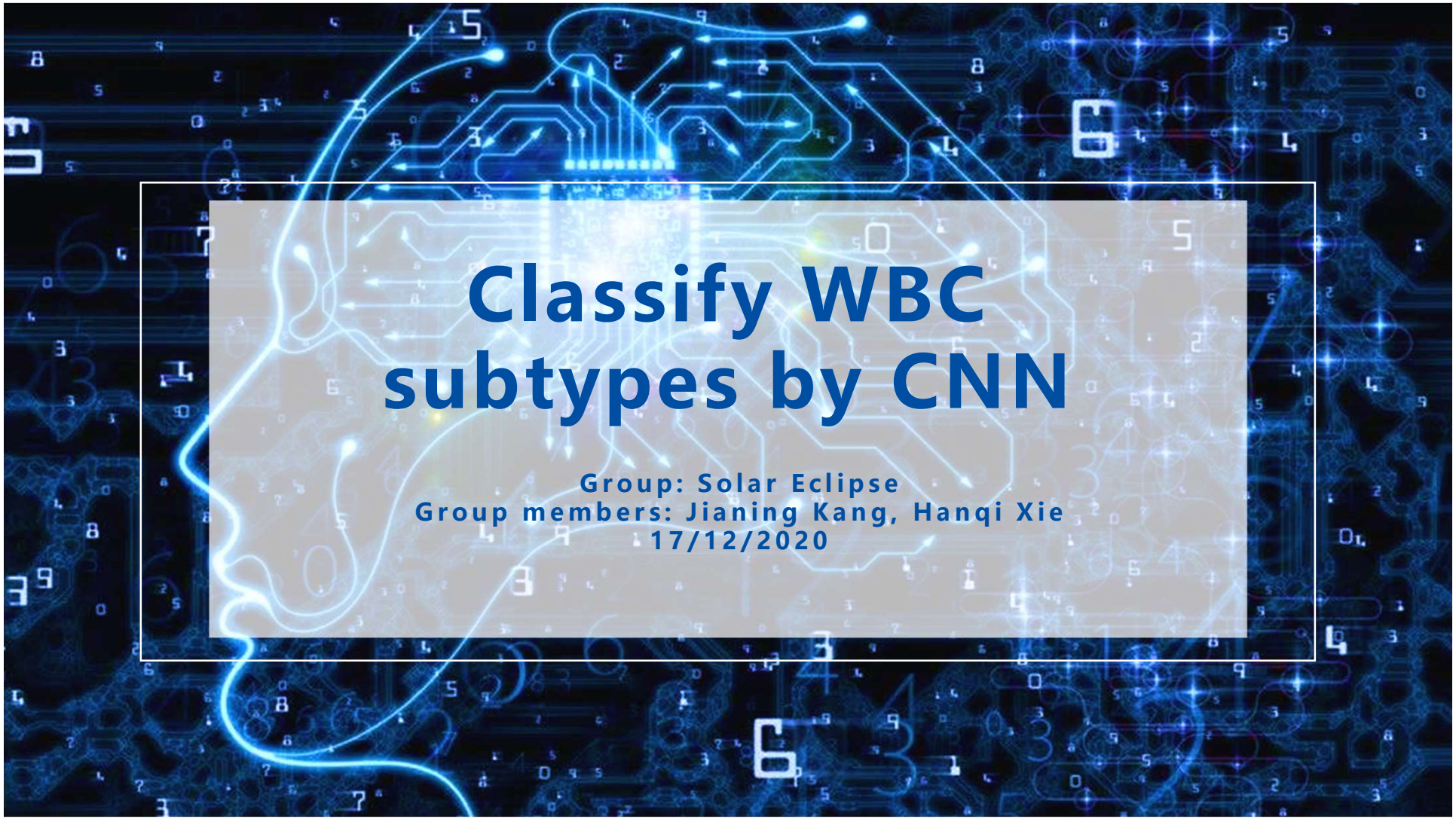


# Classify WBC subtypes by CNN

Group: Solar Eclipse
Group members: Jianing Kang, Hanqi Xie
17/12/2020

# Outline:


01 Introduction

02 Algorithm Design

03 Results

04 Discussion

# Outline:


01 Introduction

02 Algorithm Design

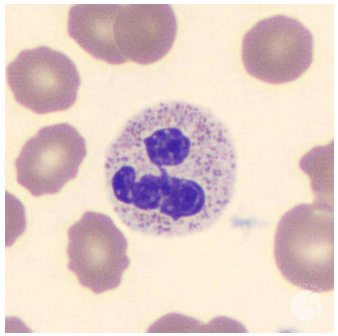
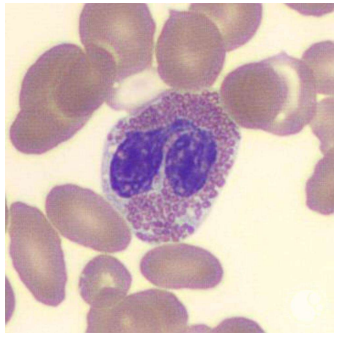
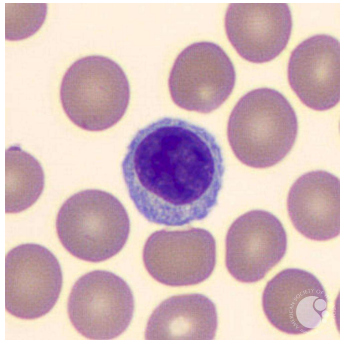
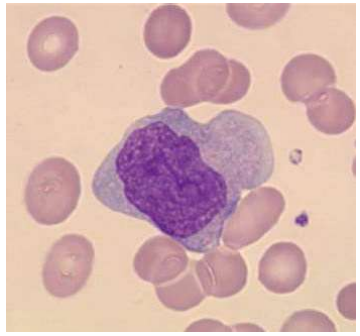
03 Results

04 Discussion

## Background and significance

- Classifying WBCs is needed during clinical diagnose of blood-based disease.
- Classification cannot be efficiently done by automated hematology analyzer.
- If the process is automated, it can largely reduce the burden of medical staff as well as improve efficiency.

## Aim of the project

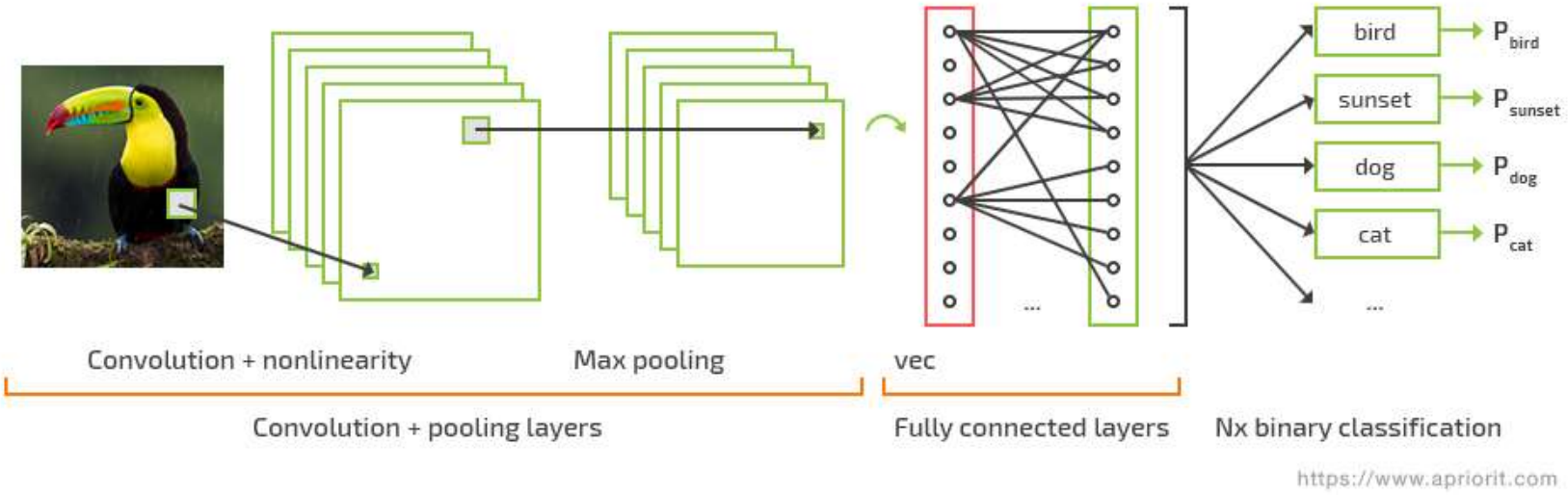
Polynuclear

Neutrophil　　Eosinophil

Mononuclear

Lymphocyte　　Monocyte

# Convolutional Neuron Network(CNN)




- The core building block in CNN is convolutional layer
- Convolutional layer uses kernel to extract features from images.

# Convolutional Neuron Network(CNN)

**Advantage:**

It can extract features from a part of the object and perform object recognition, so it has a high performance in classifying partially visible cells. CNN is good at classifying images.

**Disadvantage:**

Training CNN is computationally demanding.

# Outline:

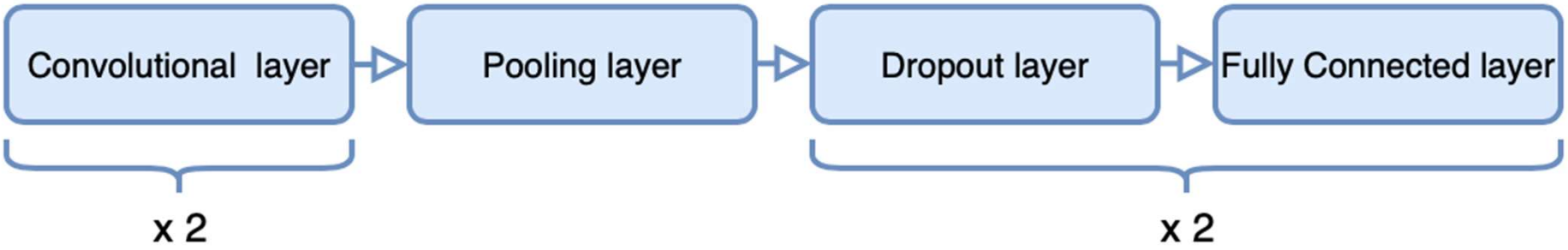

01 Introduction

02 Algorithm Design

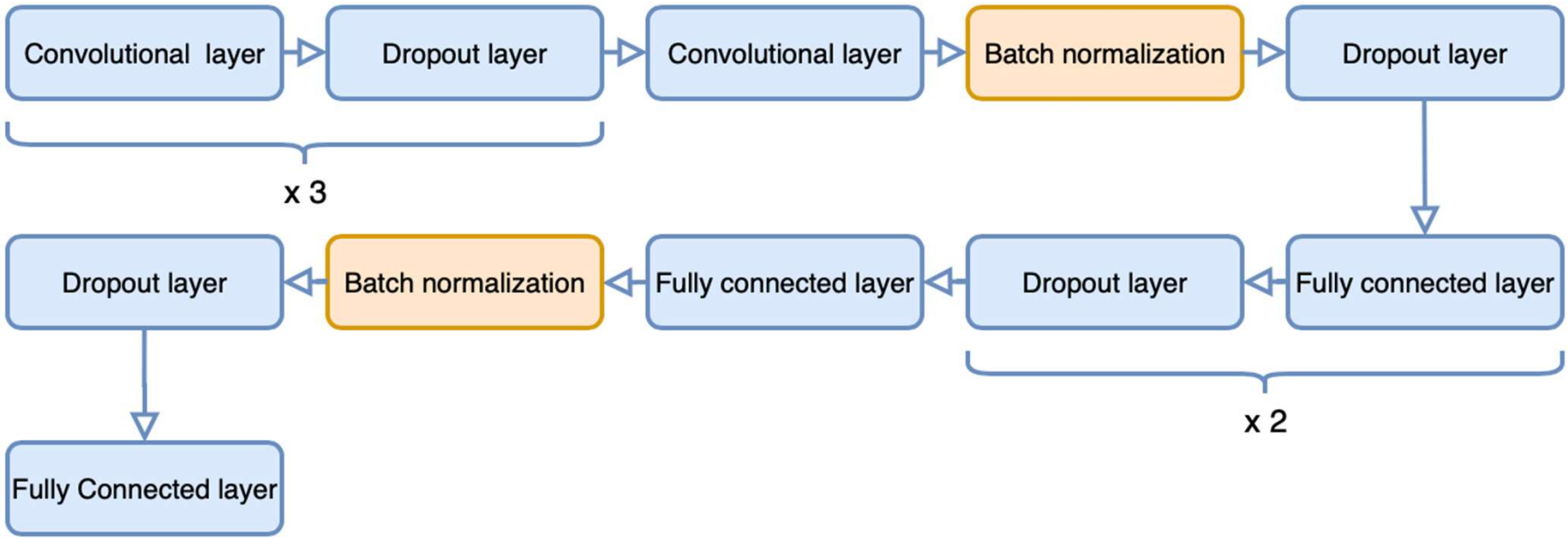
03 Results

04 Discussion

# 2-class model



# 4-class model



# Outline:

01 Introduction

02 Code implementation

03 Results

04 Discussion

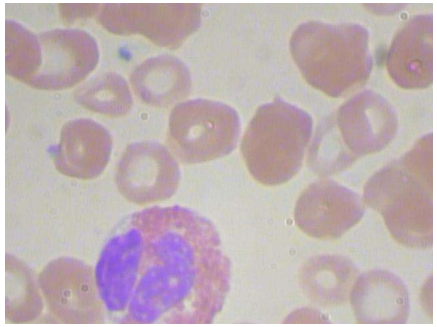
## Results

INTERNATIONAL CAMPUS
ZHEJIANG UNIVERSITY
浙江大学国际联合学院

# BCCD dataset:

410 images in total:

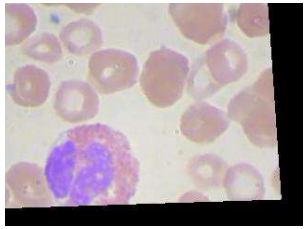
- 88 eosinophils
- 33 lymphocytes
- 21 monocytes
- 207 neutrophils

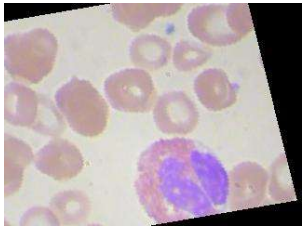


Augmentation

~125,000 images in total:

- ~3,000 images for each cell type in the training set
- ~600 images for each cell type in the testing set.





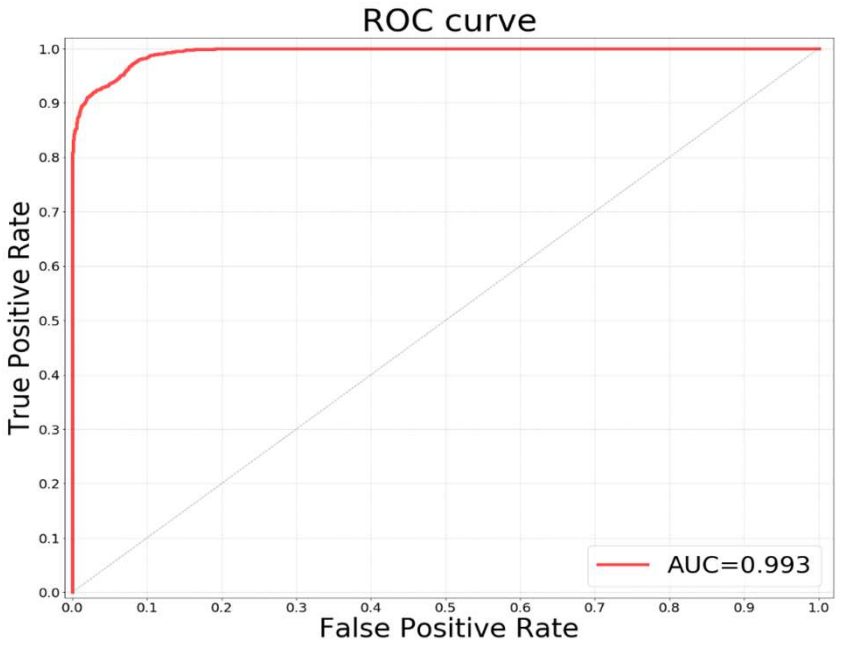
# Results — 2-class model

INTERNATIONAL CAMPUS
ZHEJIANG UNIVERSITY
浙江大学国际联合学院

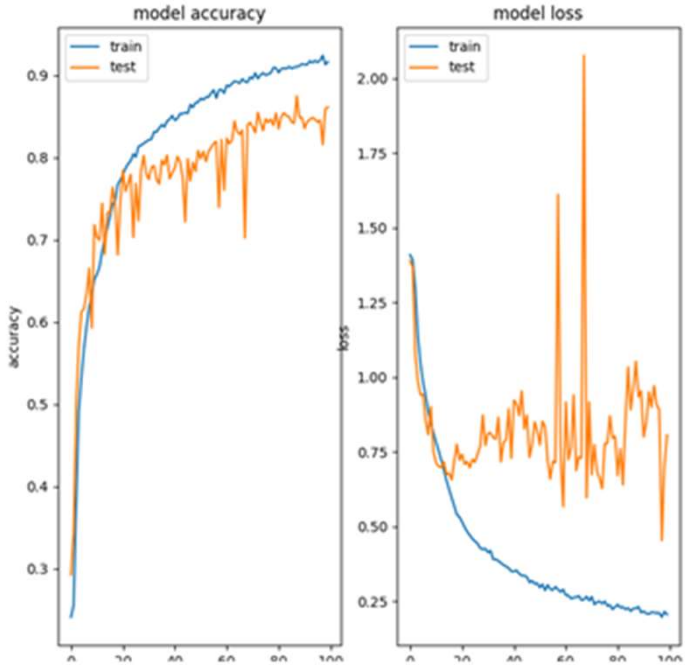
Overall accuracy: 0.9324

| | Precision |
|---|---|
| Mononuclear | 1.00 |
| Polynuclear | 0.89 |

Overall AUC: 0.993
(area under curve)



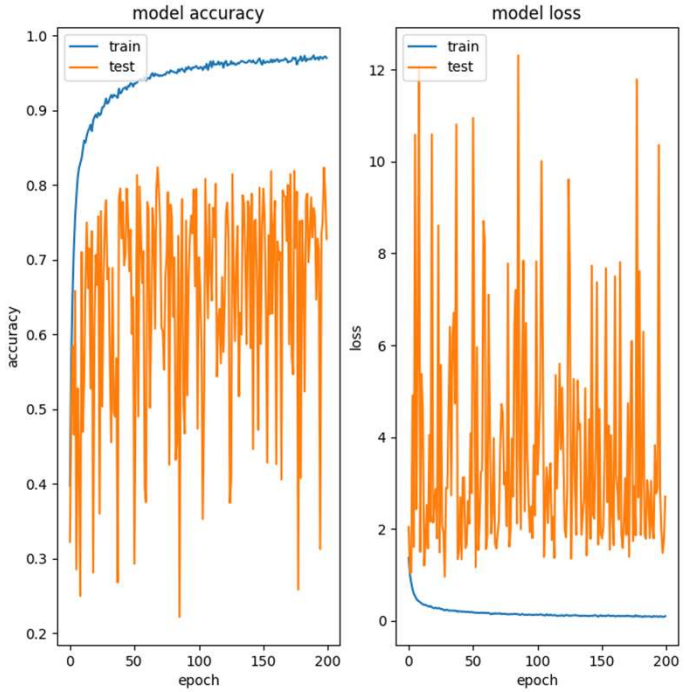
# Increase accuracy




Overall testing accuracy: 0.8343

Add two convolutional layers

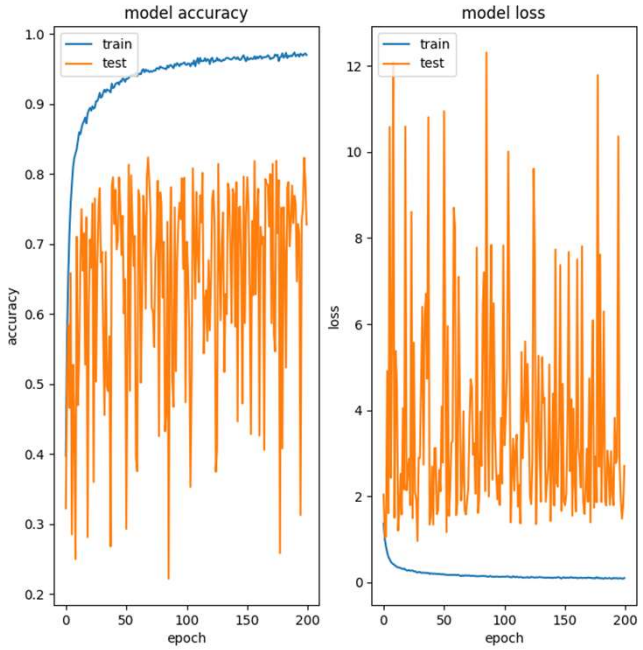
Add two fully connected layers



# Results — 4-class model

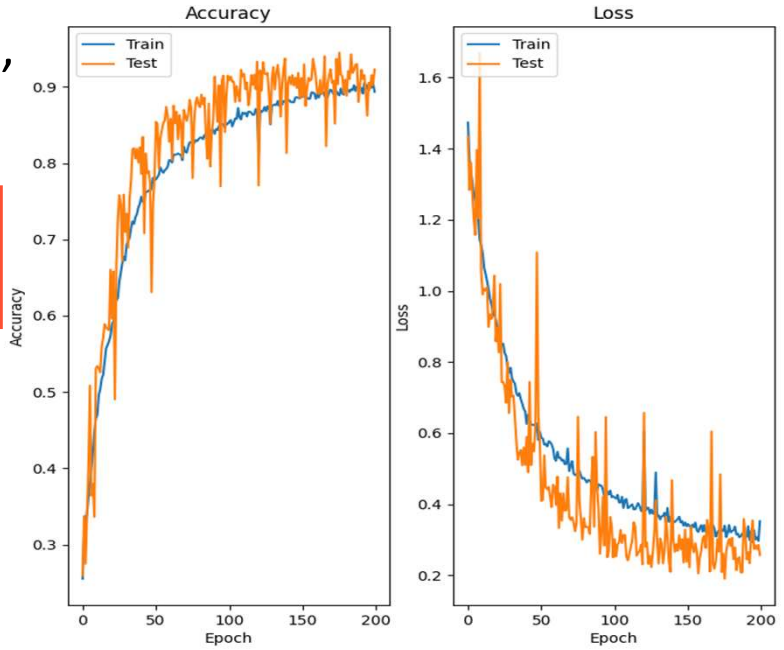
INTERNATIONAL CAMPUS
ZHEJIANG UNIVERSITY
浙江大学国际联合学院

## Reduce overfitting




Dropout, stride, kernel size

Batch normalization




(Ioffe and Szegedy, 2015)

## Results — 4-class model

INTERNATIONAL CAMPUS ZHEJIANG UNIVERSITY 浙江大学国际联合学院

### BCCD dataset + LIT dataset:

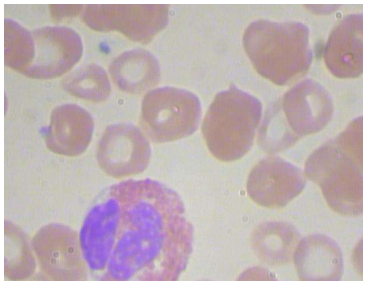
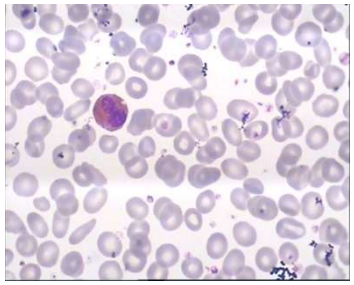
410 images + 189 images in total

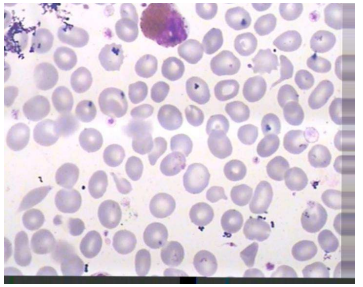
- 88 eosinophils + 39 eosinophils
- 33 lymphocytes + 52 lymphocytes
- 21 monocytes + 48 monocytes
- 207 neutrophils + 50 neutrophils

Augmentation

~ 12,400 images in total:

- ~2700 images for each cell type in the training folder
- ~600 images for each cell type in the test folder.



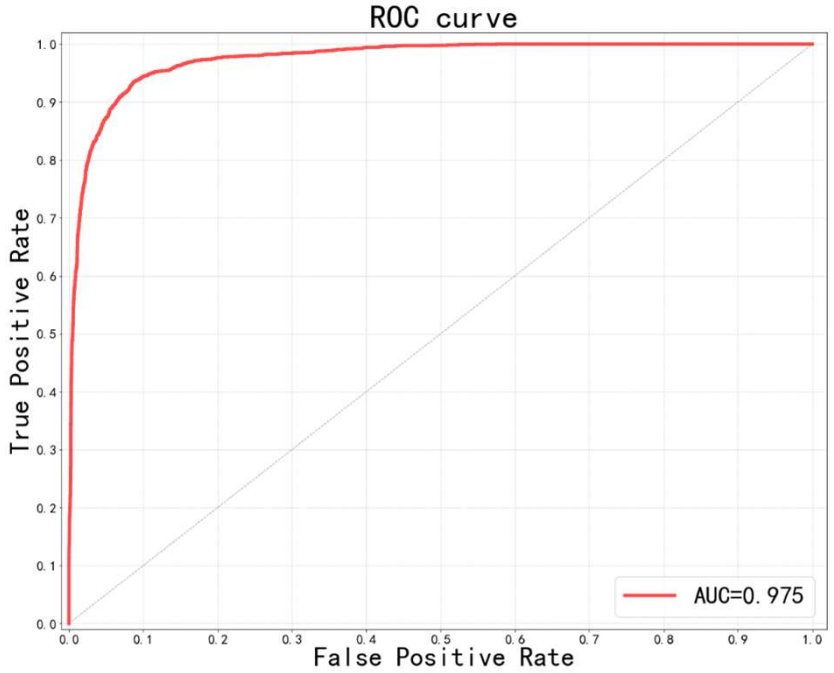
# Results — 4-class model

INTERNATIONAL CAMPUS
ZHEJIANG UNIVERSITY
浙江大学国际联合学院

Overall accuracy: 0.9226

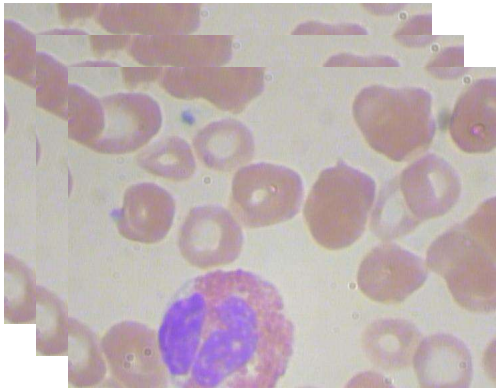
| | Precision |
|---|---|
| Neutrophil | 0.85 |
| Eosinophil | 0.99 |
| Monocyte | 0.89 |
| Lymphocyte | 0.99 |

Overall AUC: 0.975
(area under curve)



# Command line execution:

python3 classify_WBC.py [-h] [-m 1/2] [-f filename] [-d directory] [-o output_file.csv]



Input

| | Filenames | Predictions | NEUTRC | EOSINO | MONOC | LYMPHC |
|---|---|---|---|---|---|---|
| 0 | _12_2599.jpeg | EOSINOPHIL | 0.367759 | 0.631308 | 0.00041 | 0.00051 |
| 1 | _2_1226.jpeg | EOSINOPHIL | 0.397907 | 0.593848 | 0.00461 | 0.00362 |
| 2 | _1_5031.jpeg | NEUTROPHIL | 0.951597 | 0.048372 | 2.91676 | 8.49913 |
| 3 | _3_625.jpeg | EOSINOPHIL | 0.251063 | 0.747483 | 0.00043 | 0.00101 |
| 4 | _5_1744.jpeg | EOSINOPHIL | 0.027477 | 0.972497 | 3.65024 | 2.41422 |
| 5 | _11_9310.jpeg | EOSINOPHIL | 0.214300 | 0.777875 | 0.00239 | 0.00542 |

Output

# Outline:

01 Introduction

02 Code implementation

03 Results

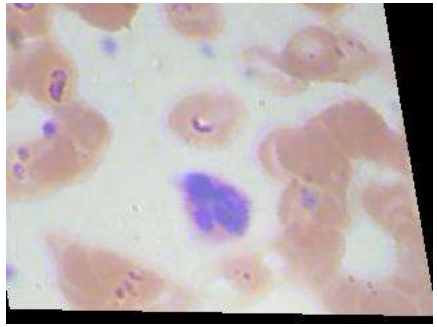
04 Discussion

# Discussion — limitation:

INTERNATIONAL CAMPUS
ZHEJIANG UNIVERSITY
浙江大学国际联合学院

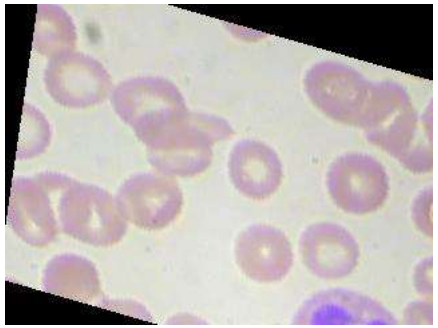
Some cells have low resolution

Neutrophil

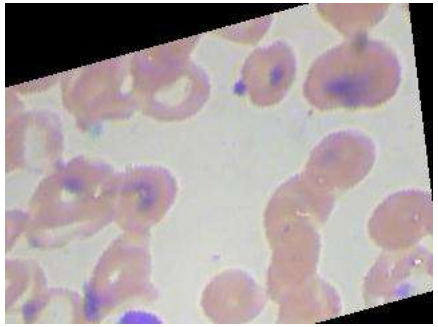


Eosinophil

Problem with augmentation



Neutrophil



Lymphocyte

# Picture segmentation:

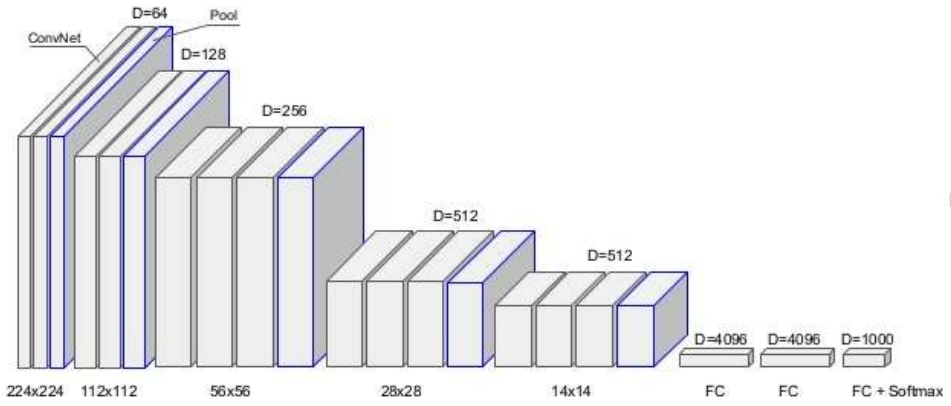
- Allows input of images containing multiple WBCs, which is more suitable for clinical use
- Take images of blood smear samples from patients as input and calculate the percentage of each WBC subtype.
- Require a new training and testing dataset

# Discussion — possible improvements

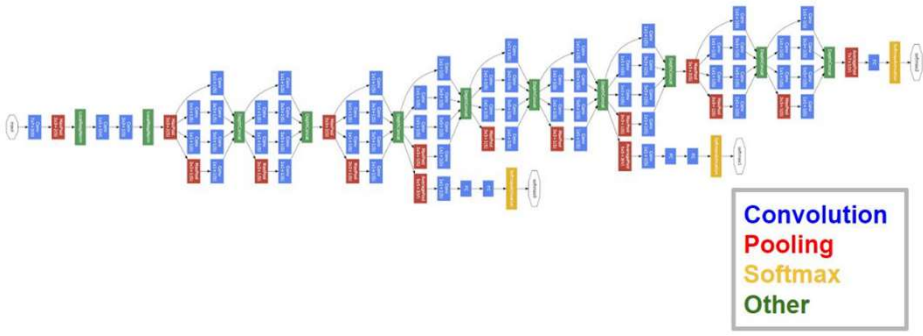


## Classical CNN designs:

### VGGNet




### GoogLeNet




Thank you for listening~
Any questions?

# Reference

https://cn.bing.com/images/search?view=detailV2&ccid=LDi4FcGp&id=0B073BBC14AC47115510C7138AC3CEE0FD7A13FE&thid=OIP.LDi4FcGpHSI9NYZVlpliZAHaCk&mediaurl=https%3a%2f%2fwww.apriorit.com%2fimages%2farticles%2faction-detection-using-dnn%2ffigure-1.jpg&exph=330&expw=950&q=convolutional+neuron+network&simid=608054270189178525&ck=3EE257942FE72F5EA90051A3B86FC956&selectedIndex=144&FORM=IRPRST&ajaxhist=0

https://cn.bing.com/images/search?view=detailV2&ccid=CeFN%2bzdp&id=141FFE42CB3A49C42672390CBCBFD0A63E244024&thid=OIP.CeFN-zdpx81x000V5Ul76AHaEK&mediaurl=http%3a%2f%2fimage.slidesharecdn.com%2fdeeplearningclass2-louismonier-160501185826%2f95%2fdeep-learning-class-2-deep-learning-for-images-i-see-what-you-mean-12-638.jpg%3fcb%3d1462253453&exph=359&expw=638&q=VGGnet&simid=608009714292754417&ck=D2598136096CF12FD94E86473E0370CD&selectedIndex=3&FORM=IRPRST&ajaxhist=0

BCCD, Blood Cell Images. 2019. Available online: URL https://www.kaggle.com/paultimothymooney/blood-cells/home. (accessed on 28.10.2020).

LISC: Leukocyte Images for Segmentation and Classification. 2019. Available online: URL http://users.cecs.anu.edu.au/~hrezatofighi/Data/Leukocyte%20Data.
htm. (accessed on 25.09.2019).

https://blog.csdn.net/shuzfan/article/details/50738394