Communication in Physical Sciences, 2022, 8(1): 9-22

Deep Convolutional Neural Network Model for Detection of Sickle Cell Anemia in Peripheral Blood Images

Emmanuel Gbenga Dada*, David Opeoluwa Oyewola and Stephen Bassi Joseph
Received: 02 February 2022/Accepted 03 March 2022/Published online: 08 March 2022

Abstract: *Sickle Cell Disease (SCD) is a disorder of red blood cells (RBC). The number of SCD patients is rising daily. The lifespan of people is reduced by this deadly disease. Statistics show that over twenty five percent of people living in the Central and West Africa region are suffering from this malady. Many of the nations in this part of the world are deficient in the essential means of detecting and treating several illnesses of which SCD is one of them. Infant mortality rates are considerably greater in these countries. The conventional techniques for SCD diagnosis are expensive, error-prone, time consuming, and require the services of medical experts. As a result, there is a pressing need to develop cost-effective and controllable approaches for the early detection and diagnosis of SCD. This paper presents novel techniques that use Plain Convolution Neural Networks (PCNN) with 15 layers and 48 layers, data augmentation of Plain Convolution Network with 48 layers (DAPN-48), Very Deep Convolutional Networks for Large Scale Image Recognition with 19 layers (VGG19), and Residual Networks with 50 layers (RESNET-50) for detecting SCD from peripheral blood image samples. Results obtained from our experiments indicated that PCNN-15 and DAPN-48 outperform PCNN-48 with sensitivity and balanced Accuracy between 99-100%. A comparison was made between the performance of PCNN-15, PCNN-48, DAPN-48, VGG19 and RESNET-50. The results attained by the proposed approaches demonstrated that our techniques are appropriate for the diagnosis of SCD, and thereby recommended for application to sickle cell image detection.*

Keywords: *Sickle cell anemia; red blood cells; erythrocytes; convolutional neural network; classification*

Emmanuel Gbenga Dada*
Department of Mathematical Sciences, Faculty of Science, University of Maiduguri, Borno State, Nigeria
Email: gbengadada@unimaid.edu.ng
Orcid id: 0000-0002-1132-5447

David Opeoluwa Oyewola
Department of Mathematics & Computer Science, Federal University Kashere P.M.B 0182, Gombe, Nigeria
Email: davidakaprof01@yahoo.com
Orcid id: 0000-0001-9638-8764

Stephen Bassi Joseph
Department of Computer Engineering, Faculty of Engineering, University of Maiduguri, Borno State, Nigeria
Email: sjbassi74@unimaid.edu.ng
Orcid id: 0000-0001-5701-2633

1.0 Introduction

Sickle cell disease (SCD) is a blood disorder that runs in families. Defective hemoglobin, an oxygen-carrying pigment and the most abundant protein in the body red blood cells (RBC) provide proof of it. Sickle cell illness reduces the amount of oxygen delivered to the body tissues (Sahu *et al*., 2015). The most common form of SCD is sickle cell anemia (SCA). This illness can be caused by a lack of hemoglobin in the RBCs. RBCs are made in the body's soft marrow, which is found in the larger bones. The RBCs that have perfect hemoglobin are usually uniform, cylindrical, and elastic, similar to doughnuts that are not perforated. They can comfortably navigate the arteries and veins (Stuart and Nagel, 2004). Cells that have been infected with

sickle cell hemoglobin are usually rigid and gummy. When the oxygen in them is exhausted, they take the shape of a sickle or hemispherical, similar to alphabet C (Sickle Cell Disease, 2020). These hematopoietic cells stay close to one other and therefore find it very difficult to travel through the network of arteries, arterioles, capillaries, venules, and veins. The resultant effect of this is the obstruction of tiny blood vessels and the mobility of disease-free, sane oxygen-carrying blood. The impediment of these blood vessels can trigger discomfort, organ failure occurring in main organs nourished by the circulatory system, incidents of neurophysiological maladies, and eventual death of the victim. Overtiredness, abnormal heartbeat, difficulty in breathing are all symptoms of SCA. People suffering from sickle cell anemic usually have hands and foot diseases, trouble inhaling or exhaling enough oxygen, persistent pain, and visual impairment. It also can increase the danger of the victim being attacked by aspergillosis or pathogenic agent from internal or external sources (Chy and Rahaman, 2019). According to the World Health Organization (WHO), around 5% of the world population has genes linked to hemoglobin disorders, the most common of which are sickle cell disease and Mediterranean anemia. In certain geographical areas of the world, as high as 25% of the populace are carriers of these genes. People having sickle cell disease are greater in number in the continent of Africa. Children born with acute hemoglobin syndromes every year are above 300,000 (WHO, 2005). The SCD is a very important public health interest. An average of 75,000 cases of admission to hospital for treatment owing to SCD happened in the US between 1989 and 1993. This attracted an economic cost of around 475 million dollars. In 2005, medical expenditures for children with sickle cell disease cost the public health insurance program for low-income families or individuals an average of 11,702 dollars. Also, 14,772 US dollars was spent for children that are subscribed under group health insurance. Nearly 40% of these two categories had spent a minimum of one night in the hospital (SCD, 2020).

It is highly necessary to keep an eye on sickle cell patients. This entails the examination of peripheral blood smear using an optical microscope which is a timewasting method. Moreover, it is a necessity that this process is carried out by an expert due to the personalised characteristic of the examination. The setback of this approach is its astronomical percentage of error. These predicaments become worsened when there is a huge number of patients. An approach to assess the direct observation and treatment of patients is to itemize the different kinds of RBCs using their structure. These cells can be classified as normal (discocyte), neonate or abnormal elongated (sickle cells). It can also have some other kinds of defects. There are many disadvantages associated with these evaluation measures. It varies from specialists' divergences of view which resulted in the problem of the inability to form a benchmark evaluation.

There are many approaches for identifying, structuring and investigating digital image frameworks for successive use in suitable classification algorithms. The procedure for removing the important attributes from the explored edge is very vital. To investigate the flaws that the erythrocytes contained in a sample of blood viewed under a microscope have, a variety of approaches have been developed.

The existing techniques used for detecting SCA are characterised by several drawbacks. It is very difficult to get the accurate classification of SCA in erythrocytes for reasons which include imbrication cells, unavailability of training data, blending of light and dark areas between cells and backdrop. Other reasons are the inhomogeneous and multifaceted structures and dimensions of cells (Alzubaidi et al., 2020). Every one of them has its attributes, strengths and weaknesses.

The following are the key contributions of this work:

i. A concise explanation of the state-of-the-art approaches for detecting and diagnosing SCA was done
ii. Deep convolutional neural network models for the classification of erythrocytes into three groups: discocyte (normal), elongated (sickle cells) and neonate (Oval) was developed.
iii. Experimental analysis on the erythrocytesIDB dataset shows that the proposed models had an average accuracy of 89.33-99.92%.

Several techniques for detecting SCA in red blood smears have been proposed in the literature. Barpanda (2013) proposed novel image processing methods that can automatically detect sickle cells anemia that exists in RBCs samples. The image dataset used was obtained from camera attached microscope. The erythrocytes and sickle cells in the microscopic smear are detected using image segmentation and clustering. Many of the state-of-the-art performance measures were not employed, therefore the effectiveness of the suggested approach was not assessed. The k-means and fuzzy c-means clustering methods were also not compared to high-performance clustering algorithms. Ademola *et al*. (2018) created an ensemble machine learning (ML) model for assessing the severity of sickle cell disease among children with SCA. The ensemble model was built using Nave Bayes, Decision trees, and Support Vector Machines (SVM). The suggested ensemble approach performance is relatively low, according to simulation findings. As a result, there is a pressing need to design more effective and high-performing methods. For the categorization of RBCs in sickle cell disease patients, Xu *et al*. (2017) developed deep convolutional neural networks. The paper presented a faster technique that permits a larger number of samples to be processed in the same, or less, time compared to other existing methods. Experiments were carried out by applying CNN on 7000 single RBC of sickle cell patients' images using 5 fold cross-validation. Simulation results demonstrated that the model has high performance in predicting eight types of sickle cell disease. There is however need to improve on the classification accuracy of the method.

Idowu *et al*.(2015) applied a fuzzy logic-based predictive model to detect and determine the probability that a patient has been infected with SCA. The degree of fetus hemoglobin, the amount of anemia, and the patient's genotype are the three input factors utilized in the prediction. Their technique's effectiveness was not assessed. Chy and Rahaman (2019) conducted a comparison study utilizing pictures of blood samples taken from patients to identify SCA using different machine learning (ML) models. To filter out the noise and convert a gray image to a high-quality image format, pre-processing activities were performed. To categorize the image into normal and sickle cells, three machine learning models were utilized. The performance of the ML models understudy is very low. It, therefore, means that develop more high performing models.

For the categorization of RBCs into sickle cells and normal blood cells, Uike and Thorat (2020) examined several automated image processing and multiclass classifier methods. The authors opined that the approaches reviewed have demonstrated good results concerning accuracy, sensitivity and execution time. It is however observed that there is still room for progressing on the effectiveness and classification accuracy of abnormal blood cells in a diverse group by classifier having more than one class. They encouraged the development of high performance image processing algorithms with enhanced execution speed and classification accuracy.

An image processing approach for automated detection of SCA was presented by Chy and Rahaman (2019). Experiments were performed on RBCs of images of blood. This is followed by the conversion of grayscale image to high quality image to enhance it. The median filter is applied to remove noise from the image. The RBCs are segmented using the threshold segmentation technique and image processing operations performed

on the image. The necessary features are extracted, and SVM is used for image classification. Statistical findings showed that the performance of the technique is satisfactory but needs to be improved upon. Moreover, only 80 images were used for the experiment. This is too small to authenticate the effectiveness of the technique. Also, the authors did not compare the performance of their system with other high performing ML algorithms.

Sharma, Rathore and Vyas (2016) applied the k-nearest neighbour algorithm to classify blood smear microscopic images into sickle cells, dacryocyte and ovalocytes. The image was segmented after several preparation procedures such as noise reduction using a median filter. Though the suggested approach performed admirably on the images, its usefulness remains in question because just 100 images were utilized for both training and testing. Also, the performance of their technique in terms of classification accuracy and sensitivity is very low. Khizra (2018) applied natural language processing and ML to investigate clinical notes in SCD patients. The authors in their work concentrated on information about the patient, pain, pain sentiment and scores of the pain using the sickle cell disease data. To extract key characteristics from the dataset, six machine learning methods were utilized. The performance of their system is good as it provides an approach to gain insight into the suffering of sickle cell patients. However, there is a need for further improvement.

Moreover, Elsalamony (2015) proposed a novel algorithm for detecting SCA. The proposed system can identify and compute the number of deformed RBCs in a blood smeared image, irrespective of whether the cells are obscure or overlapped. The images were trained and tested using artificial neural networks (ANN) and decision tree algorithms. Experimental results showed that the proposed method performed satisfactorily, though there is still room for improvement. However, the number of images used for the experiment is small and it will be difficult to ascertain that the proposed method will perform effectively with a large dataset. From the literature that has been reviewed, it is evident that much research has been conducted in the field of sickle anemia detection using different approaches. However, each work has its own set of restrictions. These flaws prompted the creation of this deep convolutional deep learning model for SCA detection in this work. The system has the ability to overcome the problems that exist in current systems.

2.0 Materials and Methods

The dataset utilized in this study is described in detail in this section. The training and testing techniques for the model were also discussed. This part also included a flowchart illustrating our innovative method. There was a discussion about the parameter settings and the values utilized in our tests. Figure 1 depicts the suggested process for this approach.

2.1 Dataset

The erythrocytesIDB dataset was used for all the experiments conducted for this work.

Blood samples with RBC pictures were obtained from individuals with SCD at the General Hospital in Santiago de Cuba (Erythrocytesidb, 2020). There are 196 images with 626 images of different cells categorised as circular (202), elongated (211), or other deformation (213). Each cell is 80 × 80 pixels in size. The doctor standards were utilized as a specialist approach for validating the findings produced by the algorithms used to classify the cells contained in the RBC images. The samples of individual cells in the dataset are shown in Fig. 1.

3.0 *Proposed Method*

Convolutional neural networks are bio-inspired networks that can categorize images and detect objects in computers. Every layer in a convolutional network is a three-dimensional grid with height, breadth, and depth. The term "depth" refers to the number of channels in each layer in the input image, such as primary color channels like blue, green, and red, or the number of hidden layers

in feature maps. The network operates similarly to a feed forward neural network, with the exception that the convolutional layers are spatially organized (Krizhevsky *et al*., 2017). Presented in Fig. 2 is the flowchart of the PCNN and DAPN.

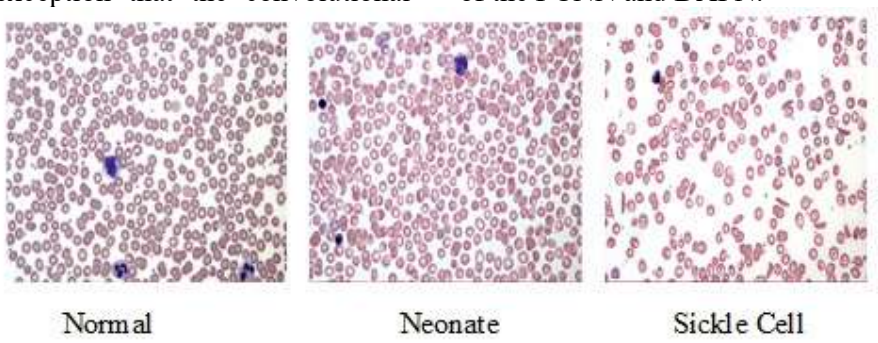


Fig. 1: Images of healthy and Unhealthy erythrocytesIDB I

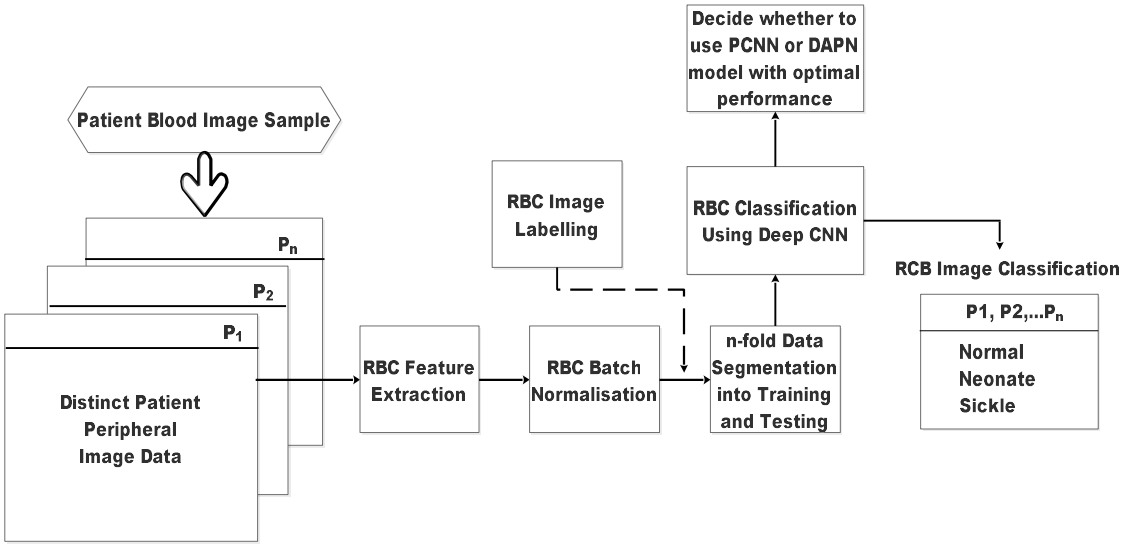
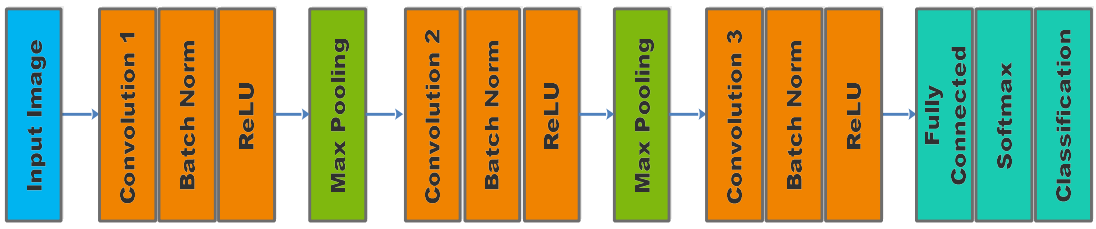


Fig. 2: Flowchart of the proposed PCNN, DAPN, VGG19, and RESNET-50 Sickle Cell Anemia classification

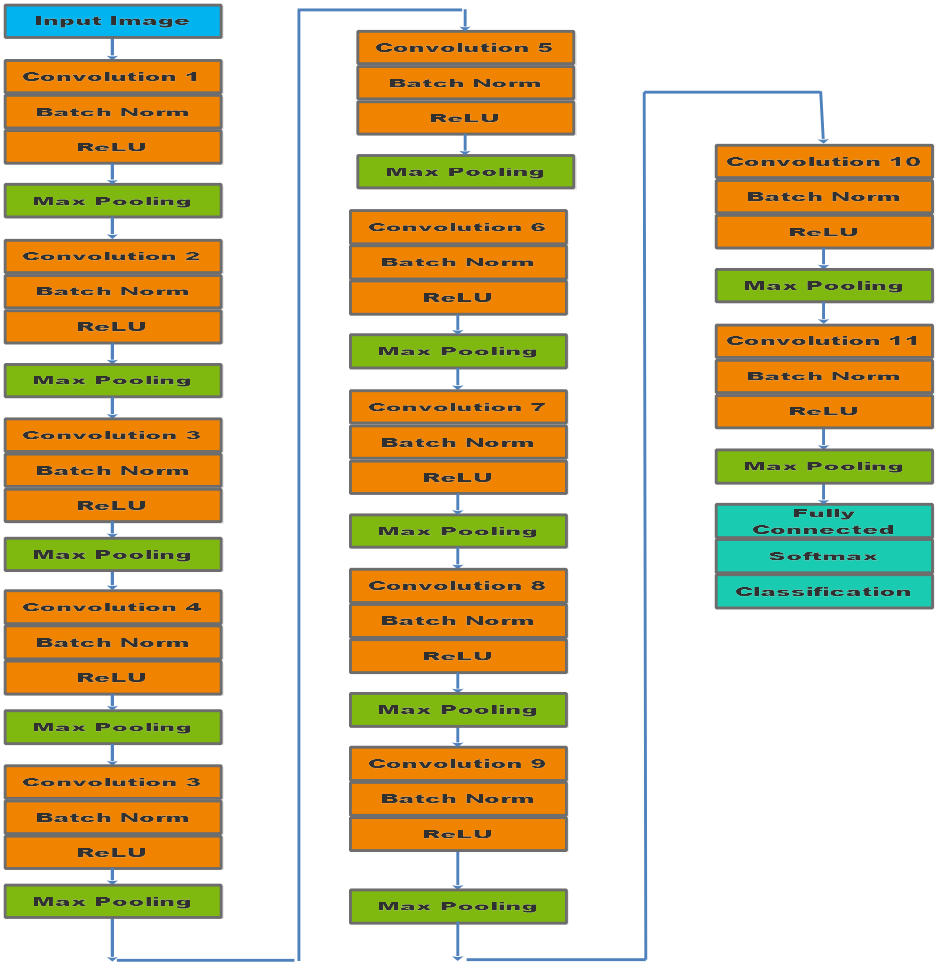
Depicted in Fig. 3 is the proposed techniques for image classification of SCA. Convolution, pooling, and rectified linear units are the three types of layers most commonly seen in convolutional neural networks. Filters or kernels are three-dimensional structural components that organize the parameters. Normally, the filters are square.



(A) Plain Convolution Neural Networks with 15 layers (PCNN-15)

The filter's usual application dimensions are much smaller than the filters themselves. The objective, methodology, and model architecture were all taken into account when

creating the model utilized in this study. In this study, we considered three different models for categorizing SCA. The models are Plain Convolution Neural Networks (PCNN-15), Plain Convolution Neural Networks (PCNN-48) and data augmentation of Plain Convolution Network (DAPN-48).



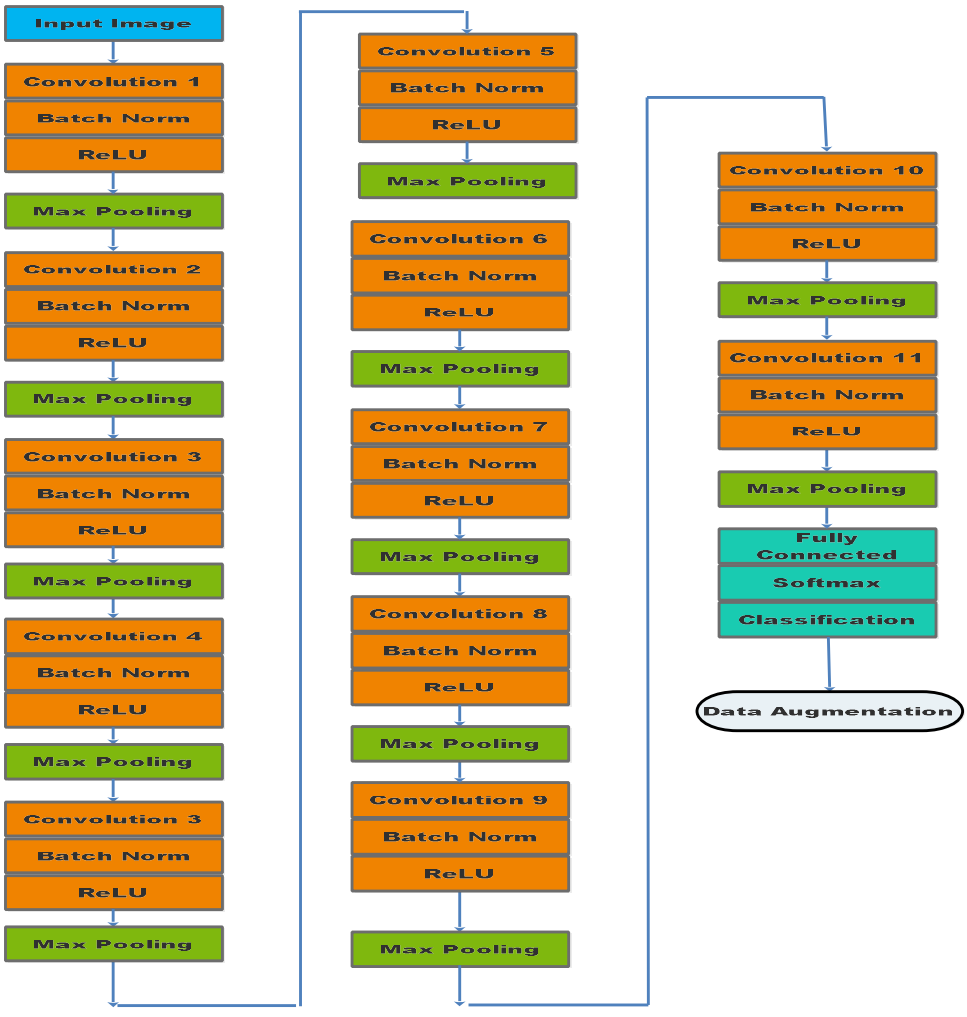
(B) Plain Convolution Neural Networks with 48 layers (PCNN-48)

3.1 Plain convolution neural networks (PCNN-15)

Normal, neonatal, and sickle cell anemia are the three types of sickle cell images utilized in this research. The image input layer, three convolution layers, three batch normalization layers, three rectified linear units (ReLU), two max-pooling layers, one fully connected layer, one softmax layer, and one classification layer are all included in Plain Convolution Neural Networks with 15 layers (PCNN-15). All of the layers of Plain Convolutional Neural Networks (PCNN-15) are linked sequentially (Yao, Xu and Zhao, 2020). In this study, the image input size was 30 heights, 30 weights, and 3 channels. Filter size, number of filters, and padding are all varied in the three convolutional layers utilized. The three convolutional layers use a 3 by 3 filter size with filter numbers ranging from 15 to 30. The spatial output size is the same as the input size with a padding of 1. Batch normalization layers equalize activation and gradients propagating throughout the network, allowing for

network training optimization (Zhang *et al*., 2018). On the other hand, ReLU layers increase network training and minimize network sensitivity. The size of max pooling layer utilized in research is 1 this enables us to down-sampling the operation (Zhang and Wallace, 2015).

Down-examining permits the quantity of estimation required per layer to be increased without expanding the number of filters. The convolutional layer and batch normalization layers are trailed by completely associated, softmax and classification layers (Chang *et al*., 2020). As the name infers, a completely associated layer fuses every one of the past layers and recognizes the bigger patterns. The completely linked layer is five which corresponds to the five classes. The softmax activation function normalizes the fully connected layer performance while the classification layer is utilized to assigns the input to a class that is unique to each of these classes based on the probabilities provided by the softmax activation function (Kanai *et al*., 2018). Presented in table 1 is the parameter settings for PCNN-15 used in our experiment:



(C) Data Augmentation of Plain Convolution Networks with 48 layers (DAPN-48)

Fig. 3: Proposed Techniques for Image Classification of SCA

3.2 Plain convolution neural networks (PCNN-48)

In the classification of sickle cell images which are normal, neonate and sickle cell anemia. The image input layer, eleven convolution layers, eleven batch normalization layers, eleven rectified linear units (ReLU), eleven max pooling levels, one fully connected layer, one softmax layer, and one classification layer are all included in this research as PCNN-48. In Plain Convolutional Neural Networks (PCNN-48) every one of the layers is linked consecutively. In this study, image input sizes of 30 heights, 30 weights, and 3 channels were employed. The three convolutional layers utilized comprises of the various filter size, several filters and padding. The three convolutional layer uses the filter size of 3 by 3 while the filter numbers increment from 15 to 30. Padding of one guarantees that the spatial yield size matches the input size. Batch normalization layers normalize the activation and gradients propagating across the network, allowing network training to be optimized (Ioffe and Szegedy, 2015). ReLU layers then again further develop network training and decrease network affectability. The size of max pooling layer used in research is one, this empowers us to down-sample the activity (Saeedan *et al*., 2018). The parameter settings and values for PCNN-48 used in our experiments are depicted in Table 2.

Table 1: **PCNN-15 parameter setting**

Parameter	Value
Layers	15
Image Input Layer	1
Convolution Layer	3
Batch Normalization Layer	3
ReLU Layer	3
Max Pooling Layer	2
Fully connected layer	1
Soft max layer	1
Classification layer	1

(a) Translation

Shifting images to the left, right, and down may be useful to prevent positional bias in the images. It is translated between 3 pixels and -3 pixels in the original image. Each augmented image is chosen at random from a uniform distribution throughout the range, translated by a pixel distance.

Table 2: PCNN-48 parameter setting

Parameter	Value
Layers	48
Image Input Layer	1
Convolution Layer	11
Batch Normalization Layer	11
ReLU Layer	11
Max Pooling Layer	11
Fully connected layer	1
Softmax layer	1
Classification layer	1

(b) Reflection

As a logical scalar, random reflection in the left-right direction is defined. Each Sickle cell image has a 50% chance of being reflected horizontally. The parameter settings and values for the DAPN-48 utilized in this experiment are shown in Table 3.

Table 3: **DAPN-48 parameter setting**

Parameter	Value
Layers	48
Image Input Layer	1
Convolution Layer	11
Batch Normalization Layer	11
ReLU Layer	11
Max Pooling Layer	11
Fully connected layer	11
Softmax layer	1
Classification layer	1

3.4 Very Deep convolutional networks for large scale image recognition (VGG19)

VGG19 is a 19-layer variant of the VGGNet network (16 convolution layers, 3 fully connected layers, 5 MaxPool layers and 1 Softmax layer). In both variants, VGGNet contains two completely connected layers with 4096 channels each, followed by another fully connected layer with 1000 channels to predict 1000 labels. In the final

fully linked layer, the Softmax layer is utilized for categorizing. The parameter settings and values for VGG19 utilized in this experiment are shown in Table 4.

Table 4: VGG19 parameter setting

Parameter	Value
Layers	47
Image Input Layer	1
Convolution Layer	16
Batch Normalization Layer	0
Dropout Layer	2
ReLU Layer	18
Max Pooling Layer	5
Fully connected layer	3
Softmax layer	1
Classification layer	1

3.5 Residual networks (RESNET-50)

ResNet-50 is a deep convolutional neural network with 50 layers. The network can categorize images into over 1000 distinct item types. As a result, the network has learned a wide range of rich feature representations for a wide range of image. The network can handle images up to 224 × 224 pixels in size. Table 5 shows the parameter settings and values for our experiment.

Table 5: RESNET-50 parameter setting

Parameter	Value
Layers	50
Image Input Layer	1
Convolution Layer	46
Batch Normalization Layer	0
ReLU Layer	0
Max Pooling Layer	0
Fully connected layer	1
Softmax layer	1
Classification layer	1

4.0 Results and Discussion

In this paper, we compared Plain Convolution Neural Network with 15 layers (PCNN-15), Plain Convolution Neural Network with 48 layers (PCNN-48) and Data Augmentation of PCNN-48 (DAPN-48). The overall performance of PCNN-15, PCNN-48, DAPN-48, and VGG19 and RESNET-50 classification model was evaluated using eight different assessment measures: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Scaled Error (MASE), Specificity, Sensitivity and Balanced Accuracy.

The comparison of sensitivity, specificity, and balancing accuracy is shown in Table 4. According to Table 4, PCNN-15 produced the greatest results when classifying RBC images as normal, neonatal, or sickle cell anemia. PCNN-15 has a sensitivity and balanced accuracy of 99-100 percent. This demonstrates that the use of 15 layers on PCNN-15 has a significant influence on getting improved results.

Table 5 shows the performance metrics of PCNN-48 model. PCNN-48 failed to classify RBC images as shown in the performance metrics such as sensitivity and balanced accuracy. This shows that increasing the number of layers of PCNN from 15 to 48 layers does have any positive influence on the classification accuracy of SCA images. In Table 6, DAPN-48 performed better than PCNN-48. This demonstrates that the combination of PCNN-48 and data augmentation has the potential to increase SCA image classification efficiency. The performance evaluation categorization of the VGG19 model is shown in table 7. VGG19 has a sensitivity and balanced accuracy range of 95-100 percent. This demonstrates that VGG19 is a promising Deep Learning method for sickle cell image classification. The results of the experiment with RESNET-50 are shown in table 9. VGG19 has a sensitivity and balanced accuracy range of 97-100 percent. This indicates that the performance of RESNET-50 is superior to that of VGG19. Table 9 depicts the performance metrics for each criterion utilized in this paper. As shown in Table 9, PCNN-15 and DAPN-48 have the lowest error when compared to the other model. An analysis of the results of the PCNN-15 and DAPN-48 model's performance, as shown in

tables 4 and 6, revealed that the generated models properly classified SCA images in 99.92 percent of actual cases, with a misclassification rate of 0.8 percent of actual cases. On the dataset employed, the findings of PCNN-15 and DAPN-48 models showed that an average of 99.92 percent of the model predictions were correct. The findings of the suggested models also show that for the classification of SCA blood images, PCNN-15 and DAPN-48 displayed the highest proficiency in properly identifying SCA images. They also have the lowest ability to misclassify the SCA blood images. PCNN-15 and DAPN-48 also have average specificity of 99.83%.

Table 4: **Metrics Evaluation of PCNN-15**

Algorithms	**Sensitivity (%)**	**Specificity (%)**	**Balanced Accuracy (%)**
Normal	100	99.67	100
Neonate	100	100	99.83
Sickle cell	100	99.83	99.92

Table 5: **Performance Evaluation Classification of PCNN-48**

Algorithms	**Sensitivity (%)**	**Specificity (%)**	**Balanced Accuracy (%)**
Normal	100	57.33	100
Neonate	100	100	78.67
Sickle cell	100	78.67	89.33

The proposed models correctly classified SCA images on average of 89.3 percent, with an incorrect classification rate of 11.7 percent, based on the performance of the PCNN-48 model presented in table 5. The results of the PCNN-48 model on the dataset used in the experiment showed that the model correctly predicted 89.3 percent. The PCNN-48 model had the worst results of all the models described in this paper.

Table 6: **Performance Evaluation Classification of DAPN-48**

Algorithms	**Sensitivity (%)**	**Specificity (%)**	**Balanced Accuracy (%)**
Normal	100	99.67	100
Neonate	100	100	99.83
Sickle cell	100	99.83	99.92

Table 7: Performance Evaluation Classification of VGG19

Algorithms	**Sensitivity (%)**	**Specificity (%)**	**Balanced Accuracy (%)**
Normal	100	96.78	100
Neonate	100	100	98.67
Sickle cell	100	95.67	99.33

In terms of the performance of the VGG19 model, as shown in Table 7, the models properly classified SCA images in an average of 97.48 percent of actual cases, with a classification error rate of 2.52 percent of actual cases. The VGG19 model results revealed that an average of 97.48 percent of the model's predictions were correct. The

average specificity of VGG19 is 97.48 percent.

Table 8: Performance Evaluation Classification of RESNET-50

Algorithms	Sensitivity (%)	Specificity (%)	Balanced Accuracy (%)
Normal	100	98.97	100
Neonate	100	100	97.83
Sickle cell	100	97.93	98.92

In terms of the RESNET-50 model performance, as shown in table 8, the models properly classified SCA images in an average of 98.99 percent of actual cases, with a classification error rate of 2.52 percent of actual cases. The VGG19 model results revealed that an average of 97.48 percent of the model predictions were correct. The average specificity of VGG19 is 97.48 percent.

Table 9: Comparative Analysis of PCNN-15, PCNN-48, DAPN-48, VGG19 and RESNET-50

Algorithms	MAE	RMSE	MASE
PCNN-15	0.0011111	0.0333333	0.124444
PCNN-48	0.1422222	0.3771236	15.92889
DAPN-48	0.0011111	0.0333333	0.124444
VGG19	0.0332123	0.2323451	4.459873
RESNET-50	0.0321111	0.2222121	4.232134

Figure 4 shows the bar charts for the results generated from PCNN-15, PCNN-48, DAPN-48, VGG19 and RESNET-50 models

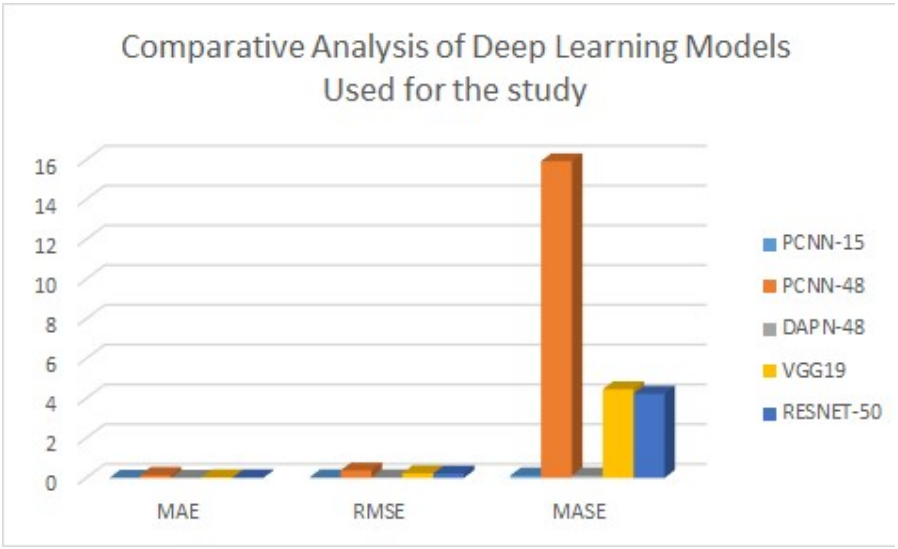


Fig. 4: Result of erythrocytes IDB dataset classification

Table 10 show evidence of the performance of our proposed models to that of other high-performing algorithms. The algorithms utilized by the authors in Chy and Rahaman (2019), such as k Nearest Neighbour (kNN), Support Vector Machine (SVM), and

Extreme Learning Machine (ELM), were compared to our proposed models. The performance of the approach utilized in Chy & Rahaman (2018) was also compared to that of our models. The comparisons reveal that our suggested models outperform the prior models utilized in the literature for SCA detection.

Table 10: Performance Comparison of Proposed Model with other Techniques

Algorithms	Accuracy (%)	Sensitivity (%)	Specificity (%)
kNN (Chy & Rahaman, 2019)	73.33	75.00	66.67
SVM (Chy & Rahaman, 2019)	83.33	83.33	83.33
ELM (Chy & Rahaman, 2019)	87.73	87.50	83.33
Image Processing [13]	95.00	96.55	92.35
PCNN-15	99.83	99.92	100
PCNN-48	78.67	89.33	100
DAPN-48	99.83	99.92	100
VGG19	97.48	97.52	100
RESNET-50	98.83	98.92	100

4.0 Conclusions

This paper presents Plain Convolution Neural Network (PCNN) and Data Augmentation Convolution Network (DAPN) techniques for classifying RBC in outlying blood samples as normal, neonate, and sickle cells. The fusion of data augmentation and CNN has the advantage of enhancing the detection and classification accuracy of SCA images. The results of our experiments showed that data augmentation of convolutional neural networks and plain convolution neural networks with 15 layers perform well than plain convolution neural networks with 48 layers. It was found that the integration of CNN with other methods had the advantage of high classification accuracy. Findings show that deep learning methods with less expertise and resources in SCA diagnosis can produce faster and more efficient results. In the future, we intend to engage in interdisciplinary academic research that will combine medical professionals' knowledge and experience with other deep learning-based systems that are yet to be exploited but have the potential to increase the effectiveness of SCA diagnosis.

5.0 Acknowledgements

We thank erythrocytes IDB Coordinator in person of Dr. Antoni Jaume-i-Capo for releasing the dataset that we used for this work.

6.0 References

Alzubaidi, L., Fadhel, M. A., Al-Shamma, O., Zhang, J., & Duan, Y. (2020). Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis. *Electronics*, *9*(3), 427. https://doi.org/10.3390/electronics9030427

Balogun, J. A., Adeyemo, T., Egejuru, N. C., & Idowu, P. A. (2018). An Ensemble Model of Machine Learning Algorithms for the Severity of Sickle Cell Disease (SCD). *Computer Reviews Journal*, *2*, pp. 331-346.

Barpanda, S. S. (2013). *Use of Image Processing Techniques to Automatically Diagnose Sickle-Cell Anemia Present in Red Blood Cells Smear* (Doctoral dissertation), National Institute of Technology Rourkela, India.

Chang, H. H., Liu, L., & Yi, Y. (2020). Deep Echo State Q-Network (DEQN) and Its Application in Dynamic Spectrum Sharing for 5G and Beyond. *IEEE Transactions on Neural Networks and Learning Systems*.

Chy, T. S., & Rahaman, M. A. (2018, November). Automatic sickle cell anemia detection using image processing technique. In *2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)* (pp. 1-4). IEEE.

Chy, T. S., & Rahaman, M. A. (2019), *A comparative analysis by knn, svm & elm classification to detect sickle cell anemia. In 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)* (pp. 455-459). IEEE.

Elsalamony, H. A. (2015). Detecting distorted and benign blood cells using the

Hough transform based on neural networks and decision trees. In *Emerging Trends in Image Processing, Computer Vision and Pattern Recognition* (pp. 457-473). Morgan Kaufmann.

Erythrocytesidb (2020). Available at http://erythrocytesidb.uib.es/ Accessed on 05/07/2020.

Idowu, A. P., Aladekomo, T. A., Williams, K. O., & Balogun, J. A. (2015). Predictive Model for Likelihood of Survival of Sickle-Cell Anaemia (SCA) among Peadiatric Patients using Fuzzy Logic. *Transactions on Networks and Communications*, *3, 1, pp.* 31, DOI: 10.14738/tnc.31.842

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448-456). PMLR.

Kanai, S., Fujiwara, Y., Yamanaka, Y., & Adachi, S. (2018). Sigsoftmax: Reanalysis of the softmax bottleneck. *arXiv preprint arXiv:1805.10829*.

Khizra, S. (2018). *Using Natural Language Processing and Machine Learning for Analyzing Clinical Notes in Sickle Cell Disease Patients* (Master Dissertation), Wright State University.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60, 6, pp. 84-90.

Saeedan, F., Weber, N., Goesele, M., & Roth, S. (2018). Detail-preserving pooling in deep networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 9108-9116).

Sahu, M., Biswas, A. K., & Uma, K. (2015). Detection of Sickle Cell Anemia in Red Blood Cell. *A. International Journal of Engineering and Applied Sciences (IJEAS)*, *2*(3), pp. 45-48.

Sharma, V., Rathore, A., & Vyas, G. (2016, August). Detection of sickle cell anaemia and thalassaemia causing abnormalities in thin smear of a human blood sample using image processing. In *2016 International Conference on Inventive Computation Technologies (ICICT),* 3, pp. 1-5.

Sickle Cell Disease(SCD). Centre for disease control and prevention (2019). Available at https://www.cdc.gov/ncbddd/sickle-cell/data.html. Accessed 2020-07-20.

Sickle Cell Disease. Johns Hopkins Medicine. (2020) Available at https://www.hopkinsmedicine.org/health/conditions-and-diseases/sickle-cell-diseaseon accessed July 10, 2020.

Stuart, M. J., & Nagel, R. L. (2004). Sickle-cell disease. *The Lancet*, *364, 9442, pp* 1343-1360.

Uike, D., & Thorat, S. (2020, March). *Implementation of Multiclass Algorithm for Sickle Cell Identification and Categorization–A Review. In 2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 300-303). IEEE.

World Health Organization (2005). Sickle-cell anaemia: report by the secretariat. In *Sickle-cell anaemia: report by the Secretariat*.

Xu, M., Papageorgiou, D. P., Abidi, S. Z., Dao, M., Zhao, H., & Karniadakis, G. E. (2017). A deep convolutional neural network for classification of red blood cells in sickle cell anemia. *PLoS computational Biology*, 13, 10, pp. e1005746, doi: 10.1371/journal.pcbi.1005746

Yao, L. S., Xu, G. M., & Zhao, F. (2020). Pooling method on PCNN in convolutional neural network. In *Journal of Physics: Conference Series*, 1486, 2, 022026). doi:10.1088/1742-6596/1486/2/022026

Zhang, B., Zhao, Q., Feng, W., & Lyu, S. (2018). AlphaMEX: A smarter global pooling method for convolutional neural networks. *Neurocomputing*, 321, pp. 36-48.

Zhang, Y., & Wallace, B. (2015). A sensitivity analysis of (and practitioners'

guide to) convolutional neural networks for sentence classification. *arXiv preprint arXiv:1510.03820*.

Zheng, Q., Yang, M., Tian, X., Jiang, N., & Wang, D. (2020). A full stage data augmentation method in deep convolutional neural network for natural image classification. *Discrete Dynamics in Nature and Society*, *2020*.

Conflict of Interest

The authors declared no conflict of interest