Sickle Cell Classification Using Deep Learning: Final Project Report

Jordan Mozebo

May 1, 2025

1 Introduction

The sickle cell is an hereditary disease caused by a mutation at the genetic level. It affects hemoglobin in red cells. The gene susceptible to mutation is the HBB gene, responsible for encoding hemoglobin. The mutation leads to hemoglobin S, which is different from regular hemoglobin. Its difference lies in the fact that hemoglobin S tends to be long in shape and is rigid rod-shaped under low oxygen [1].

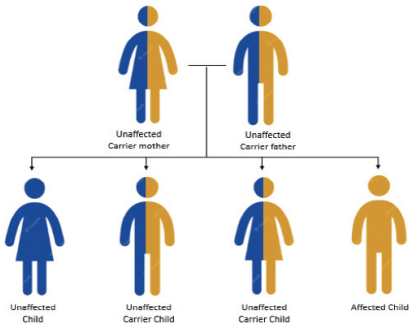

Figure 1: Inheritance Of Sickle Cell Disease
source: [1]

2 Background

So far, manual identification and classification of sickle cell are more relied on in the medical field. However, this type of approach has been relieved to be a laborious process prone to error since the approach relies on human observation. Therefore, an automated approach must be applied in the identification and classification of sickle cells. Due to technological advancement in the field of artificial intelligence, machine learning, and deep learning, there are techniques that are promising in medical images analysis. In addition, deep learning can be used to extract features automatically from images [2]. Many works have been done using CNN-based approach for automating sickle cell classification such as GoogleNet, ResNet18, and ResNet50. However, the authors did not extensively experiment with lightweight model such as MobileNetv3 small , which is known for its efficiency and fast inference [1] [3].

3 Methodology

The approach used involves data gathering, data preparation, network selection, setting training options, and training the network with different hyperparameters.

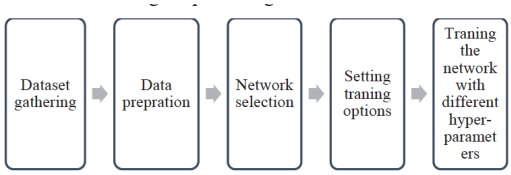

Figure 2: Workflow
source: [1]

3.1 Data pre-processing

The dataset used in this experiment came from the University college of London sickle cell dataset, and contains 1971 images. The images were converted from .tiff to jpeg format; in addition to that, some data augmentation techniques have been applied as well. Techniques such as horizontally flipping and rotating the images to add variability to the dataset. The dataset was split into 75% for training, and 25% for validation to ensure a balanced data distribution.

3.2 Model architecture

The model used in this experiment is a custom MobileNetv3 small, its architecture is mainly composed of Squeeze-Excitation blocks to summarize feature maps into single value, highlight important features, and dime the unimportant ones. The model is also made of depthwise separable convolution for feature extraction with fewer paramters. Depthwise convolution applies a separate spatial filter to each input channel independently, capturing patterns like edges without mixing channel informatiion. In addition, swish activation was used to improve training stability in deep networks by allowing gradient flow. It also preserves useful information from small negatives inputs. The final layer turns the deep features the model has learned into decision between sickle cell and normal cell [3].

3.3 Training

During the training process, Adam optimizer was used for adaptive learning rates, efficient convergence, and holding sparse gradients in deep networks like MobileNetv3 small, with a weight decay of 1e−4

4 Results

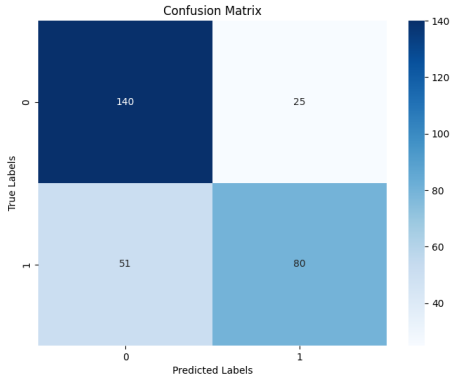

Figure 3: Confusion matrix

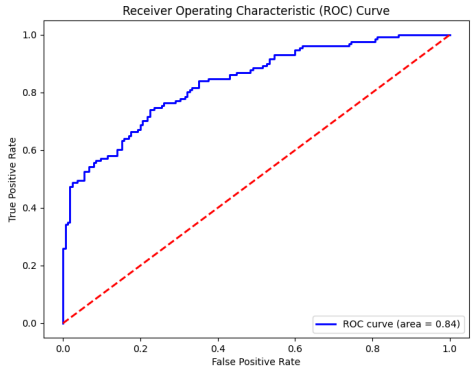

Figure 4: Workflow

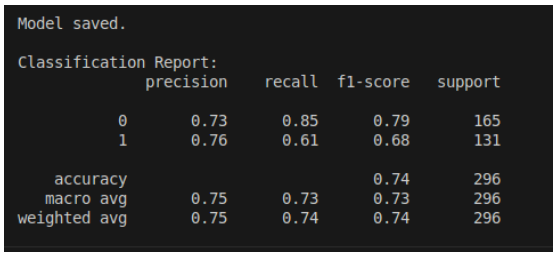

Figure 5: Classification report

5 Discussion

The model performs moderately well and shows clear learning capability. However, it is more effective in detecting normal cells than sickle cells, which is beyond the model's intended purpose. The model achieved 74% accuracy with strong and balanced performance on normal cell (class 0). Lower recall for sickle cell (0.61) suggests that the model struggles to correctly identify sickle cell, possibly due to class imbalance. Therefore, there is a need for improvement in capturing sickle cell (class 1) instances.

6 Conclusion

Sickle cell is an inherited disease caused by a mutation. Deep learning can be used to automate cell classification and make this task less laborious. However, there are considerations that need to be made to obtain a high classification accuracy such as adding more layers to the network or/and balancing the dataset.

References

[1] A. Ashley-Koch, Q. Yang, and R. S. Olney, "Sickle hemoglobin (hb s) allele and sickle cell disease: A huge review," *American Journal of Epidemiology*, vol. 151, no. 9, pp. 839–845, 2000.

[2] N. G. Goswami, A. Goswami, N. Sampathila, and G. M. Bairy, "Sickle cell classification using deep learning," in *2023 3rd International Conference on Intelligent Technologies (CONIT)*, pp. 1–6, 2023.

[3] A. Howard, M. Sandler, B. Chen, W. Wang, L. Chen, M. Tan, G. Chu, V. Vasudevan, Y. Zhu, R. Pang, H. Adam, and Q. Le, "Searching for mobilenetv3," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 1314–1324, 2019.