

HKBU at MediaEval 2017

Medico: Medical Multimedia Task

Yang Liu^{1,2}, Zhonglei Gu¹, William K. Cheung¹

¹Department of Computer Science, Hong Kong Baptist University, Hong Kong SAR, China

²Institute of Research and Continuing Education, Hong Kong Baptist University, Shenzhen, China
csygliu@comp.hkbu.edu.hk, cszlg@comp.hkbu.edu.hk, william@comp.hkbu.edu.hk

ABSTRACT

In this paper, we describe our model designed for automatic detection of diseases based on multimedia data collected in hospitals. Specifically, a two-stage learning strategy is designed to predict the diseases. In the first stage, a dimensionality reduction method called bidirectional marginal Fisher analysis (BMFA) is proposed to project the original data to the low-dimensional space, with the key discriminant information being well preserved. In the second stage, the multi-class support vector machine (SVM) is utilized on the low-dimensional space for detection. Experimental results demonstrate the efficiency of designed model.

1 INTRODUCTION

We have evidenced an increasing trend of applying multimedia processing and analysis methods, such as computer vision and medical image processing, to assisting diagnosis, detection and interpretation of medical abnormalities [7]. While there exist a number of successful applications related to healthcare decision support, predictive analytics are still considered challenging for some specific medical multimedia data like endoscopy and ultrasound images, due to the high complexity of these medical tasks [8]. Moreover, most of the existing methods make use of a limited amount of information, where possibly useful information sources such as sensory and temporal information were mostly not considered [5, 10, 11].

The 2017 MediaEval Medico Task aims to improve detection and location of abnormalities through designing an integrated approach combining information from both video and image information, as well as other sensory information and assistance of medical experts. The dataset used in this task consists of 8,000 GI tract images that are annotated and verified by medical doctors for the ground truth. More details about the task requirements and the dataset can be found in [8].

In order to perform efficient detection of diseases, we design a two-stage learning strategy. In the first stage, a manifold learning method called bidirectional marginal Fisher analysis (BMFA) is proposed to learn the compact representation of the original data, with the key discriminant information being well preserved. In the second stage, the multi-class support vector machine (SVM) is applied to the compact representation of each data point for detection.

2 METHOD

2.1 Bidirectional Marginal Fisher Analysis

Given the data matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$, where n denotes the number of data points in the training set, $\mathbf{x}_i \in \mathbb{R}^D$ the high-dimensional feature vector of the i -th data point, $l_i \in \{1, 2, \dots, 8\}$ the set of possible labels, and $\mathbf{l} = [l_1, l_2, \dots, l_n]$ the set of true labels corresponding to each data point $\mathbf{x}_i$. The proposed bidirectional marginal fisher analysis (BMFA) aims to learn the subspace of original features, in which the marginal discriminant information is well preserved.

A manifold learning algorithm typically involves the definition of a neighborhood graph. The effectiveness of the algorithm very often depends on how well the defined neighborhood graph can preserve the local topological properties of the data manifold. To achieve a robust neighborhood graph, we use the following two-way connection criterion to construct the graph [4]: we connect $\mathbf{x}_i$ and $\mathbf{x}_j$ only if $\mathbf{x}_i$ is one of the K -nearest neighbors of $\mathbf{x}_j$, and, $\mathbf{x}_j$ is also one of the K -nearest neighbors of $\mathbf{x}_i$. The above criterion adopts the “and” hypothesis, which means it agrees to connect two data points if and only if both of them are neighbors of each other.

The within-class adjacency matrix and between-class adjacency matrix are then defined as follows:

$$\mathbf{A}_{ij}^w = \begin{cases} \exp(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_F^2}{2\sigma}), & \text{if } j \in \mathcal{N}_i \text{ and } i \in \mathcal{N}_j \text{ and } l_i = l_j, \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$\mathbf{A}_{ij}^b = \begin{cases} \exp(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_F^2}{2\sigma}), & \text{if } j \in \mathcal{N}_i \text{ and } i \in \mathcal{N}_j \text{ and } l_i \neq l_j, \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where $\mathcal{N}_i$ denotes the index set of the K nearest neighbors of $\mathbf{x}_i$, and σ is empirically set by $\sigma = \sum_{i=1}^n \|\mathbf{x}_i - \mathbf{x}_{i_K}\|_F^2 / n$ where $\mathbf{x}_{i_K}$ is the K th nearest neighbor of $\mathbf{x}_i$.

BMFA aims to learn a transformation matrix $\mathbf{W} \in \mathbb{R}^{D \times d}$ where the following objective function is maximized:

$$\mathbf{W} = \arg \max_{\mathbf{W}} \frac{\sum_{i,j=1}^n \mathbf{A}_{ij}^b \|\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j\|_F^2}{\sum_{i,j=1}^n \mathbf{A}_{ij}^w \|\mathbf{W}^T \mathbf{x}_i - \mathbf{W}^T \mathbf{x}_j\|_F^2}. \quad (3)$$

The above optimization problem can be relaxed to the following ratio trace problem:

$$\mathbf{W} = \arg \max_{\mathbf{W}} \text{tr}(\frac{\mathbf{W}^T \mathbf{S}^b \mathbf{W}}{\mathbf{W}^T \mathbf{S}^w \mathbf{W}}), \quad (4)$$

where $\mathbf{S}^b = \mathbf{X} \mathbf{L}^b \mathbf{X}^T$ and $\mathbf{S}^w = \mathbf{X} \mathbf{L}^w \mathbf{X}^T$, in which $\mathbf{L}^b = \mathbf{D}^b - \mathbf{A}^b$ and $\mathbf{L}^w = \mathbf{D}^w - \mathbf{A}^w$ are the Laplacian matrices [1, 3], and $\mathbf{D}^b$ and $\mathbf{D}^w$ are diagonal matrices defined as $\mathbf{D}_{ii}^b = \sum_{j=1}^n \mathbf{A}_{ij}^b$

Table 1: Official evaluation results (provided by the organizers) of the proposed model on the Kvasir dataset.

	Recall	Specificity	Precision	Accuracy	F1	MCC	R_k statistic
Run 1 for both subtasks	0.6975	0.9568	0.6975	0.9244	0.6975	0.6543	0.6571
Run 2 for both subtasks	0.7028	0.9575	0.7028	0.9257	0.7028	0.6603	0.6626
Run 3 for both subtasks	0.6890	0.9556	0.6890	0.9223	0.6890	0.6446	0.6453
Run 4 for both subtasks	0.6988	0.9570	0.6988	0.9247	0.6988	0.6557	0.6585
Run 5 for both subtasks	0.6918	0.9560	0.6918	0.9229	0.6918	0.6477	0.6483

and $\mathbf{D}_{ii}^w = \sum_{j=1}^n \mathbf{A}_{ij}^w$ ($i = 1, \dots, n$). The optimization problem in Eq. (4) could be solved by generalized eigen-decomposition.

For each high-dimensional data point $\mathbf{x}_i$, we can obtain its low-dimensional representation by $\mathbf{y}_i = \mathbf{W}^T \mathbf{x}_i$.

2.2 Support Vector Classification

To predict diseases, we apply the multi-class support vector machine (SVM) to the low-dimensional representation of each data point. Specifically, we employ the one-against-all strategy for the multi-class classification and use the ν -SVM [9] on each class.

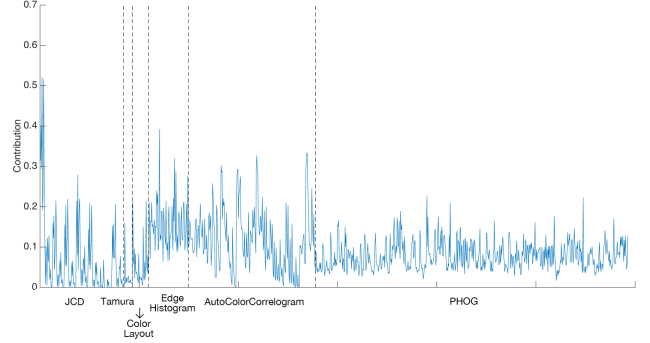
3 RESULTS

In this section, we report the experimental settings and the evaluation results. The Kvasir Dataset is used to evaluate the performance of the proposed model [6]. The size of the training set is 4,000 (500 for each of the eight classes) and the dimension of the original feature space is 1,185. We employ the same setting on both the detection subtask and the efficient detection subtask (i.e., the 5 runs of the detection subtask and those of the efficient detection subtask are identical). Specifically:

- For Run 1 of both subtasks, we randomly select 1,000 data samples (125 for each class) for training, and project the original data to the 8-dimensional subspace via BMFA;
- For Run 2 of both subtasks, we randomly select 1,000 data samples for training, and project the original data to the 9-dimensional subspace via BMFA;
- For Run 3 of both subtasks, we randomly select 800 data samples (100 for each class) for training, and project the original data to the 9-dimensional subspace via BMFA;
- For Run 4 of both subtasks, we randomly select 1,000 data samples for training, and project the original data to the 7-dimensional subspace via BMFA;
- For Run 5 of both subtasks, we randomly select 800 data samples for training, and project the original data to the 7-dimensional subspace via BMFA.

In the following step, we use the LIBSVM toolbox for classification [2]. The ν -SVC with linear kernel is selected, and the parameters are set as $\nu = 0.2$ and $cost = 1$.

Table 1 shows the official evaluation results (provided by the task organizers) of the proposed model on the Kvasir dataset. From the table, we can see that the proposed model performs well, given the task is a multi-class classification

**Figure 1: Contribution of each individual feature.**

with only a portion of the available data being utilized for training. This indicates that the learned subspace is able to capture the discriminant information of the dataset. Moreover, by reducing the original data to a very low-dimensional subspace, the learning process becomes more efficient.

In addition to the overall performance, we analyze the contribution of each dimension in the original feature space. The contribution of the i -th dimension is defined as $Contribution_i = \sum_j \lambda_j |w_{ij}|$, where λ_j denotes the j -th eigenvalue, w_{ij} denotes the (i, j) -th element of $\mathbf{W}$, and $|\cdot|$ denotes the absolute value operator. From Figure 1 we can see that the edge histogram and auto color correlogram features contribute more than the others, which indicates their importance in medical image classification.

4 CONCLUSION

This paper described the model designed for disease detection based on multimedia data. A novel dimensionality reduction algorithm dubbed bidirectional marginal Fisher analysis (BMFA) was presented to extract the discriminant information from the original feature space. After that, SVM was applied to the low-dimensional subspace for multi-class classification. Experimental results on the Kvasir dataset validated the effectiveness and efficiency of the proposed model.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 61503317, and in part by the Faculty Research Grant of Hong Kong Baptist University (HKBK) under Project FRG2/16-17/032.

REFERENCES

- [1] M. Belkin and P. Niyogi. 2001. Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering. In *Advances in Neural Information Processing Systems 14 (NIPS)*. 585–591.
- [2] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2 (2011), 27:1–27:27. Issue 3.
- [3] X. He and P. Niyogi. 2003. Locality Preserving Projections. In *Advances in Neural Information Processing Systems 16 (NIPS)*. 153–160.
- [4] Yang Liu, Yan Liu, and Keith C. C. Chan. 2011. Ordinal regression via manifold learning. In *Proceeding of the 25th AAAI Conference on Artificial Intelligence (AAAI)*. 398–403.
- [5] Konstantin Pogorelov, Sigrun Losada Eskeland, Thomas de Lange, Carsten Griwodz, Kristin Ranheim Randel, Håkon Kvale Stensland, Duc-Tien Dang-Nguyen, Concetto Spampinato, Dag Johansen, Michael Riegler, and Pål Halvorsen. 2017. A Holistic Multimedia System for Gastrointestinal Tract Disease Detection. In *Proceedings of the 8th ACM on Multimedia Systems Conference (MMSYS)*. 112–123.
- [6] Konstantin Pogorelov, Kristin Ranheim Randel, Carsten Griwodz, Sigrun Losada Eskeland, Thomas de Lange, Dag Johansen, Concetto Spampinato, Duc-Tien Dang-Nguyen, Mathias Lux, Peter Thelin Schmidt, Michael Riegler, and Pål Halvorsen. 2017. Kvasir: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection. In *Proceedings of the 8th ACM on Multimedia Systems Conference (MMSYS)*. 164–169.
- [7] Michael Riegler, Mathias Lux, Carsten Griwodz, Concetto Spampinato, Thomas de Lange, Sigrun L. Eskeland, Konstantin Pogorelov, Wallapak Tavanapong, Peter T. Schmidt, Cathal Gurrin, Dag Johansen, Håvard Johansen, and Pål Halvorsen. 2016. Multimedia and Medicine: Teammates for Better Disease Detection and Survival. In *Proceedings of the 2016 ACM on Multimedia Conference (ACMMM)*. 968–977.
- [8] Michael Riegler, Konstantin Pogorelov, Pål Halvorsen, Carsten Griwodz, Thomas de Lange, Kristin Ranheim Randel, Sigrun Losada Eskeland, Duc-Tien Dang-Nguyen, Mathias Lux, and Concetto Spampinato. 2017. Multimedia for Medicine: The Medico Task at MediaEval 2017. In *Working Notes Proceedings of the MediaEval 2017 Workshop*.
- [9] Bernhard Schölkopf, Alex J. Smola, Robert C. Williamson, and Peter L. Bartlett. 2000. New Support Vector Algorithms. *Neural Comput.* 12, 5 (2000), 1207–1245.
- [10] Y. Wang, W. Tavanapong, J. Wong, J. Oh, and P. C. de Groen. 2011. Computer-aided detection of retroflexion in colonoscopy. In *Proceedings of 24th International Symposium on Computer-Based Medical Systems (CBMS)*. 1–6.
- [11] Yi Wang, Wallapak Tavanapong, Johnny Wong, Jung Hwan Oh, and Piet C. de Groen. 2015. Polyp-Alert: Near real-time feedback during colonoscopy. *Computer Methods and Programs in Biomedicine* 120, 3 (2015), 164–179.