is compared on the recognition rates and the computation time. The *recognition rate*  $\in [0, 1]$  is the ratio between the number of correct identifying tests and the number of the total testing images. To overcome the randomness of experiments, the recognition rate is the average of 20 rounds of tests. We present the recognition rates in Table 1. Since NDML and DA-LGSP require uniform dimensions of the input data, the texture recognition task cannot be realized by the two methods.

As shown in Table 1, our proposed MRCDL not only makes a considerable improvement against RLPP, but also achieves the best recognition rate in the compared methods. This is because the use of class labels provides more between-class and within-class information for discriminative learning. In addition, the recognition rates of RCDL, RLPP and KPCA are lower than the recognition rates of the corresponding methods, namely MRCDL, RLPP+MRC, and KPCA+MRC. This result demonstrates the performance of the mixed region covariance descriptor.

On COIL-20 dataset, we compared the computational time of different methods. The time cost for each method is listed in Table 2. The NDML and DA-LGSP are faster than the methods based on covariance descriptors, since the cost of matrix logarithm operation is time-consuming. But our

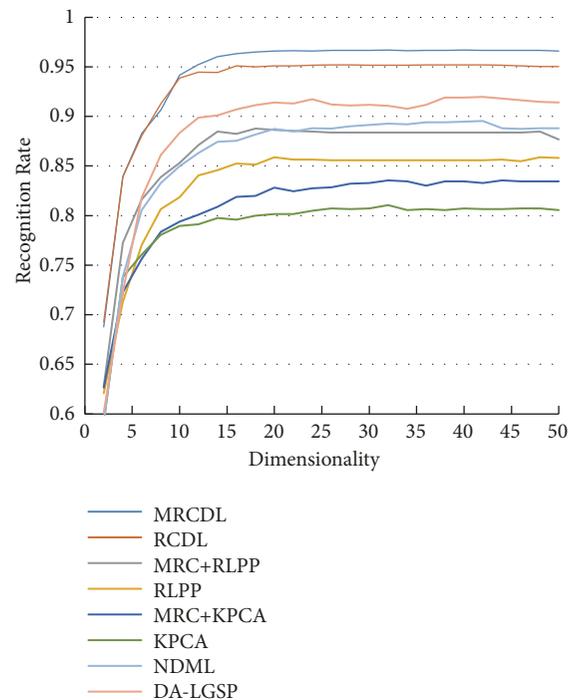


FIGURE 2: The recognition rates versus different dimensionalities on COIL-20 database.

method is much faster than RLPP. The justification for this big difference is that the geodesic distance induced by LERM is more efficient than AIRM.

**5.3. Analysis of Parameter  $m$ .** Parameter  $m$  represents the dimensionality of the projecting subspace space, which has a certain influence on the recognition rates. In this part, the recognition rates versus different dimensionality are discussed. We present the results in Figures 2, 3, and 4. It can be found that the curve of each method shows increasing trend with the increase of dimensionality. Each curve almost keeps unchanged when the recognition rate achieves its maxima. The effectiveness of mixed region covariance has been proved again by the best recognition rates on all of the three datasets.

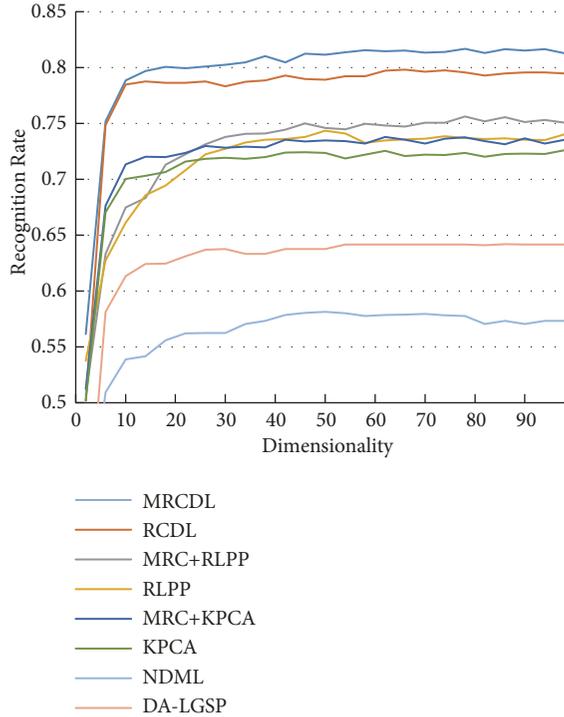


FIGURE 3: The recognition rates versus different dimensionalities on ETH-80 database.

Moreover, we can find that the dimensionality corresponding to the maxima of MRCDL and RCDL is smaller than other methods especially on Brodatz dataset, which indicates that the proposed method obtains good performance even in lower dimensionality.

5.4. Analysis of the Kernel Width  $\sigma$ . In the statistical model of a Gaussian distribution, the parameter  $\sigma$  determines the bandwidth of the distribution. As shown in (13), as long as the Riemannian metric is fixed, the parameter  $\sigma$  plays an important role in Gaussian kernel. For small values of  $\sigma$ , the neighborhoods of points on the RKHS are small. When  $\sigma$  becomes bigger, the dissimilarity of points tends ambiguous. Note that  $\sigma = \infty$  corresponding to the case when all of the elements of kernel matrix are set to 1. We suggest making  $k_{ij}$  moderately sensitive to distances and choosing the average value of distances as the value of  $\sigma$ . Experiments are conducted on COIL-20 dataset and ETH80 dataset to demonstrate the effect of  $\sigma$ . The recognition rates under different times of the average value are presented in Figure 5. It is shown that the method performs much better when  $\sigma$  is set to 0.5 to 5 times of the average value of distances.

### 6. Conclusions

In this paper, we proposed an image classification method called mixed region covariance discriminative learning (MRCDL) on Riemannian manifolds. Image regions are represented by the proposed mixed region covariance descriptors. The descriptor could improve the discriminative learning ability of covariance descriptors. This method employed

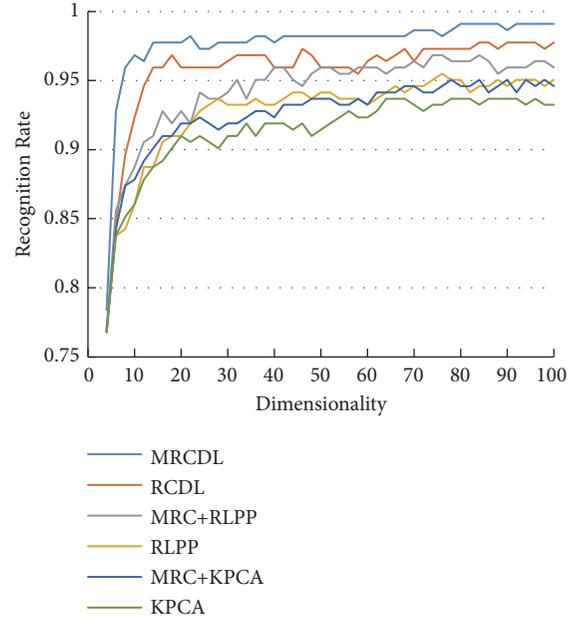


FIGURE 4: The recognition rates versus different dimensionalities on Brodatz dataset.

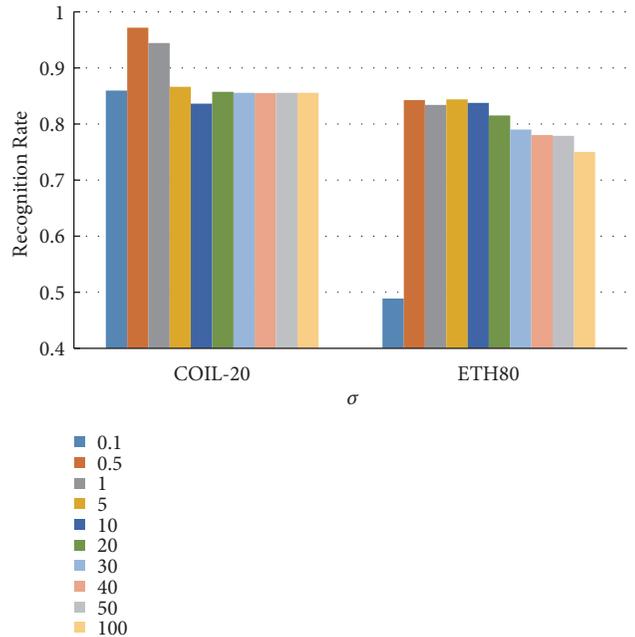


FIGURE 5: The recognition rates under different times of the average of the distance matrix.

a Log-Euclidean Gaussian kernel to embed Riemannian manifolds into RKHS and further projected the points onto a subspace of the RKHS. The kernel version of LDA is conducted on the low-dimensional subspace. Compared with other methods, the proposed method always outperforms other methods, even with low dimensionality. The justification comes from the advantages of covariance descriptors, the within-class and between-class information. Unfortunately the proposed method still suffers several drawbacks: first of

all, the computational cost of our method is not as fast as NDML and DA-LGSP; nevertheless, the time increment is relatively low. Secondly, our method is a supervised method, which means that all of the training samples should be labeled beforehand.

## Data Availability

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

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China through the Project “Research on Nonlinear Alignment Algorithm of Local Coordinates in Manifold Learning” under Grant 61773022.

## References

- [1] O. Tuzel, F. Porikli, and P. Meer, “Pedestrian detection via classification on Riemannian manifolds,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1713–1727, 2008.
- [2] X. Lan, A. J. Ma, and P. C. Yuen, “Multi-cue visual tracking using robust feature-level fusion based on joint sparse representation,” in *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR '14)*, pp. 1194–1201, Columbus, Ohio, USA, June 2014.
- [3] X. Lan, Z. Shengping, and P. C. Yuen, “Robust joint discriminative feature learning for visual tracking,” in *Proceedings of the 25th International Joint Conference on Artificial Intelligence, IJCAI 2016*, pp. 3403–3410, July 2016.
- [4] F. Porikli, O. Tuzel, and P. Meer, “Covariance tracking using model update based on Lie algebra,” in *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2006*, pp. 728–735, USA, June 2006.
- [5] D. Tosato, M. Spera, M. Cristani, and V. Murino, “Characterizing humans on riemannian manifolds,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1972–1984, 2013.
- [6] X. Wang, S. Liu, Z. Liu, and Y. Zhang, “Underwater sonar image detection: A combination of non-local spatial information and quantum-inspired shuffled frog leaping algorithm,” *PLoS ONE*, vol. 12, no. 5, p. e0177666, 2017.
- [7] U. Ozgunalp, R. Fan, X. Ai, and N. Dahnoun, “Multiple lane detection algorithm based on novel dense vanishing point estimation,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 621–632, 2017.
- [8] Q. D. Ma, Z. Ma, C. Ji, K. Yin, T. Zhu, and C. Bian, “Artificial object edge detection based on enhanced Canny algorithm for high-speed railway apparatus identification,” in *Proceedings of the 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pp. 1–6, Shanghai, October 2017.
- [9] B. Li, J. Li, and X.-P. Zhang, “Nonparametric discriminant multi-manifold learning for dimensionality reduction,” *Neurocomputing*, vol. 152, pp. 121–126, 2015.
- [10] X. Liu and Z. Ma, “Discriminant analysis with local gaussian similarity preserving for feature extraction,” *Neural Processing Letters*, vol. 47, no. 1, pp. 39–55, 2018.
- [11] M. Woźniak and D. Połap, “Adaptive neuro-heuristic hybrid model for fruit peel defects detection,” *Neural Networks*, vol. 98, pp. 16–33, 2018.
- [12] M. Woźniak and D. Połap, “Bio-inspired methods modeled for respiratory disease detection from medical images,” *Swarm and Evolutionary Computation*, vol. 41, pp. 69–96, 2018.
- [13] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [14] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 886–893, June 2005.
- [15] X. Lan, S. Zhang, P. C. Yuen, and R. Chellappa, “Learning common and feature-specific patterns: a novel multiple-sparse-representation-based tracker,” *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2022–2037, 2018.
- [16] X. Lan, A. J. Ma, P. C. Yuen, and R. Chellappa, “Joint sparse representation and robust feature-level fusion for multi-cue visual tracking,” *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 5826–5841, 2015.
- [17] O. Tuzel, F. Porikli, and P. Meer, “Region covariance: a fast descriptor for detection and classification,” in *Computer Vision—ECCV 2006*, vol. 3952 of *Lecture Notes in Computer Science*, pp. 589–600, Springer, Berlin, Germany, 2006.
- [18] R. Wang, H. Guo, L. S. Davis, and Q. Dai, “Covariance discriminative learning: A natural and efficient approach to image set classification,” in *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2496–2503, Providence, RI, USA, June 2012.
- [19] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, vol. 80, Academic Press, Inc., 1978.
- [20] A. Goh and R. Vidal, “Clustering and dimensionality reduction on Riemannian manifolds,” in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, pp. 626–632, USA, June 2008.
- [21] K. Guo, P. Ishwar, and J. Konrad, “Action recognition using sparse representation on covariance manifolds of optical flow,” in *Proceedings of the 7th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS '10)*, pp. 188–195, IEEE, Boston, Mass, USA, August–September 2010.
- [22] C. Rui, P. Martins, J. F. Henriques, F. S. Leite, and J. Batista, “Rolling riemannian manifolds to solve the multi-class classification problem,” in *Proceedings of the IEEE Conference on Computer Vision Pattern Recognition*, vol. 9, pp. 41–48, 2013.
- [23] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*, Cambridge University Press, Cambridge, UK, 2004.
- [24] M. T. Harandi, C. Sanderson, R. Hartley, and B. C. Lovell, “Sparse coding and dictionary learning for symmetric positive definite matrices: A kernel approach,” in *Proceedings of the 12th European conference on Computer Vision*, vol. 7573, pp. 216–229, 2012.
- [25] P. Li, Q. Wang, W. Zuo, and L. Zhang, “Log-euclidean kernels for sparse representation and dictionary learning,” in *Proceedings of the 2013 14th IEEE International Conference on Computer Vision, ICCV 2013*, pp. 1601–1608, Australia, December 2013.

- [26] C. Rui, J. F. Henriques, P. Martins, and J. Batista, "Semi-intrinsic mean shift on riemannian manifolds," in *Proceedings of the European Conference on Computer Vision*, pp. 342–355, 2012.
- [27] R. Vemulapalli, J. K. Pillai, and R. Chellappa, "Kernel learning for extrinsic classification of manifold features," in *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2013*, pp. 1782–1789, June 2013.
- [28] Z. Huang, R. Wang, S. Shan, and X. Chen, "Learning Euclidean-to-Riemannian metric for point-to-set classification," in *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014*, pp. 1677–1684, June 2014.
- [29] Z. Huang, R. Wang, S. Shan, X. Li, and X. Chen, "Log-euclidean metric learning on symmetric positive definite manifold with application to image set classification," in *Proceedings of the 32nd International Conference on Machine Learning, ICML 2015*, pp. 720–729, France, July 2015.
- [30] M. T. Harandi, C. Sanderson, A. Wiliem, and B. C. Lovell, "Kernel analysis over Riemannian manifolds for visual recognition of actions, pedestrians and textures," in *Proceedings of the 2012 IEEE Workshop on the Applications of Computer Vision, WACV 2012*, pp. 433–439, USA, January 2012.
- [31] X. Pennec, P. Fillard, and N. Ayache, "A riemannian framework for tensor computing," *International Journal of Computer Vision*, vol. 66, no. 1, pp. 41–66, 2006.
- [32] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi, "Kernel methods on riemannian manifolds with gaussian RBF kernels," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 12, pp. 2464–2477, 2015.
- [33] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, "Log-Euclidean metrics for fast and simple calculus on diffusion tensors," *Magnetic Resonance in Medicine*, vol. 56, no. 2, pp. 411–421, 2006.
- [34] A. Cherian, S. Sra, A. Banerjee, and N. Papanikolopoulos, "Jensen-Bregman LogDet divergence with application to efficient similarity search for covariance matrices," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 9, pp. 2161–2174, 2013.
- [35] E. Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons, New York, NY, USA, 1978.
- [36] W. Wang, R. Wang, Z. Huang, S. Shan, and X. Chen, "Discriminant analysis on Riemannian manifold of Gaussian distributions for face recognition with image sets," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 151–163, 2018.
- [37] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Oxford University Press, Oxford, UK, 2000.
- [38] M. Lovrić, M. Min-Oo, and E. A. Ruh, "Multivariate normal distributions parametrized as a Riemannian symmetric space," *Journal of Multivariate Analysis*, vol. 74, no. 1, pp. 36–48, 2000.
- [39] X. Zhou, N. Cui, Z. Li, F. Liang, and T. S. Huang, "Hierarchical Gaussianization for image classification," in *Proceedings of the 12th International Conference on Computer Vision, ICCV 2009*, pp. 1971–1977, Japan, October 2009.
- [40] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Computation*, vol. 12, no. 10, pp. 2385–2404, 2000.
- [41] S. Nene, S. Nayar, and H. Murase, "Columbia object image library (coil-20)," Technical Report, CUCS-005-96, 1996.
- [42] B. Leibe and B. Schiele, "Analyzing appearance and contour based methods for object categorization," in *Proceedings of the CVPR 2003: Computer Vision and Pattern Recognition Conference*, vol. 2, pp. 409–415, Madison, WI, USA, 2003.
- [43] T. Randen and J. H. Husøy, "Filtering for texture classification: A Comparative Study," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 291–310, 1999.



**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

