Spatial-Spectral Density Peaks Based Discriminant Projection for Classification of Membranous Nephropathy Hyperspectral Pathological Image [1]

Meng LV [a], Wei LI [a,2] and Ran TAO [a]

[a] *School of Information and Electronics, Beijing Institute of Technology, Beijing, China*

Abstract. Microscopic hyperspectral imaging has become an emerging technique for various medical applications. However, high dimensionality of hyperspectral image (HSI) makes image processing and extraction of important diagnostic information challenging. In this paper, a novel dimensionality reduction method named spatial-spectral density peaks based discriminant projection (SSDP) is proposed by considering spatial-spectral density distribution characteristics of immune complexes. The proposed SSDP coupled with support vector machine classifier (SVM) yields high-precision automatic diagnosis of membranous nephropathy (MN). Detailed ex-vivo validation of the proposed method demonstrates the potential clinical value of the system in identifying hepatitis B virus-associated membranous nephropathy (HBV-MN) and primary membranous nephropathy (PMN).

Keywords. Dimensionality reduction, feature extraction, microscopic hyperspectral imaging, membranous nephropathy diagnosis

1. Introduction

Chronic kidney disease (CKD) is a global public health problem [1] with an incidence rate of more than 10%. Early detection and treatment can often keep CKD from getting worse. Among CKDs, one of the most common pathological types of adult nephrotic syndrome is membranous nephropathy (MN) [2] which can be divided into primary membranous nephropathy (PMN) and secondary membranous nephropathy (SMN). Hepatitis B virus-associated membranous nephropathy (HBV-MN) is a popular variety of SMN. In clinical diagnosis, a clear distinction between HBV-MN and PMN is essential for MN identification. The gold standard for MN diagnosis remains tissue biopsy with pathological assessment made by pathologists using visual inspection of stained sections and immunouorescence under the optical microscope, combined with electron microscopy results. However, a certain probability of false positives is exist in the immunouores-

[2]Corresponding Author: Wei Li, 508 East, Information Science Experiment Building, Beijing Institute of Technology, Beijing; E-mail: liwei089@ieee.org.

cence and the accuracy of MN diagnosis is highly depended on the experience of the pathologists. This technique is considered expensive and time-consuming.

With the rapid development of hyperspectral cameras and artificial intelligence, hyperspectral imaging system has become promising auxiliary diagnostic tool for intelligent medicine. Differ from the bi-dimensionality image acquired by optical microscope which only contains morphological information of samples, HSI possesses abundant spatial-spectral information. Digital pathology, which utilizes the power of slide imaging and computer-aided diagnosis, has proved promising to provide rapid, consistent, and quantitative disease diagnosis from histopathology images [3]. Recently, local fisher discriminant analysis-deep neural network [4] reveals that HBV-MN and PMN have different immune complex components and show subtle spectral differences in hyperspectral images. Previous work also shows that dimensionality reduction (DR) is advantageous for acquiring an optimal reduced subspace with highly separability for features [4,5,6,7], thereby improving the accuracy of computer-based MN automatic clinical diagnosis.

In this paper, we propose a novel DR method based on density peaks (DP) based clustering algorithm [8] to make full use of the spatial-spectral density distribution characteristics of immune complexes. At first, spectral density and spatial density are defined for each pixel based on the density peaks based clustering algorithm. Then, a spatial density scatter matrix is designed to preserve the neighborhood structure in spatial domain and two spectral density scatter matrices are constructed to describe the local discriminant relationship in spectral domain. At last, a spatial-spectral density peaks based discriminant projection (SSDP) method is proposed by compacting the spatial-spectral intra-class scatter while separating the spectral interclass scatter. SSDP efficiently preserves the local intra-class local structure in HSI and enhances the discrimination power of low-dimensional embedding spatial-spectral features.

2. Proposed SSDP Method

The proposed SSDP method consists of three parts: construction of spatial density based scatter matrix, construction of spectral density based scatter matrices and the proposed spatial-spectral discriminant projection model. More details of these parts are described as follows.

2.1. Spatial density based scatter matrix

Let (a_i, b_i) denotes the coordinate of hyperspectral pixel $\mathbf{x}_i$, the local spatial neighborhood $\Omega\left(\mathbf{x}_i\right)$ centered at $\mathbf{x}_i$ is defined as:

$$\Omega\left(\mathbf{x}_i\right) = \left\{ \mathbf{x} \leftarrow (a, b) \,\middle|\, \begin{array}{l} a \in [a_i - q, a_i + q] \\ b \in [b_i - q, b_i + q] \end{array} \right\} \tag{1}$$

where $q = (T-1)/2$, and the odd number T is the size of spatial neighborhood. $\mathbf{x} \leftarrow (a, b)$ indicates that the sample locate in corresponding coordinate.

Based on the idea of density peaks based clustering algorithm, the spatial density of pixel $\mathbf{x}_i$ is defined as

$$\rho_i^{spatial} = \sum_{\mathbf{x}_j \in \Omega(\mathbf{x}_i)} \chi\left(l\left(\mathbf{x}_i\right) - l\left(\mathbf{x}_j\right)\right) \tag{2}$$

where $\chi(x) = 1$ if $x = 0$ and $\chi(x) = 0$ otherwise, and $l(\mathbf{x}_i)$ is the label of $\mathbf{x}_i$. Basically, ρ_i is equals to the number of points from $\Omega(\mathbf{x}_i)$ that have the same label with $\mathbf{x}_i$.

For training sample $\mathbf{x}_i$, the spatial density weight between $\mathbf{x}_i$ and $\mathbf{x}_j$ is defined as

$$\mathbf{W}_{ij}^{spatial} = \begin{cases} \frac{\rho_i^{spatial} \times \rho_j^{spatial}}{t^{spatial}}, & \text{if } l(\mathbf{x}_i) = l(\mathbf{x}_j) \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

in which $t^{spatial} = \max\limits_{i,j}(\rho_i^{spatial} \times \rho_j^{spatial})$. The density weight makes the points with lower spatial density not produce larger influence to the embedded space.

The intra-class spatial density based scatter matrix is defined as

$$\mathbf{S}^{sw} = \frac{1}{2} \sum_{i,j=1}^{n} \mathbf{W}_{ij}^{spatial} (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T \tag{4}$$

2.2. Spectral density based scatter matrices

Based on the idea of density peaks based clustering algorithm, the spectral density of pixel $\mathbf{x}_i$ is defined as same as [8]. That is, $\rho_i^{spectral} = \sum\limits_{j} \chi(\mathbf{d}_{ij} - d_c)$ where $\mathbf{d}_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|_2^2$, $\chi(x) = 1$ if $x = 0$ and $\chi(x) = 0$ otherwise.

For training sample $\mathbf{x}_i$, the spectral density weight between $\mathbf{x}_i$ and $\mathbf{x}_j$ is defined as

$$\mathbf{W}_{ij}^{spectral} = \frac{\rho_i^{spectral} \times \rho_j^{spectral}}{t^{spectral}} \tag{5}$$

in which $t^{spectral} = \max\limits_{i,j}(\rho_i^{spectral} \times \rho_j^{spectral})$. The density weight makes the points with lower spectral density not produce larger influence to the embedded space.

In spectral domain, the inter-class density scatter matrix $\mathbf{S}^{(lb)}$ and the intra-class density scatter matrix $\mathbf{S}^{(lw)}$ are defined as

$$\mathbf{S}^{(lb)} = \frac{1}{2} \sum_{i,j=1}^{n} \hat{\mathbf{W}}_{ij}^{(lb)} (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T \tag{6}$$

$$\mathbf{S}^{(lw)} = \frac{1}{2} \sum_{i,j=1}^{n} \hat{\mathbf{W}}_{ij}^{(lw)} (\mathbf{x}_i - \mathbf{x}_j)(\mathbf{x}_i - \mathbf{x}_j)^T \tag{7}$$

where $\hat{\mathbf{W}}^{(lb)}$ and $\hat{\mathbf{W}}^{(lw)}$ are $n \times n$ matrices defined by

$$\hat{\mathbf{W}}_{ij}^{(lb)} = \begin{cases} \mathbf{W}_{ij}^{spectral} \big/ (1/n - 1/n_c), & \text{if } l(\mathbf{x}_i) = l(\mathbf{x}_j) = c \\ 1/n, & \text{otherwise.} \end{cases} \tag{8}$$

$$\hat{\mathbf{W}}_{ij}^{(lw)} = \begin{cases} \mathbf{W}_{ij}^{spectral} \big/ n_c, & \text{if } l(\mathbf{x}_i) = l(\mathbf{x}_j) = c \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

Here $c \in [1, C]$, C represents the total number of classes and n_c denotes the number of samples belong to class c.

2.3. *Spatial-Spectral Density Peaks Based Discriminant Projection*

In this section, spatial-spectral density peaks based discriminant projection (SSDP) approach is proposed by minimizing the spatial-spectral intra-class compactness and maximizing spectral inter-class variance in embedded subspace. The intra-class spatial margin , the intra-class spectral margin and the inter-class spectral margin in embedded space can be described by

$$
\begin{cases}
\frac{1}{2} \sum\limits_{i,j=1}^{n} \left\| \mathbf{P}^T \mathbf{x}_i - \mathbf{P}^T \mathbf{x}_j \right\|_2^2 \mathbf{W}_{ik}^{spatial} = tr \left(\mathbf{P}^T \mathbf{S}^{sw} \mathbf{P} \right) \\
\frac{1}{2} \sum\limits_{i,j=1}^{n} \left\| \mathbf{P}^T \mathbf{x}_i - \mathbf{P}^T \mathbf{x}_j \right\|_2^2 \hat{\mathbf{W}}_{ij}^{(lw)} \ = tr \left(\mathbf{P}^T \mathbf{S}^{lw} \mathbf{P} \right) \\
\frac{1}{2} \sum\limits_{i,j=1}^{n} \left\| \mathbf{P}^T \mathbf{x}_i - \mathbf{P}^T \mathbf{x}_j \right\|_2^2 \hat{\mathbf{W}}_{ij}^{(lb)} \ = tr \left(\mathbf{P}^T \mathbf{S}^{lb} \mathbf{P} \right)
\end{cases}
\tag{10}
$$

The model of SSDP is defined as

$$
\max_{\mathbf{P}} \frac{tr \left(\mathbf{P}^T \mathbf{S}^{lb} \mathbf{P} \right)}{tr \left\{ \mathbf{P}^T \left[\alpha \mathbf{S}^{sw} + (1 - \alpha) \mathbf{S}^{lw} \right] \mathbf{P} \right\}}
\tag{11}
$$

where $\alpha \in [0, 1]$ is a tradeoff balancing the contribution of spatial and spectral information in projection process. The solution of the optimizing problem can be translated to solve the generalized eigenvalue problem

$$
\mathbf{S}^{lb} \mathbf{p}_i = \lambda_i \left[\alpha \mathbf{S}^{sw} + (1 - \alpha) \mathbf{S}^{lw} \right] \mathbf{p}_i
\tag{12}
$$

where λ_i is an eigenvalue and $\mathbf{p}_i$ represents the corresponding eigenvector. The projected matrix $\mathbf{P} = [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_d] \in \Re^{D \times d}$ is constructed by the first d largest eigenvalues.

SSDP makes fully use of spatial-spectral density based properties in hyeprspectral images to map the original data into a low-dimensional space, in which not only strengthen the class separability but also preserve the spatial-spectral relationship between pixels as much as possible.

3. **Experimental Results and Analysis**

In this section, real microscopic hyperspectral dataset is used to evaluate the effectiveness of the proposed approach for MN identification tasks. The dataset consists of 30 HBV-MN images and 24 PMN images, involving 10 HBV-MN patients and 9 PMN patients respectively. The performance of DR can be evaluated with classification applied to the data resulting from DR, and we use SVM due to its popularity. Three objective quality indexes (i.e., each accuracy (EA), overall accuracy (OA), average accuracy (AA))are used to evaluate the performance of hyperspectral image classification. Throughout, we report results for a variety of state-of-the-art DR methods, comparing SSDP to LFDA [9], EFDC [10], NWFE [11], BCGDA [12] as well as CDME [13].

The identification of HBV-MN and PMN is performed in fully-independent patients, divided across 90 terms. Each term consist of one HBV-MN patient and one PMN patient for training, while testing and validation is performed on remaining patients. Figure 1 shows the average OA of 90 tests. SSDP is the best performing image dimensionality
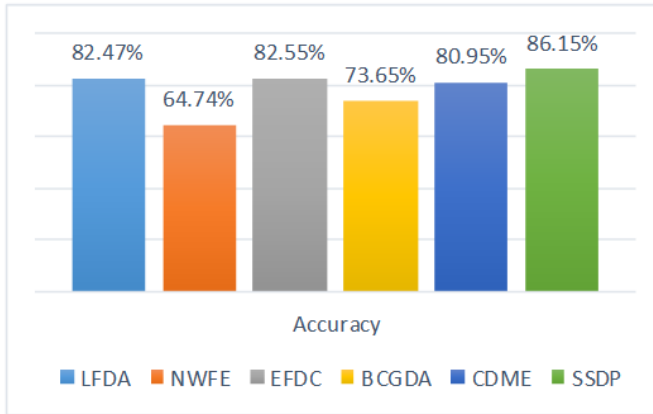
Figure 1. Average over all accuracy of MN identification.

reduction modality with accuracy of 86.15%, which meets the requirement of clinical diagnostic accuracy higher than 85%. Take one term data (one HBV-MN (ID:17472) patient and one PMN (ID:15684) patient are used as training data) as an example, we use visualization techniques to illustrate the effectiveness of SSDP for acquiring a subspace with better separability. Figure 2 shows the distribution of testing sample features before and after different DR methods process, the results confirm the potentiality of SSDP for seeking optimal subspace with high separability for features. In addition, local manifold based DR methods achieve much better performance than reconstruction based DR methods. This indicates that the potential local distribution characteristics of the immune complexes of MN are crucial for improving the discrimination of different diseases.

To further investigate the classification efficiency of the proposed algorithm as compared to other DR approaches, Table 1 and Table 2 tabulate the EA, OA and AA obtained by 1-dimensional data and 10-dimensional data respectively. In order to facilitate identification, the highest accuracy is signed on the bold. Obviously, the improvement effect of SSDP on classification accuracy is significantly better than all other DR methods. Especially, the superiority is also obvious as depicted Table 1, the proposed method still has over 6.4% increase in classification accuracy with limited data dimensionality. In summary, it proves that SSDP has the ability to extract important diagnostic information of medical hyperspectral image of MN, and can compress important diagnostic information in very limit low-dimension data.Our work provides a novel technique for the characterization and distinction of HBV-MN from IMN, especially for cases where the discrimination is not always easy with the optical microscopy, and verifies its nonnegligible potential for further application in medical field.

4. CONCLUSIONS

In this paper, a novel discriminant spatial-spectral dimensionality reduction method was proposed for medical hyperspectral image analysis. By considering the density distribution characteristics of the data, spectral density and spatial density are defined for each pixel based on the density peaks based clustering algorithm. SSDP has been confirmed to be powerful in finding the discriminative embedding space by considering spatial-
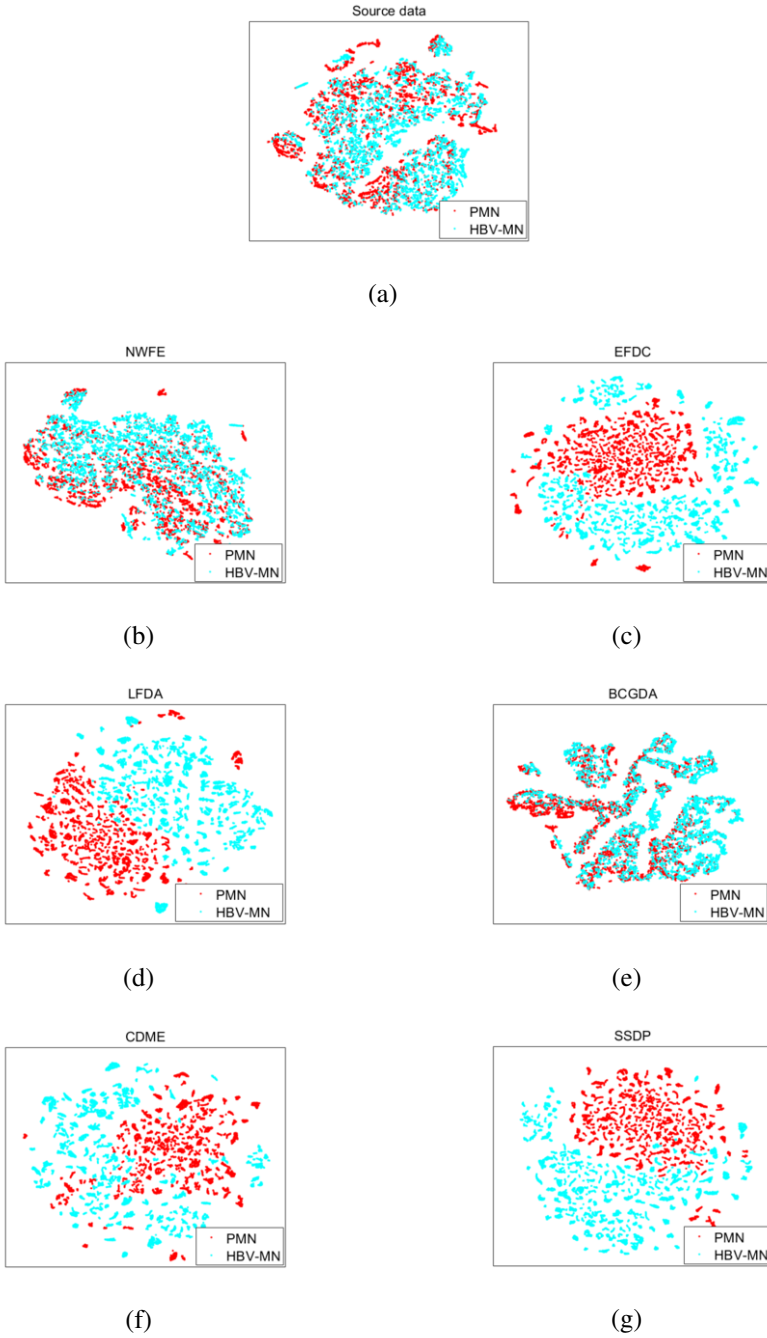
Figure 2. The distribution of (a) source data and low-dimension samples obtained by (b) NWFE, (c) EFDC, (d) LFDA, (e) BCGDA, (f) CDME and (g) SSDP respectively.

Table 1. Classification Accuracy with 1-Dimensional Data

Methods	HBV-MN(%)	PMN(%)	OA(%)	AA(%)
ORG	74.13	85.8	80.43	79.97
LFDA	49.31	44.72	46.83	47.01
EFDC	92.16	92.04	92.09	92.10
BCGDA	65.96	92.02	80.01	77.99
CDME	69.17	81.81	75.99	75.49
SSDP	**99.32**	**97.52**	**98.49**	**98.42**

Table 2. Classification Accuracy with 10-Dimensional Data

Methods	HBV-MN(%)	PMN(%)	OA(%)	AA(%)
ORG	74.13	85.8	80.43	79.97
LFDA	90.15	97.19	93.95	93.67
EFDC	92.20	95.58	94.02	93.89
BCGDA	71.40	89.95	81.41	80.68
CDME	92.55	95.47	94.13	94.01
SSDP	**99.36**	**97.85**	**98.66**	**98.61**

spectral intra-class structure and spectral inter-class structure simultaneously. The relative merits of our approach against other existing DR methods are also demonstrated in this paper, highlighting the strengths of our proposed framework in terms of MN identification.

References

[1] J. Xu, H. Lei, S. Zhang and D. O. Nephrology, Major research advances in chronic kidney disease of 2017, *Clinical Focus*, 2018.

[2] J. A. J. G. van den Brand, J. M. Hofstra and J. F. M. Wetzels, Low-Molecular-Weight Proteins as Prognostic Markers in Idiopathic Membranous Nephropathy, *Clinical Journal of the American Society of Necrology Cjasn* **6** (2011), 2846-2853.

[3] G. Lu, D. Wang, X. Qin and S. Muller, J. V. Little, X. Wang, A. Y. Chen, G. Chen and B. Feis: Histopathology Feature Mining and Association with Hyperspectral Imaging for the Detection of Squamous Neoplasia, *Scientific Reports* **9** (2019), 17863.

[4] X. Wei, T. Tu, N. Zhang, Y. Yang and W. Li, Membranous Nephropathy Identification Using Hyperspectral Microscopic Images, *Pattern Recognition and Computer Vision - Second Chinese Conference*, Xi'an, China, 2019.

[5] D. Ravi, H. Fabelo, G. M. Callic and G. Yang, Manifold Embedding and Semantic Segmentation for Intraoperative Guidance With Hyperspectral Brain Imaging, *IEEE Transactions on Medical Imaging* **36** (2017), 1845-1857.

[6] T. Li and G. Kou and Y. Peng, Improving malicious URLs detection via feature engineering: Linear and nonlinear space transformation methods, *Information Systems* **91** (2020), 101494.

[7] G. Kou, P. Yang, Y. Peng, F. Xiao, Y. Chen and F. E. Alsaadi, Evaluation of feature selection methods for text classification with small datasets using multiple criteria decision-making methods, *Applied Soft Computing* **86** (2020), 105836.

[8] A. Rodriguez and A. Laio, Clustering by fast search and find of density peaks, *Science* **344** (2014), 1492-1496.

[9] W. Li, S. Prasad, J. E. Fowler and L. M. Bruce, Locality-Preserving Dimensionality Reduction and Classification for Hyperspectral Image Analysis, *IEEE Transactions on Geoscience and Remote Sensing* **50** (2012), 1185-1198.

[10] Q. Gao, J. Liu, H. Zhang, J. Hou and X. Yang, Enhanced fisher discriminant criterion for image recognition, *Pattern Recognition* **45** (2012), 3717-3724.

[11] B. C. Kuo and D. A. Landgrebe, Nonparametric weighted feature extraction for classification, *IEEE Transactions on Geoscience and Remote Sensing* **42** (2004), 1096 - 1105.

[12] N. H. Ly, Q. Du and J. E. Fowler, Collaborative Graph-Based Discriminant Analysis for Hyperspectral Imagery, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **7** (2014), 2688-2696.

[13] M. Lv, Q. Hou and N. Deng and L. Jing, Collaborative Discriminative Manifold Embedding for Hyperspectral Imagery, *IEEE Geoscience and Remote Sensing Letters* **14** (2017), 569 - 573.