White Blood Cell Classification based on Transfer Learning

Name: Kong Yifei Student ID: A0274941X Net ID: E1124687

1. Background

White blood cell classification is a critical area of study in the field of medical diagnostics and healthcare. These vital components of the human immune system, also known as leukocytes, play a pivotal role in defending the body against infections and diseases. The classification of white blood cells holds immense importance and offers numerous advantages for medical and human health, including disease diagnosis and monitoring, infection detection, treatment personalization and early intervention.

The most challenging things in the task are insufficient dataset with proper labels, which may form the bottleneck of WBC classification.

1.1. Types of White Blood Cells

There are 5 types of white blood cells, namely Basophil, Eosinophil, Lymphocyte, Monocyte and Neutrophil.

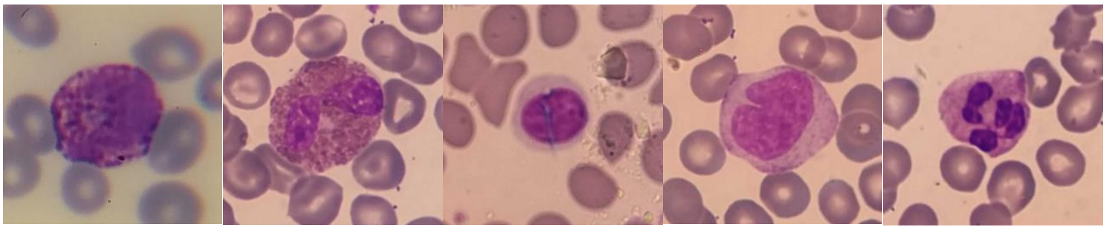

Figure 1 Left to Right: Basophil, Eosinophil, Lymphocyte, Monocyte and Neutrophil

1.2. Datasets Introduction

We use WBC dataset as WBC classification dataset. WBC_100, WBC_50, WBC_10 and WBC_1 are 100%, 50%, 10% and 1% subsets of the WBC dataset, respectively. There are train and validation folders used for training and picking up the best training epoch. Under each folder, there are mask and data folders. Masks are binary images indicating the exact position of white blood cells, where only approximately 10% of data have their masks.

There are CAM16 and pRCC datasets with abundant cell images. The former indicates tumor and normal cells and the later provides massive no-label pictures.

Table 1: Statistics of WBC_100 dataset along with its 50% segregation, 10% segregation, and 1% segregation.

Class	WBC_100			WBC_50		WBC_10		WBC_1	
	Train		Validation	Train		Train		Train	
	data	mask	data	data	mask	data	mask	data	mask
Basophil	176	17	36	88	8	17	1	1	0
Eosinophils	618	61	126	309	30	61	6	6	0
Lymphocyte	2015	201	412	1007	100	201	20	20	2
Monocyte	466	46	95	233	23	46	4	4	0
Neutrophil	5172	517	1059	2586	258	517	51	51	5
Total#	8447	842	1728	4223	419	842	82	82	7

Table 2: Statistics of CAM16 and pRCC datasets.

Class	CAM16				pRCC
	Train		Validation	Test	Train
	data	mask	data	data	data
normal	379	37	54	108	1419
tumor	378	37	54	108	
Total#	757	74	108	216	1419

1.3. Related Works[1]

There are many datasets such as Kaggle and BCCD both with 12444 images. Due to the characteristic of medical images, it's still not enough compared to other tasks. Both machine learning and deep learning methods are used in WBC classification task. For ML methods, K-mean clustering and SVM approach 99.157% and 98.3% classification accuracy. For DL methods, CNN and its variants can achieve accuracy more than 96%, even 100% using Pre trained CNN AlexNet.

2. Methodology

2.1. Model

I use ResNet-18 deep learning model to do WBC classification. The number of output features of the last, full connection layer is 5. The model occupies about 2G graphic memory.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 64 \\ 3\times3, 64 \\ 1\times1, 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1, 128 \\ 3\times3, 128 \\ 1\times1, 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1, 256 \\ 3\times3, 256 \\ 1\times1, 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1, 512 \\ 3\times3, 512 \\ 1\times1, 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Figure 2 Structure of ResNet[2]

2.2. Transfer Learning

A. Training a ResNet-18 model using CAM16 binary classification dataset.

B. Loading the pretrained binary classification model along with its weights.

C. Remove the original output layer of the binary classification model.

D. Incorporate a new output layer with output dimensions matching your 5-class classification task. Typically, this is a fully connected layer with 5 neurons and an appropriate activation function (e.g., softmax).

E. Train the new model with WBC dataset.

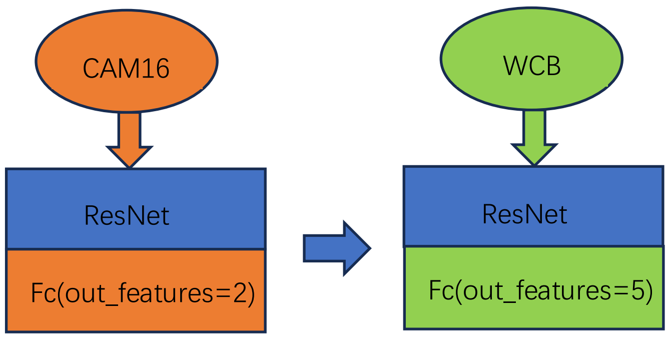

Figure 3 Transfer learning

2.3. Data Augmentation

Data augmentation is a good way to compensate for the insufficiency of data. As there are 10% of data with mask, I obtain the cell area and do random translation, rotation and flipping. For each image, 5 new images will be generated and added to new dataset.

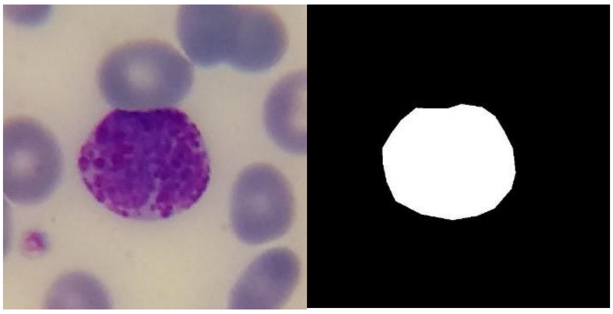

Figure 4 Original Data and Mask

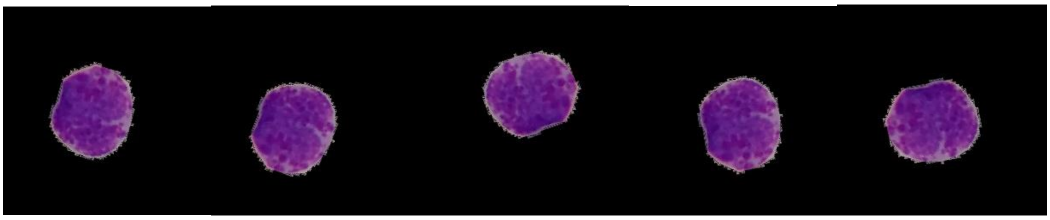

Figure 5 5 New Images

3. Results

I had 3 experiments, ResNet-18, Pretraining and Pretraining & Data Augment.

- ResNet-18: In this experiment, we used a ResNet-18 architecture for our classification task. The model was trained from scratch without any pretraining.
- Pretraining: In this experiment, we employed pretraining. I used a pretrained ResNet-18 model as the starting point and fine-tuned it on our specific classification task. The pretraining task is a binary classification task on CAM16.
- Pretraining & Data Augmentation: In this experiment, I combined pretraining with data augmentation. I started with a pretrained ResNet-18 model and applied various data augmentation techniques to improve the model's performance on our classification task.

Table 1 Results

	WBC_100	WBC_50	WBC_10	WBC_1
Resnet18	97.05%	96.59%	92.53%	84.03%
Pretrain	97.57%	95.78%	93.06%	84.32%
Pretrain & Data Augment	97.63%	96.47%	93.69%	82.93%

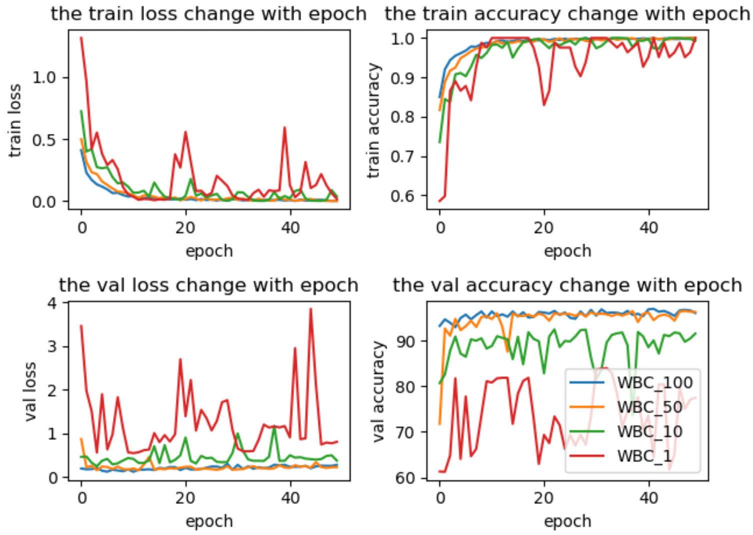

Figure 6 ResNet18

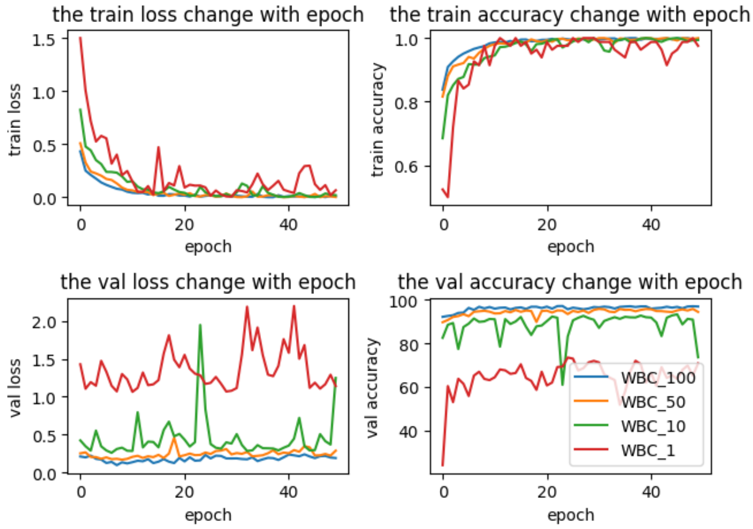

Figure 7 Pretrain

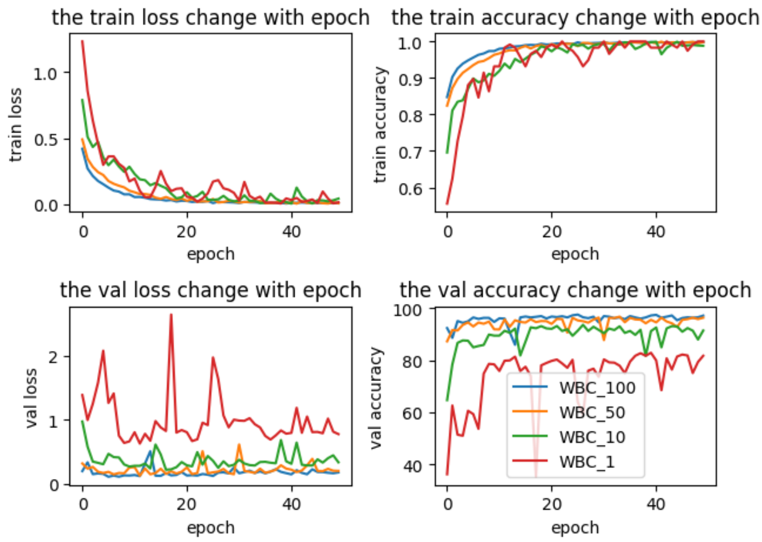

Figure 8 Pretrain & Data Augmentation

4. Conclusion

Pretraining can enhance the classification accuracy by about 0.5% percent in all datasets. Data augmentation can improve classification accuracy when dataset is large enough but reduces the accuracy on WBC_1.

First, pretraining is helpful because the model can extract common features of cells when training on the binary classification task on CAM16, which can be used in WBC classification task.

Second, data augmentation improves the performance of model on WBC_100, WBC_50 and WBC_10 but reduces the performance on WBC_1. This happens because the model can get more information with more data. However, not all data have their corresponding masks. Images with and without background may interface the model, making the model paying more attention to the difference of existence of background. And this disadvantage overcomes its advantage when quantity of data is relatively low.

There are some more ways to improve the model, which require more time and energy to verify. They are potentially useful ways and I will list them below.

Using pRCC Dataset: pRCC is a dataset without label. We can use it to train and use the weight as pretraining weight, like using a ResNet18 to encode and a symmetric one to decode, or clustering.

Getting Mask for All Data: There will be no interface of background if we have masks for all data. But data segmentation is a harder work compared to data classification. I tried both graphics and deep learning method but the performance is not good enough.

5. Reference

[1] Classification of White Blood Cells Using Machine and Deep Learning Models: A Systematic Review. (arXiv:2308.06296)

[2] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.