Faculty of Computers and Artificial Intelligence

Computer Science Department

2021/2022

CS 396 Selected Topics in CS-2

Research Project

Report Submitted for Fulfillment of the Requirements for Selected Topics in CS-2 course for Spring 2022

Team ID No. 21

	ID	Name	Grade
1.	201900794	مروان محمد محمود محمد	
2.	201900501	عمر أسامة محمد صابر	
3.	201900158	أكرم سمير سيد حسين	
4.	201900476	عز الدين أشرف محمد ابراهيم	
5.	201900426	عبد الرحمن عمرو عبد السلام	
6.	201900453	عبدالله أسامة حسن	
7.	201900619	محمد أحمد السيد أحمد	

Paper Details:

Paper title: Blood Cell Types Classification Using CNN.

Authors Name: Ishpreet Singh, Narinder Pal Singh, Harnoor Singh, Saharsh Bawankar and Alioune Ngom.

Publisher Name : Springer, Cham

Publish Year: 2020

Dataset used in the paper : Identify white Blood Cell Subtypes

Algorithm : Multiclassification using CNN

Dataset link :
https://www.kaggle.com/code/khurramabbas/identify-white-blood-cell-subtypes-from-images/log

Results:

10 epochs:

Training loss :0.85 (high loss)

Validation loss :0.82 (high loss)

Testing loss:0.80 (high loss)

Training accuracy :64.5%

Validation accuracy :66.9%

Testing accuracy :66.98%

100 epochs

Training loss :0.165

Validation loss :0.17

Testing loss:1.02

Training accuracy :93.48%

Validation accuracy :93.5%

Testing accuracy :80.9%

200 epochs

Training loss :0.06

Validation loss :0.11

Testing loss:0.94

Training accuracy :97.7%

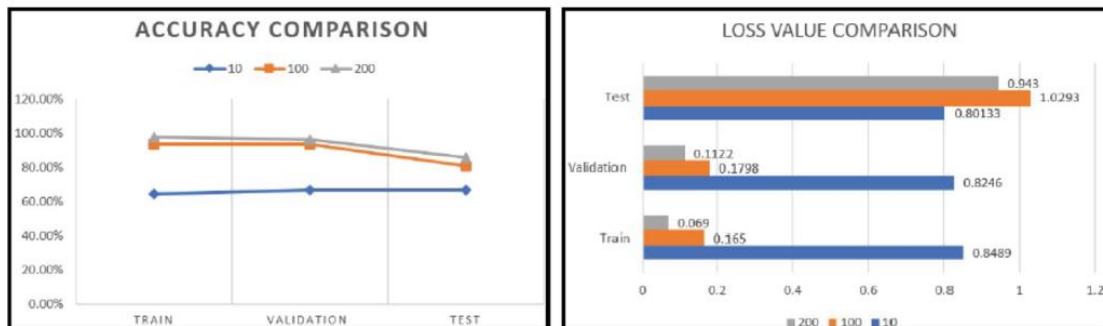Validation accuracy :96.1%

Testing accuracy :86%

Fig. 11. Comparison of accuracy and loss function value for 10, 100 and 200 epochs

Project Description:

1 – Selected dataset :

Name: [Raabin-WBC Data](http://dl.raabindata.com/WBC/Cropped_double_labeled/)
Link : http://dl.raabindata.com/WBC/Cropped_double_labeled/

Total number of samples : 14514

Train : 10,175

Test: 4,339

No of classes : 5

Classes : Basophil **,** Eosinophil , Lymphocyte , Monocyte and Neutrophil

 Image dimension : 575 * 575

2 – Implementation Details :

Training Ratio : 60%   no of images : 9159

Testing Ratio : 30%   no of images: 4339

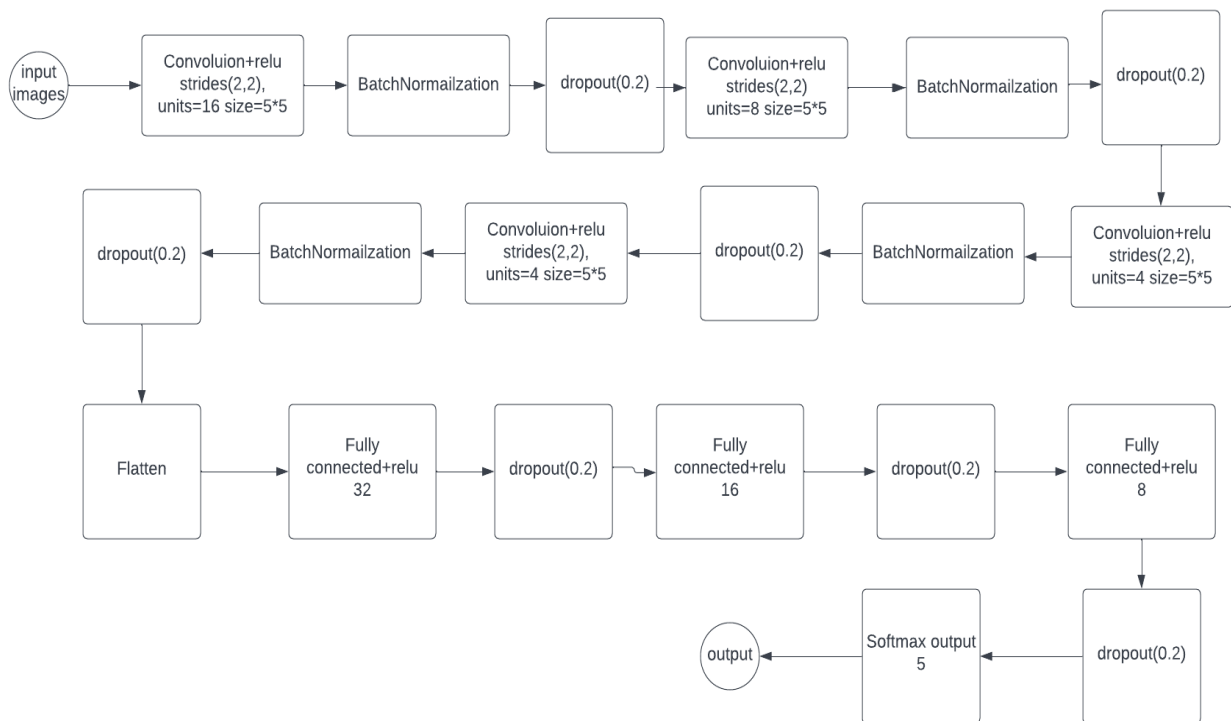Validation Ratio: 10%   no of images: 1016

Model block diagram:

Input shape : 120 ,160 ,3

```
Total params: 16,741
Trainable params: 16,677
Non-trainable params: 64
```


Hyper Parameters:

First Layer:

(Convolution(16,5) , activation function -> relu, strides(2,2),inputs size shape(120,160,3) ,padding ((Input width) / (stride width) + 1.,(Input height) / (stride height) + 1.) , Dropout = 0.2)

Second Layer:

(Convolution(8,5) , activation function -> relu, strides(2,2),padding ((Input width) / (stride width) + 1.,(Input height) / (stride height) + 1.) , Dropout = 0.2)

Third Layer:

(Convolution(4,5) , activation function -> relu, strides(2,2),padding ((Input width) / (stride width) + 1.,(Input height) / (stride height) + 1.) , Dropout = 0.2)

Fourth Layer:

(Convolution(4,5) , activation function -> relu, strides(2,2),padding ((Input width) / (stride width) + 1.,(Input height) / (stride height) + 1.) , Dropout = 0.2)

First Dense Layer:

(Size of Dense Layer = 32,activation = relu, Dropout = 0.2)

Second Dense Layer:

(Size of Dense Layer = 16,activation = relu, Dropout = 0.2)

Third Dense Layer:

(Size of Dense Layer = 8,activation = relu, Dropout = 0.2)

Output Dense Layer:

(Size of Dense Layer = 5,activation = SoftMax)

Optimizer : RMSprop

Loss :categorical_crossentropy

Metrics : accuracy

Results :

We trained our model for 100 epochs but it started to overfit after the epoch number 20 so we used early stopping function to stop after 20 more epochs and save the best model

So we stopped at 40 epoch and the best model was at 20 epoch so we used it.

```python
from keras import callbacks
earlystopping = callbacks.EarlyStopping(monitor ="val_loss",
                                        mode ="min", patience = 20,
                                        restore_best_weights = True)

history = model.fit(train_images,
    validation_data=val_images,
    epochs=100,
                    callbacks =[earlystopping])
```
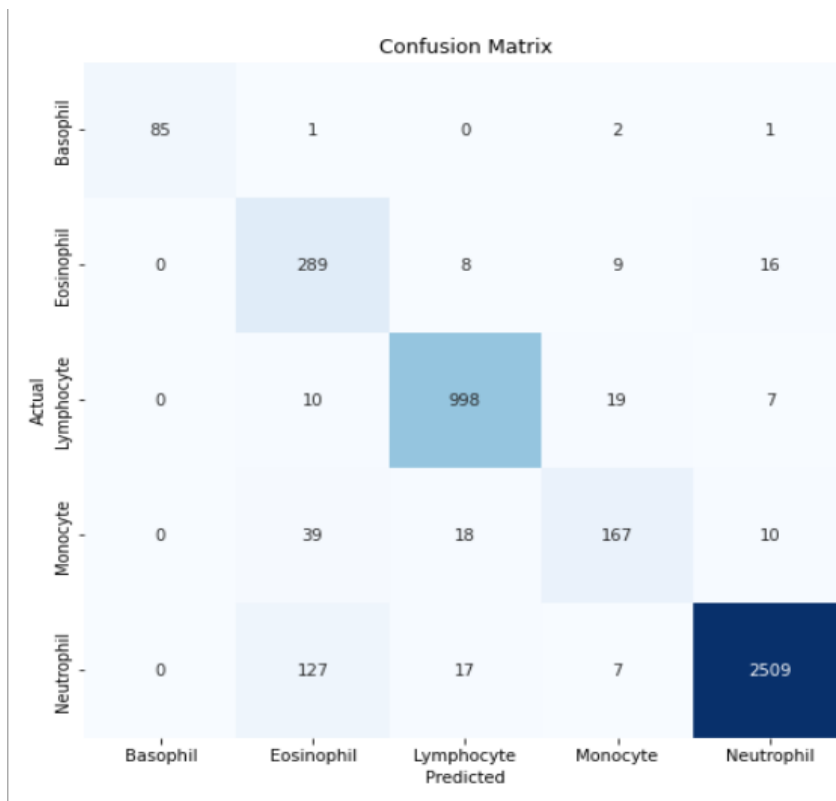
Train loss : 0.23

Validation loss: 0.36

Train accuracy: 92.3%

Testing accuracy: 93.293%

Validation accuracy: 92.6%

Confusion Matrix :

Classification Report :

```
Classification Report:
----------------------
              precision    recall  f1-score   support

    Basophil       1.00      0.96      0.98        89
  Eosinophil       0.62      0.90      0.73       322
  Lymphocyte       0.96      0.97      0.96      1034
    Monocyte       0.82      0.71      0.76       234
  Neutrophil       0.99      0.94      0.96      2660

    accuracy                          0.93      4339
   macro avg       0.88      0.89      0.88      4339
weighted avg       0.94      0.93      0.94      4339
```
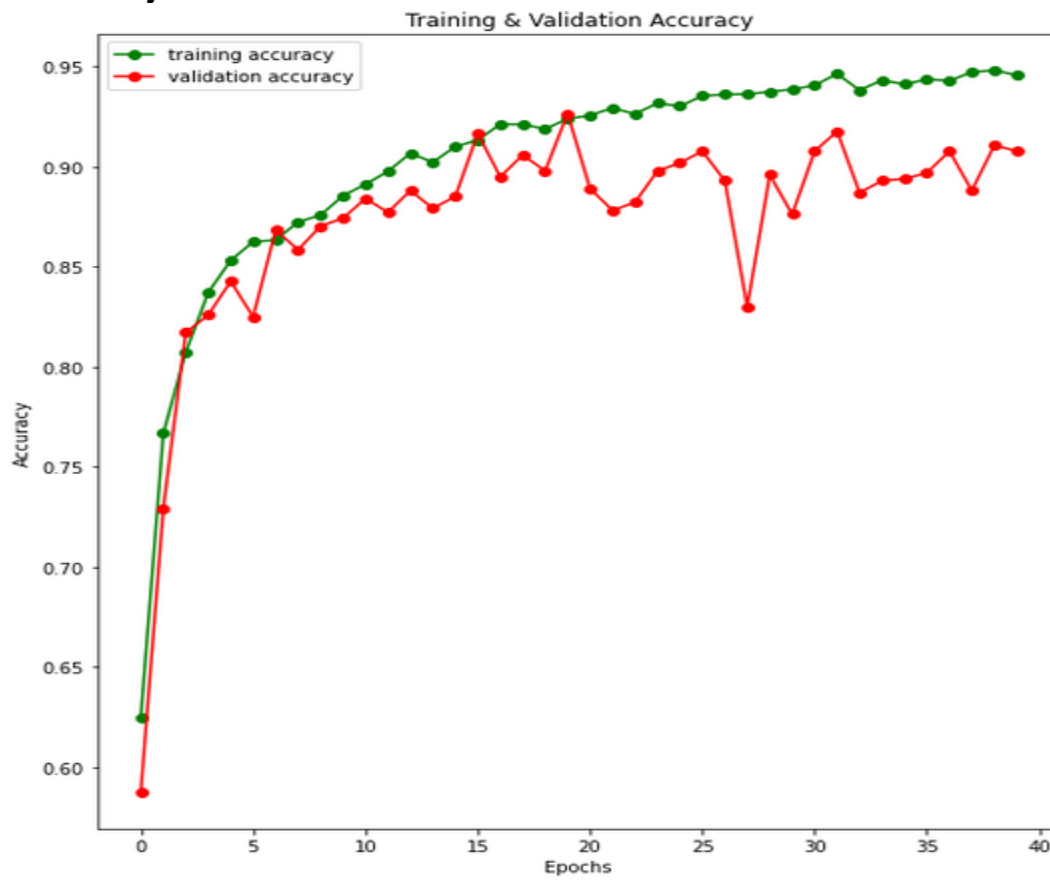
Accuracy curve:

Loss curve :