INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

Sickle Cell Anemia Detection Using Convolutional Neural Network

Dethe Krushna[1], Mungase Nilesh[2], Shinde Tushar[3], Prof. N. S. Sapike[4]

Department of Computer Engineering, Vishwabharati Academy's College of Engineering, Ahmednagar, India[1,2,3,4]

Professor, Department of Computer Engineering, Vishwabharati Academy's College of Engineering, Ahmednagar, India[5]

ABSTRACT: Sickle Cell Anemia (SCA) is a prevalent genetic blood disorder characterized by the abnormal shape of red blood cells due to abnormal hemoglobin. Timely and precise detection of SCA is crucial for effective treatment and management. This research introduces an innovative methodology for SCA detection utilizing Convolutional Neural Networks (CNNs), a category of deep learning algorithms renowned for their efficacy in image analysis tasks. The proposed approach involves leveraging microscopic images of blood smears obtained from patients. These images undergo preprocessing to optimize contrast, normalize intensities, and eliminate noise. Subsequently, CNN architectures are deployed to automatically extract hierarchical features from the blood smear images. The trained CNN model acquires discriminative features directly from the images, adeptly discerning subtle distinctions between normal and sickled red blood cells (RBCs).

KEYWORDS: Sickle Cell Anemia, convolutional neural network, deep learning, image analysis, blood smear images, early detection, diagnostic tool.

I. INTRODUCTION

Sickle Cell Anemia (SCA), a hereditary blood disorder, is characterized by the transformation of normally circular red blood cells into a sickle shape, impairing their smooth movement through blood vessels and leading to vessel occlusions. RBCs play

a vital role in oxygen transportation to various body parts, and the entrapment of sickle-shaped RBCs disrupts oxygen supply, resulting in chronic anemia. Notably, the lifespan of sickle cell RBCs is dramatically shorter, lasting only 10-20 days compared to the normal RBC lifespan of 120 days. This condition can adversely affect vital organs such as the heart, lungs, and kidneys, sometimes culminating in fatal consequences. Sickle Cell Disease (SCD) is caused by mutations in genes coding for the hemoglobin protein, and if both parents carry this gene, there is a one in four chance that the child will have SCD.

While significant strides have been made in treating SCD, the disease remains incurable and can only be managed through early detection and appropriate medication. Conventional diagnostic methods involve manual laboratory tests performed by technicians or the use of a hemocytometer. Manual detection is time-consuming and relies heavily on the expertise of technicians. The alternative, using a hemocytometer, is expensive and may not be accessible in remote or underdeveloped areas.

Consequently, active research has focused on developing automated detection algorithms for SCD. Numerous image processing, machine learning, and deep learning algorithms have been explored for successful RBC counting, segmentation, and identification from erythrocytes. However, many approaches face challenges when dealing with variations in cell size, color, and other parameters. Machine learning classifiers often require manual feature extraction, whereas deep learning networks autonomously extract features from datasets. Deep learning, with its proven success in various image processing applications and biomedical image classification, continues to drive research in the SCD detection domain, seeking higher accuracies despite increased complexities.

II. RELATED WORK

In related works, various studies have employed machine learning techniques, specifically Support Vector Machines (SVMs), for the diagnosis of sickle cell anemia based on patient datasets. Researchers have emphasized the importance of selecting relevant features that significantly influence the presence of sickle cell anemia. Commonly utilized features include hematological parameters, genetic markers, and clinical indicators. The process typically involves the careful

curation of a dataset, incorporating diverse patient profiles. The dataset is then partitioned into training and testing sets to facilitate the SVM model's learning and evaluation phases. Researchers have explored different kernel functions for SVM, such as linear, polynomial, and radial basis function (RBF) kernels, to discern the most suitable for accurate classification.

The SVM model is trained on the training set, learning to distinguish between patients with and without sickle cell anemia based on the selected features. Evaluation metrics like accuracy, precision, recall, F1-score, and ROC curves are commonly employed to assess the model's performance using the testing data. Successful models demonstrate the potential for predicting the presence of sickle cell anemia in new patients based on their feature data, highlighting the significance of machine learning in enhancing diagnostic capabilities in this domain.

The existing methods for diagnosing Malaria and Anemia, including Sickle Cell Anemia, Megaloblastic Anemia, and Thalassemia, often rely on traditional protocols such as the Complete Blood Count (CBC) test. While these tests have proven efficacy, they are time-consuming and may involve complex laboratory procedures. Furthermore, they can be costly and may not be readily accessible in resource-constrained environments. Additionally, traditional diagnostic approaches may require a significant amount of human intervention, leading to the potential for inter-observer variability and delays in obtaining results. Therefore, there is a recognized need for alternative, efficient, and cost-effective diagnostic methods that can enhance the speed and accessibility of blood sample analysis.

The use of Convolutional Neural Networks (CNNs) in the proposed method represents a departure from the conventional reliance on standard laboratory protocols. Although CNNs have shown tremendous success in various image analysis tasks, including medical image classification, there is a need to thoroughly evaluate their performance and accuracy in the context of diagnosing blood disorders. The challenge lies in ensuring that the CNN-based approach not only achieves high accuracy but also addresses potential limitations, such as the generalization of results across diverse datasets and variations in image quality.

Moreover, it is imperative to consider the implications of false positives or false negatives in the diagnosis of blood disorders. A thorough review of existing literature should be conducted to assess the strengths and limitations of CNN-based diagnostic approaches in comparison to traditional methods. Additionally, the sensitivity and specificity of the proposed method need to be critically examined to ensure its reliability across diverse patient populations. As the proposed method claims a total accuracy of 93.4%, it is crucial to scrutinize the methodology used for testing and validating the CNN model to ascertain the robustness of the results.

In summary, while the presented alternative method holds promise in revolutionizing the diagnosis of Malaria and various forms of Anemia, a comprehensive understanding of its advantages and potential drawbacks is essential. A critical examination of existing systems and their limitations will contribute to the refinement and validation of the proposed CNN-based diagnostic approach.

III. PROPOSED METHODOLOGY

The proposed Convolutional Neural Network (CNN) model, comprising a total of 18 layers, is designed to effectively detect and classify medical conditions, specifically focusing on its application to the diagnosis of certain blood disorders. The architecture incorporates five convolutional layers, each utilizing a filter size of (3×3), aimed at processing input data and extracting relevant features. The convolution layers are followed by Rectified Linear Unit (ReLU) activation functions to introduce non-linearity and enhance processing speed. This design is inspired by the concept of neurons' responses to stimuli in the visual cortex, where each convolutional neuron processes data for its assigned receptive field. The subsequent application of max pooling layers serves to further reduce the dimensionality of the data by merging outputs from neuron clusters into single neurons at the next layer.

To prevent overfitting, the architecture includes three fully connected dense layers, with two dropout layers strategically placed between them. The flattening process is employed to convert the output of the convolutional layers into a one-dimensional array, creating a continuous feature vector. This vector serves as input for the densely connected layers, where each neuron receives input from all neurons in the previous layer, ensuring comprehensive connectivity. The dense layers play a crucial role in capturing intricate patterns and relationships within the data. Finally, the model utilizes a softmax function for classification, allowing the output to be categorized into one of three classes, each representing a distinct blood disorder.

This proposed system leverages the hierarchical feature extraction capabilities of CNNs to discern subtle patterns in microscopic images of blood samples. The combination of convolutional layers, activation functions, and pooling layers facilitates the extraction of relevant features, enabling the model to learn discriminative representations for accurate classification. The incorporation of dropout layers mitigates the risk of overfitting, ensuring the model's generalizability. Overall, this CNN-based system offers a robust framework for automated blood disorder diagnosis, contributing to the advancement of medical image analysis and disease detection methodologies.
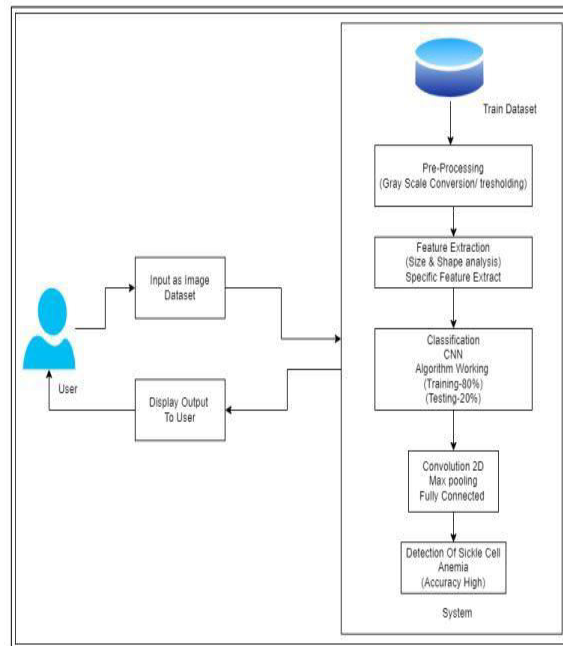
Fig 1. Proposed System Architecture

1.   Image Input Processing:

Preprocessing: Microscopic images of blood smears are subjected to preprocessing steps, including contrast enhancement, intensity normalization, and noise removal. These steps ensure that the input images are optimized for analysis, facilitating the CNN's ability to focus on relevant features associated with Sickle Cell Anemia (SCA).

Standardization: The preprocessing steps contribute to creating a standardized dataset. Standardization is crucial for ensuring that the CNN can effectively learn and differentiate between normal and sickled red blood cells (RBCs). This standardized input serves as the foundation for accurate feature extraction.

2.   Feature Extraction by CNN:

Hierarchical Feature Learning: Convolutional Neural Networks (CNNs) are employed to automatically extract hierarchical features from the preprocessed blood smear images. CNNs excel in learning intricate patterns and structures within images. The model learns to recognize discriminative features directly from the microscopic images, capturing subtle differences in the shapes and characteristics of normal and sickled RBCs.

Discriminative Features: The trained CNN model identifies discriminative features that are indicative of SCA. These features might include the distinct sickle shape of RBCs affected by the disorder. The hierarchical nature of feature extraction allows the CNN to understand complex relationships within the data, enhancing its ability to make accurate predictions.

3.   Classification and Prediction:

Trained Model as a Classifier: The CNN, having learned from a labeled dataset during training, serves as a sophisticated classifier. It categorizes blood smear images into normal or indicative of SCA based on the features it has extracted. The model's ability to generalize from the learned features enables it to analyze new, unseen images and make predictions. Continuous refinement and optimization through additional data contribute to the CNN's accuracy and reliability in the early detection of SCA, providing valuable insights for effective medical intervention and management.

IV. WORKING MODULE

The proposed methodology for the detection of Sickle Cell Anemia (SCA) leverages Convolutional Neural Networks (CNNs), a powerful class of deep learning algorithms widely recognized for their proficiency in image analysis tasks. The system's workflow begins with the acquisition of microscopic images of blood smears obtained from patients. These images are crucial as they provide detailed information about the morphology and characteristics of red blood cells (RBCs).

In the preprocessing stage, the images undergo several enhancements to ensure optimal input for the CNN. This includes contrast adjustment, intensity normalization, and noise removal. These steps are essential for creating a standardized and clean dataset, allowing the CNN to focus on the relevant features associated with SCA. Preprocessing is a crucial step to enhance the network's ability to differentiate between normal and sickled RBCs.

The core of the proposed methodology lies in the utilization of CNN architectures for feature extraction. CNNs are designed to automatically learn hierarchical features from images. In the context of SCA detection, the trained CNN model performs feature extraction directly from the blood smear images. This involves the network learning discriminative features that distinguish normal RBCs from those affected by SCA. The hierarchical nature of feature extraction enables the model to capture subtle differences in the shapes and structures of RBCs, contributing to the accuracy of the diagnostic process.

The trained CNN serves as a sophisticated classifier, effectively categorizing blood smear images into normal or indicative of SCA. The model's ability to generalize from the learned features allows it to analyze new, unseen images and make predictions based on the patterns it has identified during training. Continuous refinement and optimization of the CNN model with additional data contribute to its robustness and accuracy in SCA detection.
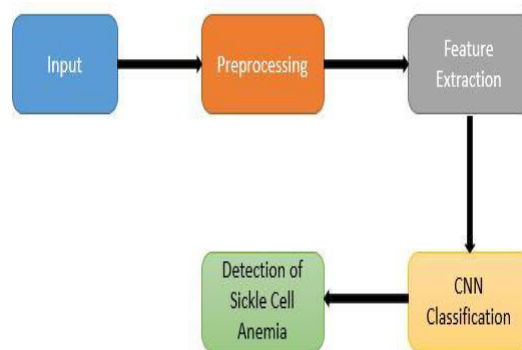
Fig 2: Working Flow Diagram

In summary, the proposed methodology combines advanced image preprocessing techniques with the power of CNNs to automatically extract and learn discriminative features from microscopic blood smear images. This approach holds significant promise for the early and accurate detection of Sickle Cell Anemia, providing a valuable tool for effective medical intervention and management.

V. CONCLUSION

In conclusion, the proposed deep learning model, integrating histogram equalization and gray scaling for preprocessing along with data augmentation, demonstrates a promising approach for the precise classification of blood smear images into normal, sickle-shaped, and other blood components. The layered convolutional neural network effectively addresses the challenge of dataset scarcity through transformations applied via data augmentation techniques. The incorporation of histogram equalization enhances image contrast, contributing to improved model performance. The resulting 5-layer convolutional neural network proves to be a lightweight and low-complexity solution for the accurate classification of red blood cells, consequently enabling the prediction of Sickle Cell Disease. The model's capability for automatic detection and classification suggests its potential utility in clinical settings. As part of future work, enhancements in efficiency are anticipated through the implementation of advanced data augmentation techniques such as transfer learning, maintaining a commitment to maintaining a manageable level of model complexity.

REFERENCES

1. Alhumaid, A. Aleidi, A. Alfakhri, N. Alosaimi, Y. Ali, and M. Alzahrani, "Clinical features and outcome of sickle cell anemia in a tertiary center: A retrospective cohort study," Journal of Applied Hematology, vol. 9, p. 22, 01 2018.
2. M. J. Stuart and R. L. Nagel, "Sickle-cell disease," Lancet (London, England), vol. 364(9442), p. 1343–1360, 2004. [Online]. Available: https://doi.org/10.1016/S0140-6736(04)17192-4

3. Anglin, "Sickle cell disease," Journal of Consumer Health on the Internet, vol. 19, no. 2, pp. 122–131, 2015. [Online]. Available: https://doi.org/10.1080/15398285.2015.1026706

4. L. Alzubaidi, M. A. Fadhel, O. Al-Shamma, J. Zhang, and Y. Duan, "Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis," Electronics, vol. 9, no. 3, p. Article number: 427, March 2020. [Online]. Available: https://eprints.qut.edu.au/204816/

5. P. K. Das, S. Meher, R. Panda, and A. Abraham, "A review of automated methods for the detection of sickle cell disease," IEEE Reviews in Biomedical Engineering, vol. 13, pp. 309–324, 2020.

6. T. S. Chy and M. A. Rahaman, "Automatic sickle cell anemia detection using image processing technique," in 2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE), 2018, pp. 1–4.

7. ACHARYA AND P. KUMAR, "IDENTIFICATION AND RED BLOOD CELL CLASSIFICATION USING COMPUTER AIDED SYSTEM TO DIAGNOSE BLOOD DISORDERS," IN 2017 INTERNATIONAL CONFERENCE ON ADVANCES IN COMPUTING, COMMUNICATIONS AND INFORMATICS (ICACCI), 2017, PP. 2098–2104.

8. S. Knowlton, I. Sencan, Y. Aytar, J. Khoory, M. Heeney, I. Ghiran, and S. Tasoglu, "Sickle cell detection using a smartphone," Scientific reports, vol. 5, p. 15022, 10 2015.

9. H. A. Elsalamony, "Anaemia cells detection based on shape signature using neural networks," Measurement, vol. 104, pp. 50–59, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0263224117301707

10. P. Panneerselvam, "Application of embedded system for a genetic disease, sickle cell anemia," in 2014 International Conference on Advances in Electrical Engineering (ICAEE), 2014, pp. 1–4.

11. P. Rakshit and K. Bhowmik, "Detection of abnormal findings in human rbc in diagnosing sickle cell anaemia using image processing," Procedia Technology, vol. 10, pp. 28–36, 2013, first International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212017313004878

12. M. Abdulraheem Fadhel, A. J. Humaidi, and S. R. Oleiwi, "Image processing-based diagnosis of sickle cell anemia in erythrocytes," in 2017 Annual Conference on New Trends in Information Communications Technology Applications (NTICT), 2017, pp. 203–207.

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

14. M. Khalaf, A. J. Hussain, R. Keight, D. Al-Jumeily, R. Keenan, P. Fergus, and I. O. Idowu, "The utilisation of composite machine learning models for the classification of medical datasets for sickle cell disease," in 2016 Sixth International Conference on Digital Information Processing and Communications (ICDIPC), 2016, pp. 37–41.

15. T. S. Chy and M. Anisur Rahaman, "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), 2019, pp. 455–459.

INNO **SPACE**
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross**ref**

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

📱 **9940 572 462**   🟢 **6381 907 438**   ✉ **ijircce@gmail.com**

Scan to save the contact details