Research on Biomedical Image Analysis

Kaikai Zhao zhaojiayi11@nudt.edu.cn

1 Overview

At present, I work as a PhD candidate in the Machine Learning Group of Department of Computer Science and Technology, National University of Defense Technology. I have been involved in the research about biomedical image analysis for half a year. In this document, I summarize our research work on how we apply machine learning algorithms, including deep learning, to bioinformatics problems. Specifically, we have proposed a new framework which could predict tuberculosis(TB) drug resistance automatically and presented a high accuracy automatic HEp-2 classification method with small datasets via VGGNET. Also, I present some details about our implementation in Sect.4. Meanwhile, I list my research interests and specific goals at the end of this document.

2 Tuberculosis multi-drug resistant(MDR) detection based on CT image analysis

2.1 Introduction

About 130 years after the discovery of Mycobacterium tuberculosis, the disease remains a persistent threat and a leading cause of death worldwide. The greatest disaster that can happen to a patient with tuberculosis (TB) is that the organisms become resistant to two or more of the standard drugs. In contrast to drug sensitive (DS) tuberculosis, its multi-drug resistant (MDR) form is much more difficult and expensive to recover from. Thus, early detection of the drug resistance (DR) status is of great importance for effective treatment. The most commonly used methods of DR detection are either expensive or take too much time (up to several months). Therefore there is a need for quick and at the same time cheap methods of DR detection. One of the possible approaches for this task is based on Computed Tomography (CT) image analysis.

The Table 1, Fig.1 and Fig.2 are a brief description about the Task 1 dataset in the ImageCLEF2017 competition. For Task 1, a dataset of 3-D CT images is used along with a set of clinically relevant metadata. The dataset includes only HIV-negative patients with no relapses and having one of the two forms of tuberculosis: drug sensitive (DS) or multi-drug resistant (MDR).

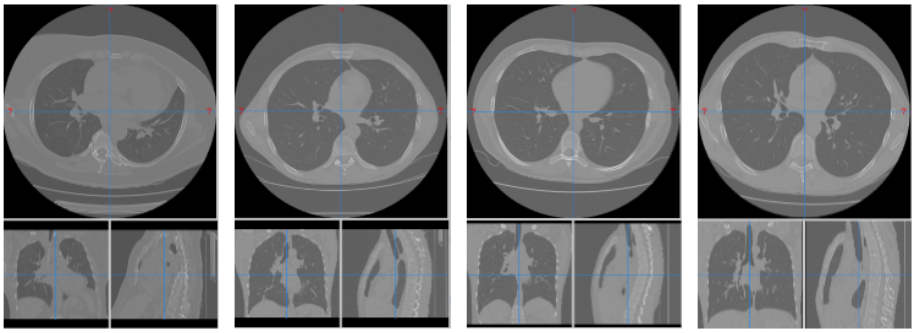

Figure 1: ImageCLEF2017 dataset for lung CT images

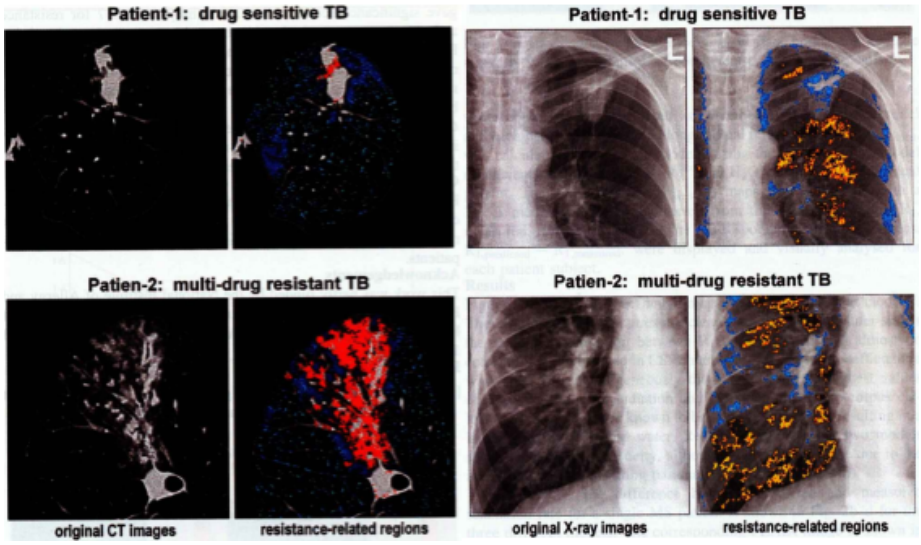

Figure 2: Lung CT&Xray images with drug sensitive and multi-drug resistant TB

Table 1: Dataset for TB MDR

Patients	Train	Test
DS	134	101
MDR	96	113
Total patients	230	214

We survey the development of TB drug resistance, lung region segmentation, and feature extraction and selection in different radiological images(CT and Xray images). Then we designed a new framework to detect TB multi-drug resistance. Comprehensive experiments on the ImageCLEF2017 competition dataset demonstrate that our method significantly outperforms the best results which have been submitted by MedGIFT Group who ranked the first in the ImageCLEF2017 competition Task 1 - Multi-drug resistance detection.

After the literature review, we have summarized TB drug resistance detection methods in a framework(see Fig.3). As we can see from Fig.3, both semi-automatic and automatic prediction are shown in the framework. Researchers first do preprocessing, such as resizing images to uniform sizes, contrast enhancement, etc.. Secondly, they segment lung region in order to extract specific features. After extracting features, it is necessary to use the Principal Component Analysis(PCA) to reduce the number of features. Then they get new features, and combine them with other features(gathered manually). And they train classifiers with the new features. Then they apply these well-trained classifiers to predict lung images in the test dataset, and finally get the results.

2.2 Our method and experiments

Now we present our framework with deep learning method. The Fig.4 shows that our framework consists of three parts, i.e. image preprocessing, feature extraction(VGG16 model) as well as principal component analysis(PCA), and classification(SVM).

Preprocessing CT Images: In order to better use the features of 3-D image, it is necessary to convert the 3-D image into multi-view and multi-level 2-D images. Masks are used to obtain the foreground part (lung region), so images are enhanced to strengthen the texture information.

Feature Extraction: The preprocessed images are input into the pre-trained VGG16 model on ImageNet. And we will give you a detailed explanation of this part later.

Classification: In this part, we use two ways to classify input images: SVM and SOFTMAX. This is because in the subsequent comparison experiments, we find that SVM and SOFTMAX have good performance in some specific ways respectively. For SVM classifier, since feature extraction and classification are divided into two steps, we can add manual features, and combine them with features extracted before. For SOFTMAX classifier, this

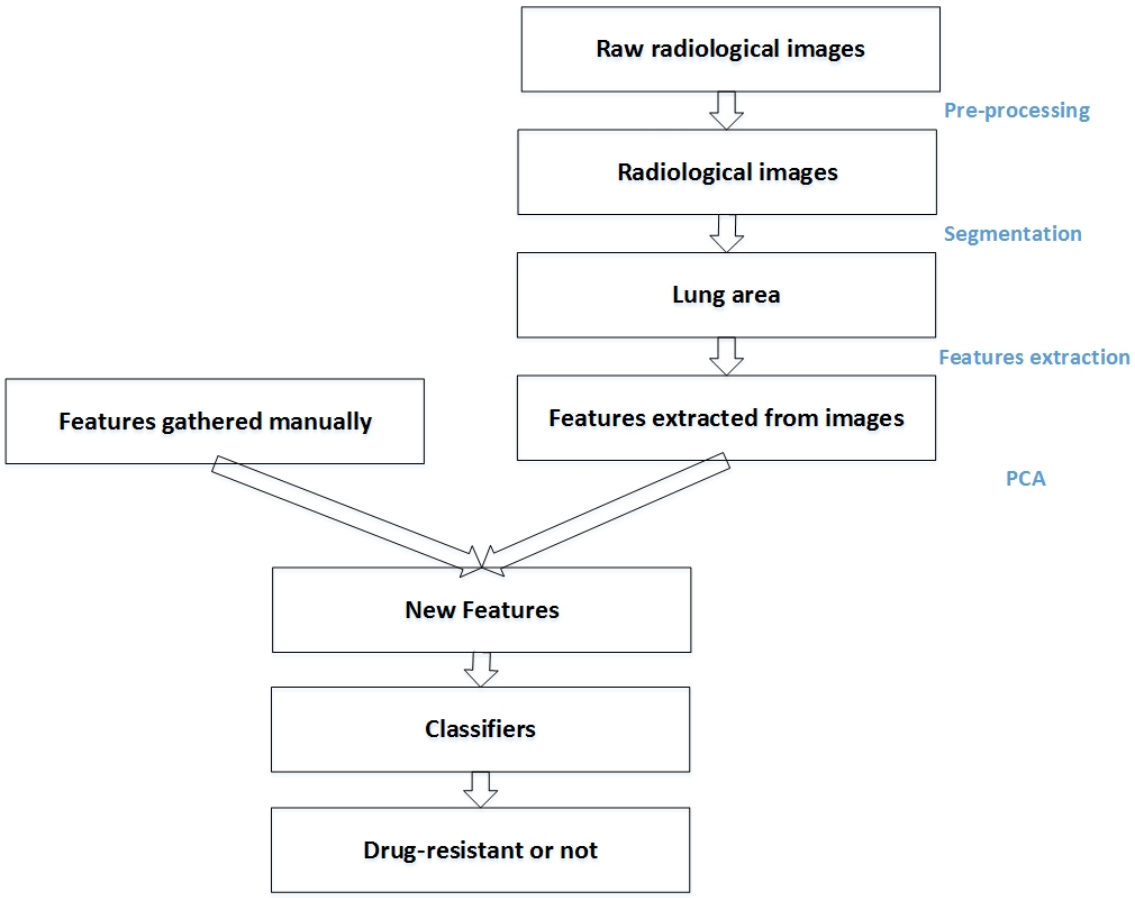

Figure 3: The framework of general TB drug resistance detection

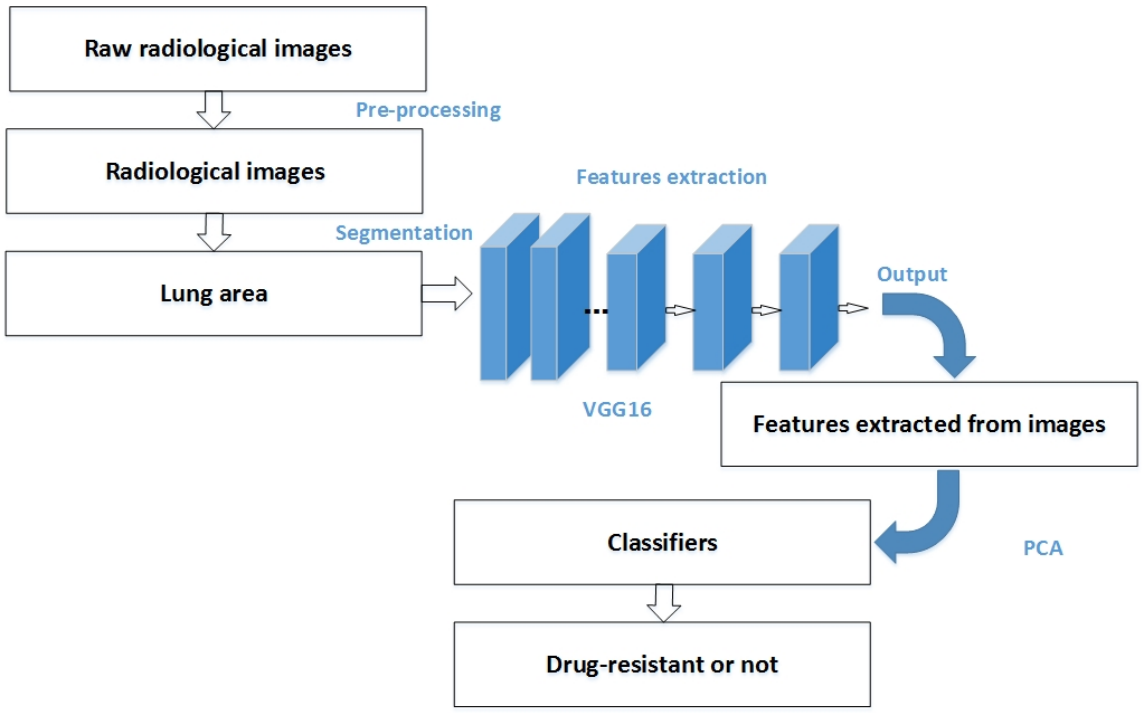

Figure 4: Our framework of TB drug resistance detection using deep learning method

method is more efficient, because it combines feature extraction and classification together. In the following experiments, we will conduct experiments on both classification methods.

Since the number of patients is not large, we choose a relative simple model, i.e. the VGG16 model, to avoid overfitting. In the preprocessing stage, we sample 9 images from each patient's 3-D CT images as the input of the VGG16 model. After extracting features with VGG16, a high dimensional and sparse matrix is obtained for each patient. To get a compact feature representation, we obtain new compact features via PCA.

Our experiments can be divided into three types: classification via SVM with features obtained from original VGG16 model(see Fig.5), classification via SVM with features obtained from fine-tuned VGG16 model(see Fig.6), and classification via softmax(see Fig.7). Next, I will demonstrate each type and its corresponding experiment.

2.2.1 TB MDR detection with original VGG16 and SVM

For the first type, we use the output of the VGG16 model's last three layers as features and train SVM with these features. To make full use of the context information, we combine certain layers to train SVM, and Fig.5 is an example corresponding to the fourth experiment result in Table 2 . In terms of using the original VGG16 model, we conducted experiments in six different ways and the results are presented in Table 2. The evaluation standard is the average precision(RCC), which refers to the definition of the average accuracy of the whole dataset, as in 1.

From Table 2, we can find detection performances with features extracted from two layers are better than those with a single layer, but not the more, the better, because the context information would improve the performance and too much context information would not help.

$$RCC = \frac{1}{n} * \sum_{i=1}^{n} RCC_i \tag{1}$$

Where n is the number of experiments repeated.

Table 2: TB MDR detection results with original VGG16 model and SVM

Methods	RCC(%)
FC1024+PCA+SVM	57.5
FC4096+PCA+SVM	53.0
FC4096(2)+PCA+SVM	51.4
FC1024+FC4096+PCA+SVM	57.6
FC4096+FC4096(2)+PCA+SVM	**58.7**
FC1024+FC4096+FC4096(2)+PCA+SVM	58.2

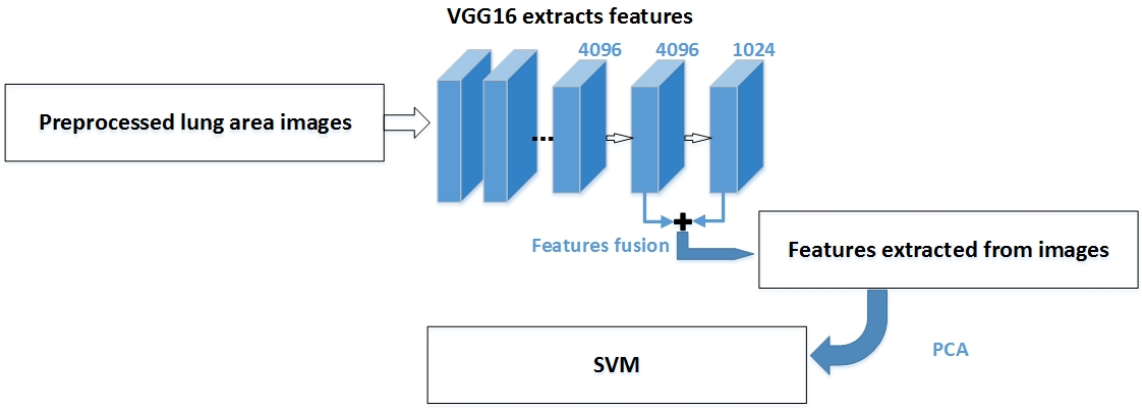

Figure 5: TB MDR detection with the original VGG16 model and SVM

2.2.2 TB MDR detection with fine-tuned VGG16 and SVM

In this section, we will fine-tune the VGG16 model. More specifically, we replace the original fully connected layers with FC4096+FC1024 and use the training dataset of TB images to train the new model with loading the ConvNet parameters pre-trained on ImageNet. It is denoted by fine-tuned1. Similarly, we replace the fully connected layers with FC1024+FC1024 and FC1024, and denoted by fine-tuned2 and fine-tuned3 respectively. Then we conduct extraction fusion of the two new layers' features and get the classification results via SVM. The fine-tuned2 model is shown in Fig.6. The Table 3 demonstrates the experiments results for fine-tuned models.

Table 3: TB MDR detection results with fine-tuned models and SVM

Methods	RCC(%)
fine-tuned1+PCA+SVM	49.1
fine-tuned2+PCA+SVM	**54.7**
fine-tuned3+PCA+SVM	51.4

2.2.3 TB MDR detection via deep learning with softmax

In this part, we use the above fine-tuned models to extract features and then put them into softmax for classification. The Fig.7 shows the fine-tuned2+softmax model and Table 4 illustrates the corresponding experiments results. From the results we find that fine-tuned model with softmax is more time-efficient and the accuracy is higher as well.

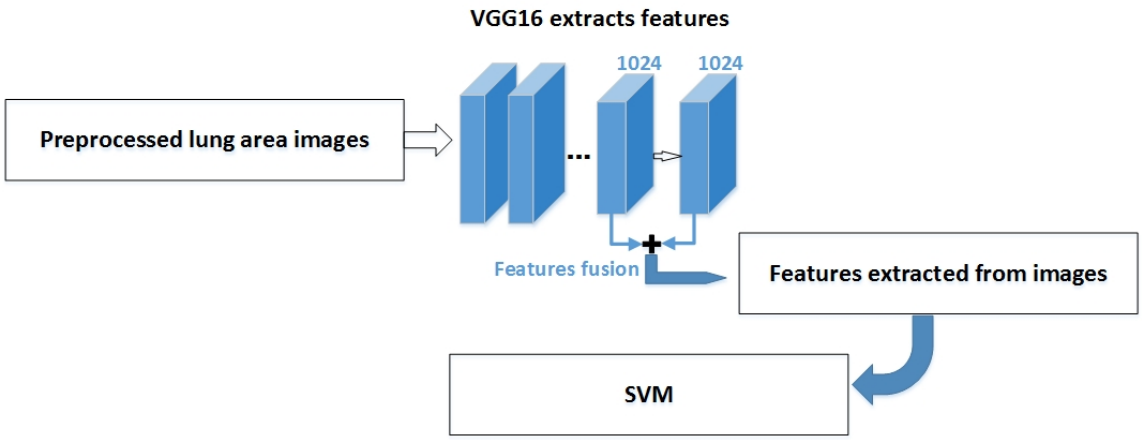

Figure 6: TB MDR detection with the fine-tuned2 model and SVM

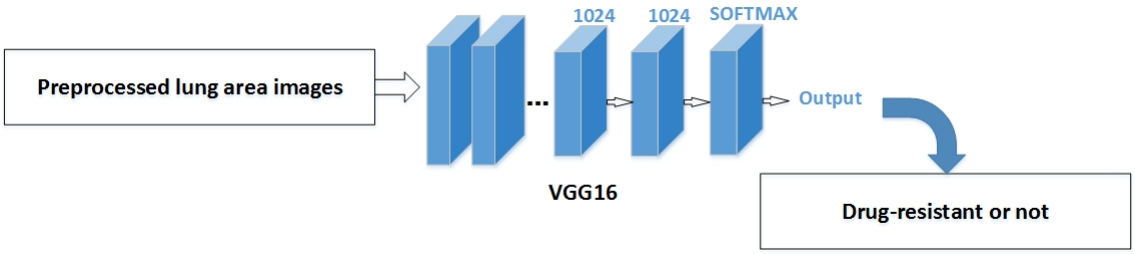

Figure 7: TB MDR detection with the fine-tuned model and softmax

Table 4: TB MDR detection results with fine-tuned models and softmax

Methods	RCC(%)
fine-tuned1+softmax	58.5
fine-tuned2+softmax	60.8
fine-tuned3+softmax	**64.0**

Task 1 - Multi-drug resistance detection

Group Name	Run	Run Type	AUC	ACC	Rank
MedGIFT	MDR_Top1_correct.csv	Automatic	0.5825	0.5164	1
MedGIFT	MDR_submitted_topBest3_correct.csv	Automatic	0.5727	0.4648	2
MedGIFT	MDR_submitted_topBest5_correct.csv	Automatic	0.5624	0.4836	3
SGEast	MDR_LSTM_6_probs.txt	Not applicable	0.5620	0.5493	4
SGEast	MDR_resnet_full.txt	Not applicable	0.5591	0.5493	5
SGEast	MDR_BiLSTM_25_wcrop_probs.txt	Not applicable	0.5501	0.5399	6
UIIP	MDR_supervoxels_run_1.txt	Automatic	0.5415	0.4930	7
SGEast	MDR_LSTM_18_wcrop_probs.txt	Not applicable	0.5404	0.5540	8
SGEast	MDR_LSTM_21wcrop_probs.txt	Not applicable	0.5360	0.5070	9
MedGIFT	MDR_Top2_correct.csv	Automatic	0.5337	0.4883	10
HHU DBS	MDR_basecnndo_212.csv	Automatic	0.5297	0.5681	11
SGEast	MDR_LSTM_25_wcrop_probs.txt	Not applicable	0.5297	0.5211	12
BatmanLab	MDR_submitted_top5.csv	Automatic	0.5241	0.5164	13
HHU DBS	MDR_basecnndo_113.csv	Automatic	0.5237	0.5540	14
MEDGIFT UPB	MDR_TST_RUN_1.txt	Automatic	0.5184	0.5352	15
BatmanLab	MDR_submitted_top4_0.656522.csv	Automatic	0.5130	0.5024	16
MedGIFT	MDR_Top3_correct.csv	Automatic	0.5112	0.4413	17
HHU DBS	MDR_basecnndo_132.csv	Automatic	0.5054	0.5305	18

Figure 8: The ImageCLEF competition participants ranking

2.2.4 Comparison with the participants in the ImageCLEF2017 competition

We list the ImageCLEF2017 participants' submission results in Fig.8. Our best performance is 64% which is higher than the MedGIFT Group(58.25%).

3 HEp-2 Cell Image Classification Method Based on Very Deep Convolutional Networks with Small Datasets

The Indirect Immunofluorescence(IIF) test using human epithelial type 2(HEp-2) cells has been a critical standard to diagnose autoimmune diseases1 because HEp-2 cells have a wide range of antigens. However, the test is still conducted and analyzed manually, which is time-consuming, labor intensive, and low efficient. This section summarizes our work on HEp-2 cell image classification based on Very Deep Convolutional Networks(VGGNET) with small datasets.

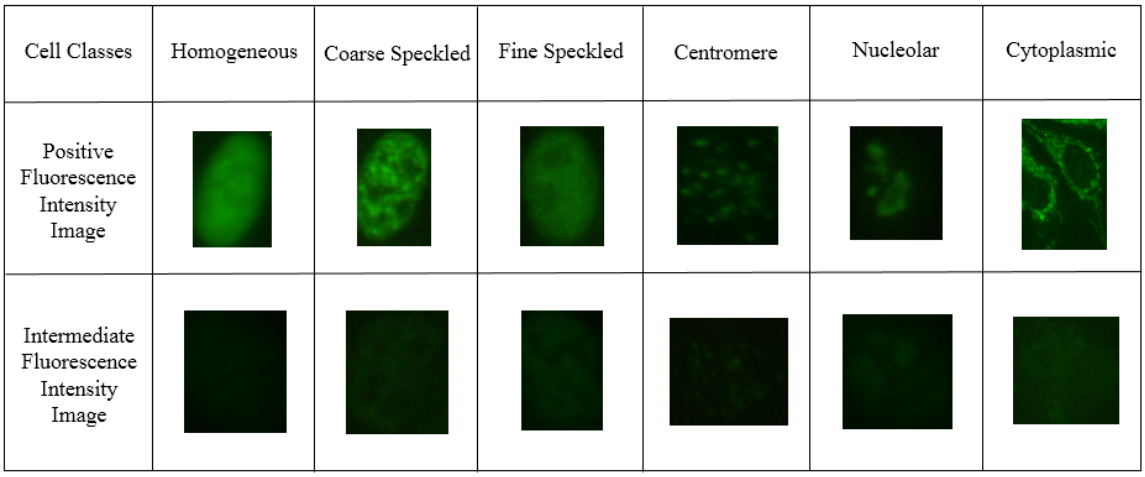

Figure 9: The HEp-2 cell image datasets

Our experiments were conducted on two competition datasets: ICPR2012 and ICPR2014. All images in these two datasets are divided into two parts: positive and intermediate, as shown in Fig.9. More specifically, the training and test sets contain (721, 734) and (10877, 2719) images respectively.

3.1 Our method and experiments

Our framework is shown in Fig.10. It can be divided into three parts: preprocessing, features extraction, and classification.

Preprocessing has a significant influence on the quality of feature extraction, which determines classification accuracy. To reduce the variance and enhance the contrast, we first use mask to keep the foreground of each cell image, and then resize each image to 224*224 pixels to guarantee a uniform scale. Furthermore, after normalizing all images, we obtain the preprocessed cell images which will be used to train VGGNet.

As for the feature extraction stage, since the datasets are small, we make our network as simple as possible, so we remove FC4096+FC4096 to reduce the number of parameters in order to prevent overfitting. As you can see in Fig.10, besides Conv and Pooling layers, only FC1024 and softmax are left. We fine-tune the new model and use the features of the last convolution layer to train the SVM. Then we validate our model on the test datsets.

Table 5 and 6 shows the comparison experiments on ICPR2012 and ICPR2014 datasets, respectively.

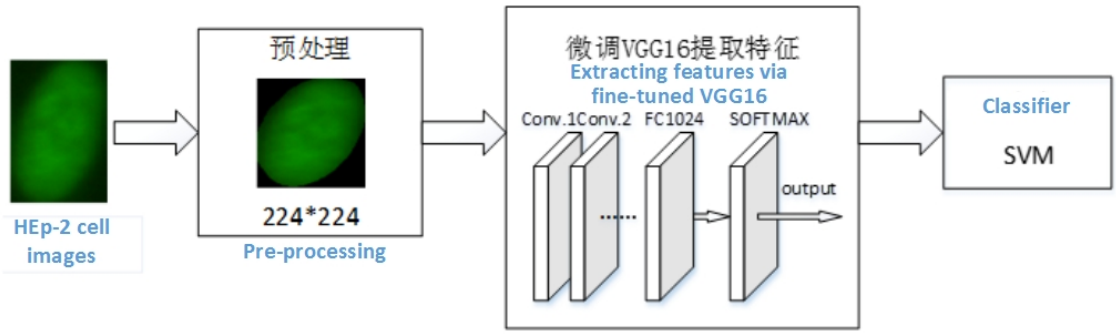

Figure 10: Our HEp-2 cell classification framework

Table 5: The comparison with other HEp-2 cell classification methods on ICPR2012 dataset

Methods	Accuracy(%)
2012 contest	68.7
2012 contest CNN	59.8
Shen(2014)	74.4
Larsen(2014)	71.5
Gao Zhimin(2016)	74.8
Fine-tuned VGGNET(ours)	**92.2**

4 Details about our experiment implementation

Our implementation codes consist of two main parts: feature extraction and classification. More specifically, the feature extraction code was written in python, while we used LIBSVM matlab packages to classify the TB CT images and HEp-2 cell images. As for LIBSVM hyperparameters, we use 5-fold cross validation to choose hyperparameters, e.g. g(gamma) and c(cost). More details about implementation have be released on my Github account(`https://github.com/KaikaiZhao`). In addition, we also conducted some experiments on the Kaggle 2017 lung cancer recognition dataset, but some hyperparameters affect the quality of the model recovered by the training process and its ability to infer correct results when deployed on new inputs. As for lung cancer recognitiontaskin Kaggle 2017 competition, firstly we need to do lung area segmentation, and this time we do not use deep learning techniques to segment lung regions. On the other hand, we use morphological characteristics(e.g. border, prior shape, low level features, etc.) and filters(e.g. shadow filter, multilevel thresholding etc.) to get regions of interest.

Table 6: The comparison with other HEp-2 cell classification methods on ICPR2014 dataset

Methods	Accuracy(%)
Manivannan(2014)	87.10
Paisitkriangkrai(2014)	81.55
Theodorakopoulos(2014)	83.33
Gao Zhimin(2016)	96.76
Fine-tuned VGGNET(ours)	**98.0**

5 Research interests

My main research interests are machine learning, computer vision, bioinformatics, etc.. More specifically, I am very interested in large-scale machine learning, esp. large-scale kernel learning. Also, I have spent a large part of my time on deep learning. I hope that I could apply machine learning algorithms to medical area, e.g. to analyze and understand medical images or clinical records. In my view, the medical applications via AI methods are quite meaningful, which can help reduce the burdens on doctors and improve the accuracy of diagnosis.

Honestly, if I could get a PhD candidate offer, I would combine my research interests with my supervisor's projects to do some solid and meaningful work. I love machine learning and real applications with ML methods are my specific goals.

Most of the figures in this document were drawn manually by myself with Microsoft Visio and that's why they look not very perfect, but they can convey the main ideas. If you have some questions about our work or other questions, feel free to contact me via email. My email addresses are as follows:

zhaojiayi11@nudt.edu.cn
kevin2035@qq.com