Image-set classification using Discriminant Neighborhood Preserving Embedding on Symmetric Positive Definite Manifold

1st Ying Hu
College of Computer and Information Science
Chongqing Normal University
Chongqing, China
Email: 2024210516043@stu.cqnu.edu.cn

2nd Benchao Li
School of Computing and Artificial Intelligence
Southwest Jiaotong University
Chengdu, China
Email: libenchao517@my.swjtu.edu.cn

3rd Ruisheng Ran *
College of Computer and Information Science
Chongqing Normal University
Chongqing, China
Email: rshran@cqnu.edu.cn

4th Ji Feng
College of Computer and Information Science
Chongqing Normal University
Chongqing, China
Email: jifeng@cqnu.edu.cn

***Abstract*—Symmetric positive definite (SPD) manifolds provide a fundamental representation for high-dimensional image sets and covariance descriptors by capturing second-order statistics and intrinsic geometric structures. However, in high-dimensional scenarios, SPD manifold data commonly suffer from the curse of dimensionality and geometric distortion, while conventional Euclidean dimensionality reduction methods fail to preserve Riemannian consistency. To address these challenges, this paper proposes a discriminative neighborhood preserving embedding on SPD manifolds (SPD-DNPE) with dual-mode neighborhood weighting strategies. By extending neighborhood preserving embedding(NPE) to the SPD Riemannian manifold and incorporating class label information, the proposed method enhances discriminability in the low-dimensional tangent space while preserving local manifold geometry. Extensive experiments on multiple benchmark datasets demonstrate that SPD-DNPE achieves superior classification accuracy and robustness compared with existing Grassmann and SPD manifold methods, kernel-based approaches, and several deep manifold models.**

***Index Terms*—Manifold Learning, Symmetric Positive Definite Manifold, Image-Set Classification, Neighborhood Preserving Embedding, Unsupervised Learning, Discriminant Learning, Machine Learning.**

I. Introduction

Image set classification effectively utilizes spatiotemporal information but suffers from high-dimensionality and nonlinearity. While SPD manifolds capture intrinsic geometric structures, traditional dimensionality reduction often neglects discriminative local information. To bridge this, Jayasumana et al. [1] proposed kernel learning for RKHS embedding, while Harandi et al. [2] introduced more scalable manifold-to-manifold mappings. The field has since shifted toward deep learning; SPDNet [3] pioneered end-to-end bilinear mapping, which was further enhanced by DreamNet [18] for generalization via reconstruction constraints. Recent advances include multi-scale local modeling [19] for improved discriminative power and Bures–Wasserstein-based batch normalization [20] to ensure numerical stability under ill-conditioned scenarios.

Unsupervised manifold dimensionality reduction preserves intrinsic geometry without label information. On Grassmann manifolds, Grassmann Locality Preserving Projections (GLPP) [5] and Neighborhood Preserving Embedding on the Grassmann Manifold (GNPE) [6] extend locality-preserving principles, while Grassmannian Adaptive Local Learning (GALL) [21] introduces adaptive neighborhood learning. Further, Embedding Graphs on the Grassmann Manifold (EGG) [22] enables graph-structured embedding in Grassmann space. For SPD manifolds, Locality Preserving Projections on the SPD Matrix Lie Group (Lie-LPP) [4] captures symmetric transformations, and Riemannian Manifold Tangent Spaces and Local Affinity for SPD data (RMTSLA-SPDDR) [23] leverages tangent spaces with local affinities. Finally, Ran et al. [13] further extended classical manifold dimensionality reduction methods through a unified matrix function modeling framework.

Supervised manifold dimensionality reduction leverages labels to optimize class geometry and discriminability. Specifically, Wang et al. [7] and Harandi et al. [8] proposed Covariance Discriminative Learning (CDL) and Dimensionality Reduction Metric Learning (DRML) for SPD embedding and joint metric learning, while Zhu et al. [9] generalized this via Riemannian Manifold Metric Learning (RMML). For SPD data, Gao et al. [10] and Xu et al. [11] developed Point-to-Set and Set-to-Set Distance Measure (PSSSD) and Polynomial Manifold Learning (PML) using sub-manifolds and geodesics.Furthermore, Ran et al. employed autoencoder-based bidirectional mappings for reliable low-dimensional

* Ruisheng Ran is the corresponding author.

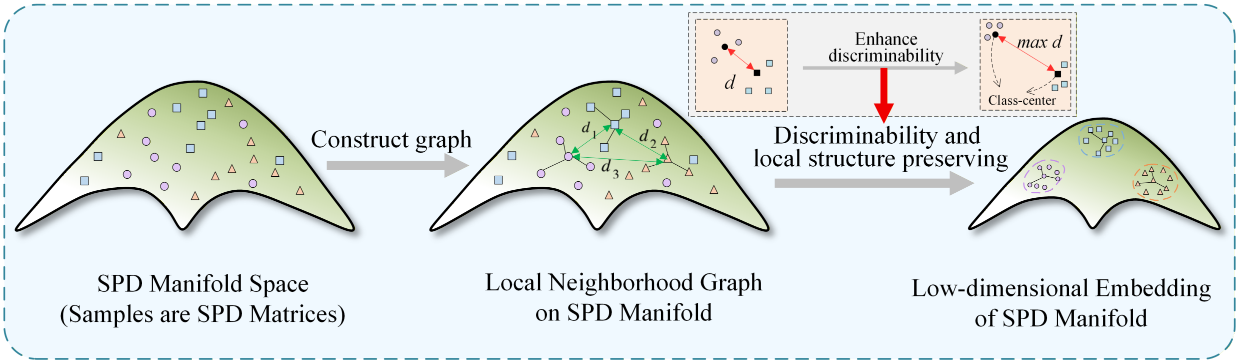


Fig. 1: Conceptual illustration of the proposed discriminant neighborhood preserving embedding on SPD Manifold.

embeddings [14]. Finally, Li et al. introduced Discriminant Neighborhood Preserving Embedding (GDNPE) [12] on Grassmann manifolds to improve convergence and classification performance.

However, existing neighborhood-preserving methods focus primarily on Euclidean or Grassmann spaces. While some studies have introduced locality preservation or discriminative learning to SPD manifolds, they typically rely on global projections or kernel frameworks, failing to characterize local geometry through neighborhood reconstruction. To address this issue, this study proposes the extension of Neighborhood Preserving Embedding (NPE) to the SPD Riemannian manifold, and further develops its supervised discriminative variant, Discriminant Neighborhood Preserving Embedding on SPD Manifold (SPD-DNPE), enabling the unified modeling of both the local geometric structure and discriminative information of high-dimensional covariance data. The fundamental concepts of SPD-DNPE are schematically illustrated in Fig. 1.

The main contributions of this paper are summarized as follows:

- We propose dual-mode weighting on SPD manifolds to enhance local geometry preservation and robustness against noisy data distributions.
- We develop SPD-DNPE, extending NPE to SPD manifolds with discriminative constraints to jointly optimize intra-class compactness and inter-class separability.
- We bridge Riemannian and Euclidean spaces via a Calibrated Log-Euclidean (CLE) representation, enabling geometry-consistent SVM classification of SPD-DNPE outputs.

II. Preliminary Knowledge

A. SPD Manifold

Image sets and video sequences are frequently modeled as covariance matrices on the Symmetric Positive Definite (SPD) manifold. Formally, the set of $d \times d$ SPD matrices, denoted by $\mathcal{S}_{++}^d$, is defined as:

$$\mathcal{S}_{++}^d = \left\{ \boldsymbol{X} \in \mathbb{R}^{d\times d} \mid \boldsymbol{X} = \boldsymbol{X}^T, \boldsymbol{v}^T X \boldsymbol{v} > 0, \forall \boldsymbol{v} \in \mathbb{R}^d, \boldsymbol{v} \neq \boldsymbol{0} \right\} \quad (1)$$

Geometrically, $\mathcal{S}_{++}^d$ is a smooth Riemannian manifold rather than a flat Euclidean space. Standard Euclidean operations (e.g., arithmetic mean) are inapplicable here as they may violate positive definiteness. Therefore, processing SPD data necessitates a non-Euclidean Riemannian geometric framework.

B. Projection metric on SPD manifold

The Log-Euclidean Metric (LEM) utilizes the matrix logarithm $\log(\cdot) : \mathcal{S}_{++}^D \to Sym(D)$ to map the curved SPD manifold onto the flat tangent space of symmetric matrices. This isomorphic mapping linearizes the manifold structure, allowing Riemannian operations to be performed via Euclidean linear algebra. Specifically, the geodesic distance d_g between $\boldsymbol{X}_i$ and $\boldsymbol{X}_j$ is computed as the Frobenius norm in the tangent space:

$$d_g^2(\boldsymbol{X}_i, \boldsymbol{X}_j) = \|\log(\boldsymbol{X}_i) - \log(\boldsymbol{X}_j)\|_F^2 \quad (2)$$

By converting nonlinear manifold operations into efficient linear ones, LEM provides the metric foundation for the proposed dimensionality reduction and classification framework.

C. NPE

Neighborhood Preserving Embedding (NPE) is an unsupervised linear dimensionality reduction method that approximates LLE to preserve local manifold structures. It reconstructs each sample x_i as a linear combination of its k nearest neighbors using weights W. NPE seeks an optimal projection matrix A that maintains these local relationships in the low-dimensional subspace by minimizing the following objective:

$$\Phi_{NPE} = \min_{\boldsymbol{A}} \sum_{i=1}^{N} \left\| \boldsymbol{A}^T x_i - \sum_{j=1}^{N} \boldsymbol{W}_{ij} \boldsymbol{A}^T x_j \right\|^2 \quad (3)$$

The aforementioned optimization problem is ultimately reduced to solving a generalized eigenvalue problem.

III. The Proposed Methods

This study first draws inspiration from the concept of NPE and proposes an unsupervised SPD-NPE algorithm by designing strategies for local linear embedding and regularized weight learning. Furthermore, to overcome the limitations of unsupervised learning in classification tasks, this study places particular emphasis on proposing a supervised SPD-DNPE algorithm. This is achieved by constructing a discriminative neighborhood graph with the constraint of class labels and optimizing a trace-ratio objective function, thereby minimizing the local reconstruction error within the same class while maximizing the inter-class center distance.

A. Dual-Mode Neighborhood Weighting Strategies on SPD Manifolds

For each sample $\boldsymbol{X}_i \in S_{++}^D$, we construct a neighborhood set $\mathcal{N}(i)$ consisting of its k nearest neighbors based on the specific task requirements and compute the corresponding reconstruction weights s_{ij}.We design two reconstruction weighting strategies tailored to different data distribution characteristics:

Mode 1: Local Linear Embedding (LLE). This mode relies on the standard assumption of local linear reconstruction, positing that the geometric relationship between a sample and its neighbors remains invariant across the dimensionality reduction process. The weights s_{ij} are derived by minimizing the reconstruction error within the tangent space:

$$\min_{s_i} \sum_i \left\| \log(\boldsymbol{X}_i) - \sum_{j \in N_k(i)} s_{ij} \log(\boldsymbol{X}_j) \right\|_F^2 \tag{4}$$

To preserve the intrinsic local geometry and shift-invariance, we impose a strict sum-to-one constraint on the weights, denoted as $\sum_{j \in N_k(i)} s_{ij} = 1$. This mode is optimal for data exhibiting uniform distribution and low noise.

Mode 2: L2-Regularization. In practical applications, the neighborhood of a sample may exhibit multicollinearity, resulting in a singular or unstable local Gram matrix. To mitigate this, we introduce Ridge Regression to prevent overfitting. The weights s_{ij} are obtained by solving the following regularized least-squares problem:

$$\min_{s_i} \sum_i \left(\frac{1}{2} \left\| \log(\boldsymbol{X}_i) - \sum_{j=1}^{N} s_{ij} \log(\boldsymbol{X}_j) \right\|_F^2 + \frac{\mu}{2} \|\boldsymbol{s_i}\|_2^2 \right) \tag{5}$$

In this formulation, we enforce the constraint $s_{ii} = 0$ to preclude the trivial solution where a sample reconstructs itself. The regularization parameter μ balances the reconstruction error against the smoothness of the weight vector, thereby enhancing algorithm robustness when dealing with sparse or highly noisy data.

B. SPD-DNPE

Given a sample $\boldsymbol{X}_i \in S_{++}^D$ on the high-dimensional SPD manifold, the objective is to learn a mapping matrix $\boldsymbol{A}$ to obtain its projection $\boldsymbol{Y}_i \in S_{++}^d$ on the low-dimensional SPD manifold. Following Harandi et al. [15], the projection is performed via a bilateral projection, as follows:

$$\log(\boldsymbol{Y}_i) \approx \boldsymbol{A}^T \log(\boldsymbol{X}_i) \boldsymbol{A} \tag{6}$$

Since unsupervised NPE cannot leverage class labels, samples from distinct classes may overlap in the reduced space during dimensionality reduction, thereby degrading classification performance. To address this, SPD-DNPE introduces a Discriminative Constraint. The core tenet is to restrict neighborhood selection during graph construction exclusively to samples belonging to the same class, thereby ensuring the algorithm focuses on learning intra-class local reconstruction relationships. The schematic of this method is illustrated in Fig. 1. The optimization objective of SPD-DNPE is twofold: to minimize the local reconstruction error among intra-class samples while simultaneously maximizing the distance between centers of different classes. Drawing upon Linear Discriminant Analysis (LDA), we formulate a Generalized Rayleigh Quotient objective function.

Assume the dataset comprises C classes, where the c-th class contains N_c samples with a class center $\boldsymbol{Z}_c$. Let $\boldsymbol{Y}_i^c$ denote the low-dimensional embedding of the i-th sample $\boldsymbol{X}_i^c$ in class c, and let $\boldsymbol{F}_c$ represent the low-dimensional embedding of the class center $\boldsymbol{Z}_c$. The objective function is defined as follows:

$$\min_{\boldsymbol{A}} J(\boldsymbol{A}) = \frac{\sum_{c=1}^{C} \sum_{i=1}^{N_c} \left\| \log(\boldsymbol{Y}_i^c) - \sum_{j=1}^{N_c} w_{ij}^c \log(\boldsymbol{Y}_j^c) \right\|_F^2}{\sum_{i=1}^{C} \left\| \log(\boldsymbol{F}_i) - \sum_{j=1}^{C} b_{ij} \log(\boldsymbol{F}_j) \right\|_F^2} \tag{7}$$

Here, w_{ij}^c represents the intra-class local reconstruction weights for the c-th class, designed to preserve the compactness of intra-class data. Conversely, b_{ij} denotes the reconstruction weights between distinct class centers, serving to quantify the inter-class dispersion. The numerator of this objective function aims to preserve intra-class neighborhood structures by minimizing the reconstruction error, while the denominator seeks to maximize the separation between the centers of different classes, thereby enhancing the discriminative power of the features.

By substituting (6) into (7), the objective function can be further simplified as follows:

$$\min_{\boldsymbol{A}} \frac{\sum_{c=1}^{C} \sum_{i=1}^{N_c} \left\| \boldsymbol{A}^T \boldsymbol{G}_i^c \boldsymbol{A} \right\|_F^2}{\sum_{i=1}^{C} \left\| \boldsymbol{A}^T \boldsymbol{H}_i \boldsymbol{A} \right\|_F^2} \tag{8}$$

where the intra-class difference matrix is defined as $\boldsymbol{G}_i^c = \log(\boldsymbol{X}_i^c) - \sum_{j=1}^{N_c} w_{ij}^c \log(\boldsymbol{X}_j^c)$, and the inter-class difference matrix is given by $\boldsymbol{H}_i = \log(\boldsymbol{Z}_i) - \sum_{j=1}^{C} b_{ij} \log(\boldsymbol{Z}_j)$. Note that $\log(\boldsymbol{Z}_i) = \frac{1}{N_i} \sum_k \log(\boldsymbol{X}_k^i)$ represents the logarithmic mean of the i-th class. Leveraging the property of the Frobenius norm, $\|\boldsymbol{M}\|_F^2 = tr(\boldsymbol{M}^T \boldsymbol{M})$, and the symmetry of $\boldsymbol{G}_i^c$, the objective function of SPD-DNPE can be further reformulated as:

$$\min_{\boldsymbol{A}} J(\boldsymbol{A}) = \frac{\sum_{c=1}^{C} \sum_{i=1}^{N_c} tr\left(\boldsymbol{A}^T \boldsymbol{G}_i^c \boldsymbol{A} \boldsymbol{A}^T \boldsymbol{G}_i^c \boldsymbol{A}\right)}{\sum_{i=1}^{C} tr\left(\boldsymbol{A}^T \boldsymbol{H}_i \boldsymbol{A} \boldsymbol{A}^T \boldsymbol{H}_i \boldsymbol{A}\right)} \tag{9}$$

At this stage, the optimization variable $\boldsymbol{A}$ appears in a complex Quartic Form, rendering direct derivation for optimization computationally intractable. To address this, we introduce an Iterative Approximation strategy, which utilizes the projection matrix $\boldsymbol{A}^{(t-1)}$ obtained from the previous iteration to approximate the current intermediate term:

$$\boldsymbol{A}\boldsymbol{A}^T \approx \boldsymbol{A}^{(t-1)}(\boldsymbol{A}^{(t-1)})^T \tag{10}$$

Substituting (10) into (9), the iterative objective function of SPD-DNPE can be approximated as:

$$\min_{\boldsymbol{A}} J(\boldsymbol{A}) = \frac{tr\left(\boldsymbol{A}^T\left[\sum_{c=1}^{C}\sum_{i=1}^{N_c}\boldsymbol{G}_i^c\boldsymbol{A}^{(t-1)}(\boldsymbol{A}^{(t-1)})^T\boldsymbol{G}_i^c\right]\boldsymbol{A}\right)}{tr\left(\boldsymbol{A}^T\left[\sum_{i=1}^{C}\boldsymbol{H}_i\boldsymbol{A}^{(t-1)}(\boldsymbol{A}^{(t-1)})^T\boldsymbol{H}_i\right]\boldsymbol{A}\right)} \tag{11}$$

Consequently, we explicitly define the intra-class scatter matrix, $\boldsymbol{S}_w$, as:

$$\boldsymbol{S}_w = \sum_{c=1}^{C}\sum_{i=1}^{N_c}\boldsymbol{G}_i^c\boldsymbol{A}^{(t-1)}(\boldsymbol{A}^{(t-1)})^T\boldsymbol{G}_i^c \tag{12}$$

and the inter-class scatter matrix, $\boldsymbol{S}_b$, as:

$$\boldsymbol{S}_b = \sum_{i=1}^{C}\boldsymbol{H}_i\boldsymbol{A}^{(t-1)}(\boldsymbol{A}^{(t-1)})^T\boldsymbol{H}_i \tag{13}$$

Ultimately, following the iterative approximation process, the original non-convex optimization problem is transformed into a standard trace ratio problem:

$$\min_{\boldsymbol{A}} \frac{tr(\boldsymbol{A}^T\boldsymbol{S}_w\boldsymbol{A})}{tr(\boldsymbol{A}^T\boldsymbol{S}_b\boldsymbol{A})} \tag{14}$$

The optimal solution to this problem is obtained by solving the generalized eigenvalue problem $\boldsymbol{S}_w\mathbf{a} = \lambda\boldsymbol{S}_b\mathbf{a}$. The dimensionality reduction process of SPD-DNPE can be summarized in Algorithm 1.

C. Tangent Space Classifier Adaptation via CLE Framework

SPD-DNPE features reside in a non-Euclidean Riemannian space, creating a mismatch with Euclidean-based classifiers like SVM. We bridge this gap via the Calibrated Log-Euclidean (CLE) framework, ensuring a geometrically consistent and computationally efficient classification pipeline.

To prevent metric distortion, we utilize an isometric vectorization operator $vec_{iso}(\cdot)$, which scales off-diagonal elements by $\sqrt{2}$. This ensures the Euclidean norm of the vectorized feature remains equivalent to the Frobenius norm of the tangent vector:

$$\|vec_{iso}(\boldsymbol{S})\|_2 = \|\boldsymbol{S}\|_F \tag{15}$$

This guarantees distance invariance under the Riemannian metric within the vector space.

Furthermore, we introduce a Reference Point Calibration mechanism to minimize linearization errors. Rather than using the identity matrix, we adopt the Riemannian mean $\bar{\boldsymbol{Y}}$ of the training samples as the reference origin for tangent space mapping. The final calibrated feature representation is:

$$z = vec_{iso}(\log(\boldsymbol{Y}) - \log(\bar{\boldsymbol{Y}})) \tag{16}$$

Algorithm 1 Discriminant Neighborhood Preserving Embedding on SPD Manifold

Require: A set of labeled SPD matrices $\boldsymbol{X} = \{\boldsymbol{X}_1, \ldots, \boldsymbol{X}_N\}$ with labels $Y = \{y_1, \ldots, y_N\}$, where $\boldsymbol{X}_i \in S_{++}^D$; Target dimension d; Number of nearest neighbors k.
Ensure: Optimal mapping matrix $\boldsymbol{A} \in \mathbb{R}^{D\times d}$
1: Initialize the mapping matrix $\boldsymbol{A}^{(0)}$;
2: Graph Construction: For each sample, find k nearest neighbors within the same class to compute intra-class reconstruction weights w_{ij}^c ; Simultaneously, construct inter-class center weights b_{ij} using (4) or (5);
3: Preprocessing: Compute intra-class local difference matrix $\boldsymbol{G}_i$ and inter-class center difference matrix $\boldsymbol{H}_i$;
4: **while** not converged and not max epoch **do**
5: Update Scatter Matrices: Using the projection matrix $\boldsymbol{A}^{(t-1)}$ from the previous iteration, compute intra-class scatter matrix $\boldsymbol{S}_w$ using (12) and inter-class scatter matrix $\boldsymbol{S}_b$ using (13);
6: Solve Generalized Eigenvalue Problem: Solve the equation $\boldsymbol{S}_w\mathbf{a} = \lambda\boldsymbol{S}_b\mathbf{a}$;
7: Update Mapping Matrix $\boldsymbol{A}$: Form the new $\boldsymbol{A}^{(t)}$ by selecting the eigenvectors corresponding to the d smallest non-zero eigenvalues;
8: **end while**
9: **return** The final mapping matrix $\boldsymbol{A}$ and low-dimensional embeddings.

Through the aforementioned centering operation, samples originally residing on the curved low-dimensional SPD submanifold are mapped into Euclidean “calibrated vectors” possessing geometric fidelity. Consequently, these vectors can be directly input into classifiers—specifically Support Vector Machines (SVM) in this study—for efficient classification, thereby ensuring that the discriminative geometric structures learned by SPD-DNPE are maximally preserved and utilized during the classification stage.

IV. EXPERIMENT

To comprehensively validate the effectiveness of the proposed SPD-DNPE algorithm in Riemannian manifold dimensionality reduction and image classification tasks, as well as to verify the efficacy of the adaptive optimization framework, we designed a series of multi-dimensional comparative experiments; accordingly, this section will elaborate on the experimental settings and dataset characteristics, followed by a detailed analysis conducted from the perspectives of classification performance comparison, parameter sensitivity and geometric exploration, and a discussion of the results.

A. Experimental Setting

We employed four metrics to evaluate the classification performance of the algorithms: Accuracy, Precision, Recall, and F1-score. Superior algorithm performance is indicated by metric values approaching 1, whereas inferior performance is characterized by values approaching 0.

The target dimensionality, d, of the low-dimensional manifold samples is determined based on the cumulative contribution rate of the eigenvectors. For a given cumulative contribution threshold $\eta \in (0,1)$ and the maximum dimension D, the target dimension d is defined as:

$$d = \arg\min_{d^*} \left\{ 0 < d^* \leq D \;\middle|\; \sum_{i=1}^{d^*} \sigma_i \geq \eta \cdot \sum_{i=1}^{D} \sigma_i \right\} \tag{17}$$

where σ_i denotes the contribution rate of the i-th eigenvalue.

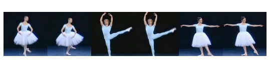
We denote the SPD-DNPE method utilizing Mode 1 as SPD-DNPE-I, and that utilizing Mode 2 as SPD-DNPE-II. The comparative study includes both shallow learning and deep learning approaches. The shallow learning methods include SPD-SVM, GNPE-I [6], GNPE-II [6], Lie-LPP [4], GDNPE-I [12], GDNPE-II [12], SPDML-AIM [15], and SPDML-Stein [15]. The deep learning methods include SPDNet [3], SymNet [16], and SPDNetBN [17].

Regarding implementation, for methods with publicly available code, we utilized their default parameter settings or officially recommended configurations. For approaches lacking available source code or requiring adaptation to SPD data, we reproduced the algorithms based on the pseudocode or mathematical derivations provided in their respective original papers. For all datasets, we randomly selected 50% of the samples as the training set, with the remaining samples serving as the test set. We executed our proposed algorithm and all comparative methods under this data partition protocol.

To validate the reliability and reproducibility of our results, we performed repeated experiments across different hardware platforms, yielding consistent and stable conclusions.

B. Datasets

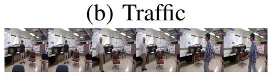
We evaluate our proposed method on four public datasets: Ballet, Traffic, UCF-Sports, and UT-Kinect. Representative sample images from these datasets are illustrated in Fig. 2, and detailed statistics are summarized in Table I.



(a) Ballet

(b) Traffic

(c) UCF-Sports



(d) UT-Kinect

Fig. 2: Dataset presentation

The Ballet dataset comprises 59 image sets derived from 44 instructional videos, characterized by significant intra-class variations in speed and appearance. The Traffic dataset includes 254 highway sequences captured under diverse weather conditions. UCF-Sports features 150 video sequences covering 13 action categories in natural scenes, while UT-Kinect provides 200 RGB-D and skeletal sequences from 10 subjects performing 10 actions. Regarding preprocessing, frames in Ballet, Traffic, and UCF-Sports are all normalized to 20×20 grayscale images.

C. Classification Performance

Table II presents the accuracy (ACC), precision (PRE), recall (REC), and F1-score achieved by SPD-DNPE-I, SPD-DNPE-II, and all baseline methods across the Ballet, Traffic, UCF-Sports, and UT-Kinect datasets.

Experimental results indicate that SPD-DNPE significantly outperforms traditional methods and competes favorably with deep learning baselines across most datasets. Specifically, SPD-DNPE-II establishes a dominant lead on the Traffic dataset, while SPD-DNPE-I achieves top performance on Ballet and UT-Kinect, verifying the framework's robustness.

A comparative analysis reveals distinct adaptability patterns: SPD-DNPE-I excels on datasets like Ballet (involving gestures) by leveraging sparse graph construction to preserve local geometric structures. Conversely, SPD-DNPE-II dominates on complex, noisy datasets like Traffic and UCF-Sports, where its fully connected strategy effectively utilizes global topological information to mitigate noise. This dual-mode design ensures the algorithm's flexibility in handling diverse manifold distributions.

TABLE I: Summary of Datasets

Name	**Samples**	**Train**	**Test**	**Classes**	$\boldsymbol{S}_d^+$
Ballet	59	31	28	8	400×400
Traffic	254	128	126	3	400×400
UCF-Sports	150	77	73	13	400×400
UT-Kinect	199	100	99	10	400×400

D. Parameters Analysis

The most critical parameter of the SPD-DNPE algorithm is the target dimension, d, of the low-dimensional SPD manifold, while the classifier hyperparameters (C, γ) determine the upper bound of the model's performance in image set classification tasks. We conducted a multi-stage experimental validation across various datasets to deeply elucidate the compatibility mechanisms between the intrinsic geometric characteristics of the data and the algorithm parameters. We select the Traffic dataset as the representative case for demonstration in this section. Among all datasets employed in this study, Traffic possesses the largest sample size and the most complex spatiotemporal data distribution, making it the most challenging and robust scenario for evaluating the algorithm's parameter sensitivity.

To investigate the impact of parameter settings on the performance of SPD-DNPE—specifically the influence of the Energy Preservation Ratio, η, on the determination of the intrinsic dimension d, as well as the applicability of different weighting construction modes Mode 1 and Mode 2—we evaluated SPD-DNPE using both weighting modes while traversing the cumulative contribution rate $\eta \in [0.50, 0.99]$. The results are recorded in Fig. 3. Experimental results indicate that the intrinsic dimension d increases significantly with the increase of η, and the specific characteristics of the data distribution dictate the selection of the optimal weighting mode.

TABLE II: Comparison of the Average Classification Performance across Different Datasets.

Method	Ballet				Traffic				UCF-Sports				UT-Kinect			
	ACC	PRE	F1	REC	ACC	PRE	F1	REC	ACC	PRE	F1	REC	ACC	PRE	F1	REC
SPD-SVM	0.286	0.196	0.158	0.276	0.735	0.715	0.664	0.731	0.228	0.164	0.140	0.235	0.556	0.567	0.545	0.557
GNPE-I	0.214	0.140	0.148	0.187	0.799	0.745	0.723	0.734	0.374	0.416	0.350	0.360	0.421	0.400	0.400	0.420
GNPE-II	0.321	0.233	0.231	0.313	0.804	0.725	0.707	0.715	0.452	0.469	0.415	0.424	0.428	0.433	0.405	0.429
Lie-LPP	0.450	0.398	0.3809	0.392	0.802	0.711	0.708	0.721	0.237	0.277	0.120	0.212	0.394	0.426	0.384	0.396
GDNPE-I	0.357	0.378	0.319	0.363	0.855	0.775	0.750	0.741	0.566	0.558	0.530	0.553	0.451	0.461	0.440	0.452
GDNPE-II	0.238	0.246	0.187	0.212	0.870	0.805	0.784	0.771	0.539	0.522	0.481	0.501	0.387	0.428	0.397	0.387
SPDML-AIM	0.642	0.672	0.594	0.608	0.889	0.824	0.815	0.810	0.424	0.500	0.445	0.432	0.475	0.494	0.477	0.477
SPDML-Stein	0.607	0.625	0.540	0.563	0.865	0.778	0.771	0.765	0.411	0.518	0.441	0.419	0.455	0.466	0.456	0.454
SPDNet	0.750	0.796	0.720	0.754	0.810	0.770	0.669	0.636	0.548	0.544	0.590	0.534	0.556	0.600	0.561	0.552
SymNet	0.679	0.773	0.653	0.650	0.817	0.728	0.679	0.652	0.523	0.525	0.478	0.498	0.515	0.541	0.509	0.516
SPDNetBN	0.771	0.843	0.757	0.767	0.912	0.919	0.913	0.913	0.575	0.582	0.563	0.600	0.535	0.559	0.543	0.540
SPD-DNPE-I	**0.786**	**0.857**	**0.784**	**0.775**	0.823	0.827	0.810	0.823	0.571	0.626	0.560	0.571	**0.761**	**0.818**	**0.755**	**0.761**
SPD-DNPE-II	0.619	0.655	0.594	0.617	**0.966**	**0.967**	**0.965**	**0.959**	**0.662**	**0.758**	**0.676**	**0.658**	0.660	0.708	0.649	0.661

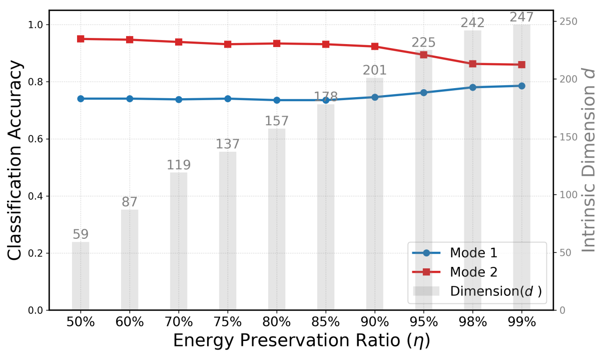


Fig. 3: Classification accuracy and intrinsic dimension (d) vs. energy preservation ratio (η).

It is observable that the estimated intrinsic dimension, d, increases significantly with the rise of η, and the high-dimensional nature of the data dictates the selection of the optimal mode. Specifically, Mode 2 exhibits exceptional adaptability and robustness on this dataset, maintaining a classification accuracy consistently above 90%, which is significantly superior to Mode 1, which hovers around 75%. This result strongly suggests that the fully connected strategy of Mode 2 more effectively captures the complex topological structures of the high-dimensional manifold inherent to the Traffic dataset.

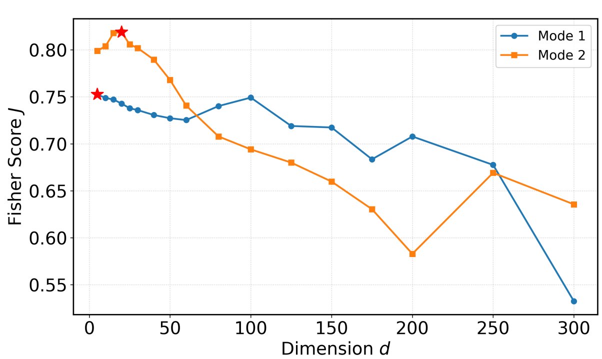


Fig. 4: Fisher Score (J) vs. dimension (d) for Mode 1 and Mode 2.

Following the preliminary determination of the effective intrinsic dimension range based on the energy threshold, we introduced the Fisher Score (J) as a discriminative screening metric to further pinpoint the optimal subspace dimension, d, that is most conducive to subsequent SVM classification, while simultaneously reducing the computational complexity of the hyperparameter search. As illustrated in Fig. 4, the J value for the Traffic dataset exhibits a typical "rise-then-fall" trend with respect to dimension variations. Notably, the performance of Mode 2 is significantly superior to that of Mode 1, achieving the highest Fisher Score at $d = 20$. This further confirms the applicability of Mode 2 to more complex data scenarios where sample neighborhoods may exhibit multicollinearity or instability.

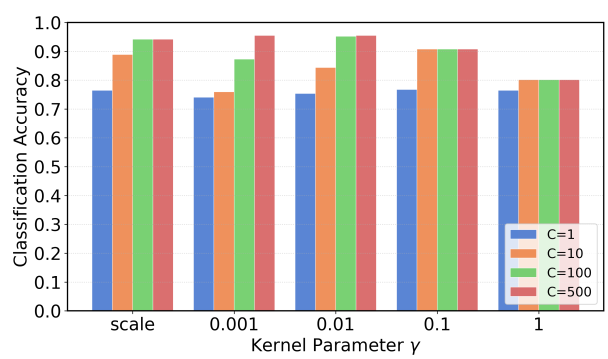


Fig. 5: Impact of SVM hyperparameters (C and γ) on classification accuracy.

Adhering to the principle of maximizing the Fisher Score, we fixed $d = 20$. Finally, to further exploit the classification potential within this discriminative subspace, we conducted a fine-grained grid search for the SVM regularization parameter C and the kernel parameter γ, with results presented in Fig. 5. It can be observed that the model performance is relatively sensitive to the value of C; as C increases, classification accuracy demonstrates a significant stepwise improvement. This indicates that the high-dimensional manifold data requires strong regularization constraints to optimize decision boundaries following low-dimensional projection, further attesting to the high discriminability of the SPD-DNPE method. Regarding the selection of kernel width, $\gamma = 0.01$ exhibited the best generalization capability. Ultimately, the model achieved global optimal performance under the configuration of $\gamma = 0.01$ and $C = 500$. Through the sequential determination of the optimal cumulative energy contribution rate (physical dimension truncation), Fisher Score optimization (discriminative subspace locking), and SVM hyperparameter tuning, the SPD-DNPE algorithm successfully achieved the optimal alignment between feature extraction and the classifier, thereby obtaining the highest classification accuracy.

E. Discussion

In contrast to deep learning approaches characterized by massive parameters, SPD-DNPE exhibits significantly lower model complexity, enabling efficient training and inference without the necessity for pre-training on large-scale datasets. This efficiency, coupled with its capability to precisely capture high-dimensional geometric structures, renders SPD-DNPE uniquely valuable for real-time action recognition tasks, particularly in scenarios involving limited computational resources or small sample sizes.

V. Conclusion

This paper proposes SPD-DNPE, a supervised dimensionality reduction algorithm based on the SPD manifold, which significantly enhances the performance of feature extraction and classification tasks for image sets and videos by leveraging Dual-Mode Neighborhood Weighting Strategies on SPD Manifolds and the CLE kernel strategy. Future work will be dedicated to addressing the challenge of label scarcity in practical applications by exploring the incorporation of semi-supervised and self-supervised learning mechanisms into the manifold dimensionality reduction process to fully exploit unlabeled data for bolstering the robustness of the manifold topological structure; concurrently, we plan to further investigate the deep integration of this Riemannian geometric theoretical framework with deep neural networks to construct end-to-end nonlinear feature learning models capable of uncovering deeper intrinsic data representations.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China 62376042, the Key Project of Chongqing Municipal Education Commission KJZD-K202100505, the Project of the Ministry of Education in China 20YJAZH084, the Chongqing Technology Innovation and Application Development Project under Grant cstc2020jscx-msxmX0190 and Grant cstc2019jscxmbdxX0061, the Natural Science Foundation of Chongqing CSTB2024NSCQ-KJFZZDX0036, the Key Special Project for Technological Innovation and Application Development in Chongqing CSTB2023TIAD-KPX0050 & CSTB2022TIAD-KPX0176.

References

[1] Jayasumana S, Hartley R, Salzmann M, et al. Kernel methods on the Riemannian manifold of symmetric positive definite matrices[C]//proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013: 73-80.

[2] Harandi M T, Salzmann M, Hartley R. From manifold to manifold: Geometry-aware dimensionality reduction for SPD matrices[C]//European conference on computer vision. Cham: Springer International Publishing, 2014: 17-32.

[3] Huang Z, Wan C, Probst T, et al. Deep learning on lie groups for skeleton-based action recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 6099-6108.

[4] Li Y, Lu R. Locality preserving projection on SPD matrix Lie group: algorithm and analysis[J]. Science China Information Sciences, 2018, 61: 1-15.

[5] Wang B, Hu Y, Gao J, et al. Locality preserving projections for Grassmann manifold[C]. Proceedings of the International Joint Conference on Artificial Intelligence. 2017:2893–2900.

[6] Wei D, Shen X, Sun Q, et al. Neighborhood preserving embedding on Grassmann manifold for image-set analysis[J]. Pattern Recognition, 2022, 122: 108335.

[7] Wang R, Guo H, Davis L S, et al. Covariance discriminative learning: A natural and efficient approach to image set classification[C]//2012 IEEE conference on computer vision and pattern recognition. IEEE, 2012: 2496-2503.

[8] Harandi M, Salzmann M, Hartley R. Joint dimensionality reduction and metric learning: A geometric take[C]//International Conference on Machine Learning. PMLR, 2017: 1404-1413.

[9] Zhu P, Cheng H, Hu Q, et al. Towards generalized and efficient metric learning on Riemannian manifold[C]. Proceedings of the International Joint Conference on Artificial Intelligence. 2018:3235–3241.

[10] Gao Z, Wu Y, Harandi M, et al. A robust distance measure for similarity-based classification on the SPD manifold[J]. IEEE transactions on neural networks and learning systems, 2019, 31(9): 3230-3244.

[11] Xu H. Unsupervised manifold learning with polynomial mapping on symmetric positive definite matrices[J]. Information Sciences, 2022, 609: 215-227.

[12] Li B, Zheng Y, Ran R, et al. Image-set classification using Discriminant Neighborhood Preserving Embedding on Grassmann manifold[J]. Signal Processing, 2025, 235: 110028.

[13] Ran R, Feng J, Zhang S, et al. A general matrix function dimensionality reduction framework and extension for manifold learning[J]. IEEE Trans. on Cybernetics, 2022, 52(4): 2137-2148.

[14] Ran R, Feng J, Li Z, et al. Locality preserving projections with autoencoder[J]. Expert Systems with Applications, 2024, 242: 122750.

[15] Harandi M, Salzmann M, Hartley R. Dimensionality reduction on SPD manifolds: The emergence of geometry-aware methods[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(1): 48–62.

[16] Wang R, Wu X J, Kittler J. SymNet: A simple symmetric positive definite manifold deep learning method for image set classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(5): 2208–2222.

[17] Brooks D, Schwander O, Barbaresco F, et al. Riemannian batch normalization for SPD neural networks[J]. Advances in Neural Information Processing Systems, 2019, 32.

[18] Wang R, Wu X J, Chen Z, et al. Dreamnet: A deep riemannian manifold network for spd matrix learning[C]//Proceedings of the Asian conference on computer vision. 2022: 3241-3257.

[19] Chen Z, Xu T, Wu X J, et al. Riemannian local mechanism for SPD neural networks[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2023, 37(6): 7104-7112.

[20] Wang R, Jin S, Chen Z, et al. Learning to Normalize on the SPD Manifold under Bures-Wasserstein Geometry[C]//Proceedings of the Computer Vision and Pattern Recognition Conference. 2025: 8289-8298.

[21] Wei D, Shen X, Sun Q, et al. Learning adaptive Grassmann neighbors for image-set analysis[J]. Expert Systems with Applications, 2024, 247: 123316.

[22] Zhou B, Zheng X, Wang Y G, et al. Embedding graphs on Grassmann manifold[J]. Neural Networks, 2022, 152: 322-331.

[23] Gao W, Ma Z, Xiong C, et al. Dimensionality reduction of SPD data based on Riemannian manifold tangent spaces and local affinity[J]. Applied Intelligence, 2023, 53(2): 1887-1911.