A Survey on Sickle Cell Disease Detection and Analysis

Mrs. Kodur Srividya., Sanjana Jagannatha., Shrusti L., Shreya S Upadhya

Department of Computer Science and Engineering, K S Institute of Technology, Bengaluru, India

DOI: https://doi.org/10.51244/IJRSI.2026.1305000124

Received: 09 May 2026; Accepted: 14 May 2026; Published: 03 June 2026

ABSTRACT

With increasing population numbers worldwide, sickle cell disease (SCD) continues to pose a serious global health problem, especially in sub-Saharan Africa, where close to 240,000 babies are born every year with the disease. SCD refers to a genetic disorder that leads to the distortion of hemoglobin molecules and subsequent deformity of red blood cells into rigid, sickle-like forms. The irregular blood cells block blood vessels and degenerate prematurely, causing health problems such as anemia, pain crises, infections, and organ dysfunction. It is critical to diagnose the condition early and correctly to manage and control it effectively.

Advancements in artificial intelligence and image processing have facilitated the development of automatic detection systems for SCD. Deep learning methods have shown great promise in recognizing deformities in microscopic images of blood smears with high accuracy and speed. Automatic detection models can aid healthcare practitioners through shortened diagnostic duration, reduced error rates, and easy-to-use tests in resource-limited areas. This survey paper provides an extensive analysis of deep learning solutions for the detection of sickle cell disease.

Keywords: Sickle Cell Disease, Hemoglobin Gene Mutation, Diagnostic Techniques, Red Blood Cell (RBC) Analysis, Medical Image Analysis, Image Processing, Machine Learning, Deep Learning, Convolutional Neural Networks (CNN).

INTRODUCTION

Recent studies have indicated that the early detection of Sickle Cell Disease (SCD) is necessary for increasing the survival rate and minimizing mortality. Over 90 percent of patients diagnosed with SCD are newborn babies, while more than 50 percent of individuals affected by SCD are located in areas such as Nigeria, India, and the Democratic Republic of Congo (DRC). In the traditional approach to diagnosing diseases, sickling test and solubility test were utilized for detecting SCD, but these methods proved to be inefficient compared to the gold standard hemoglobin electrophoresis. Although hemoglobin electrophoresis offers great precision, the method is rarely available in areas that require the procedure the most, owing to resource constraints.

Through technological improvements, image processing techniques and machine learning algorithms have been used to increase classification performance through the extraction of morphological properties of RBCs. Due to their effectiveness, deep learning methods, including Convolutional Neural Network (CNN), have proven to be efficient in learning from data and images and offer better results in classification tasks. Nevertheless, most systems developed for diagnosing SCD are primarily concerned with detection and fail to provide analysis, such as evaluating severity and predicting risks of Vaso-occlusive.

BACKGROUND

Sickle cell anemia is an inherited disease characterized by abnormal haemoglobin which is the major carrier of oxygen in the blood. Red blood cells have a disc-like shape and are flexible allowing them to flow freely within the vessels. Sickle cell diseases cause red blood cells to assume a curved appearance known as the "sickle" shape because of changes in genes which affect haemoglobin. The formation of these sickled cells makes it difficult for them to bend and hence obstruct blood flow to other parts of the body.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue V May 2026

Sickle Cell Disease (SCD) was first documented in Western medicine in 1910 by James B. Herrick, who described abnormal, sickle-shaped red blood cells in a patient. However, the disease had already been recognized in African regions under different local names. In 1949, Linus Pauling identified SCD as the first molecular disease, linking it to abnormal haemoglobin. Later, in 1956, Vernon Ingram discovered the specific amino acid substitution in the β-globin chain responsible for the formation of haemoglobin S (HbS).

SCD results from a mutation in the β-globin (HBB) gene, encoding one subunit of HaA hemoglobin. The mutation leads to the replacement of glutamic acid by valine at position 6 of the β-chain of hemoglobin. Hemoglobin S is produced and forms rigid structures when there is a shortage of oxygen in the environment. Red blood cells change their shape and appear as sickle cells due to the polymerization of hemoglobin S in the absence of oxygen. This condition presents as an autosomal recessive disease since patients who inherit two genes with mutations from both parents suffer from SCD, while those carrying only one gene from one parent suffer from sickle cell trait (SCT).

The disease exists in several forms depending on genetic variations. The most severe type is Haemoglobin SS (HbSS), where both genes produce haemoglobin S. Haemoglobin SC (HbSC) involves one haemoglobin S gene and one haemoglobin C gene, typically resulting in milder symptoms. Sickle cell beta-thalassemia (HbS/β-thalassemia) occurs when one gene produces haemoglobin S and the other causes reduced haemoglobin production. Rare forms include combinations such as HbSD, HbSE, and HbSO.

Common symptoms of SCD include sudden episodes of severe pain due to blocked blood flow, chronic anaemia caused by rapid breakdown of red blood cells, fatigue, and shortness of breath. Patients are also prone to frequent infections due to spleen damage. Additional symptoms include delayed growth in children and vision problems caused by damage to blood vessels in the eyes.

SCD can lead to severe complications such as stroke, acute chest syndrome, organ damage, leg ulcers, priapism, and pregnancy-related complications. The obstruction of blood flow by sickled cells results in reduced oxygen supply to tissues, causing long-term damage and increased risk of early mortality, especially in regions with limited healthcare access.

Diagnosis is primarily performed through blood tests that detect abnormal hemoglobin. Hemoglobin electrophoresis remains the gold standard for confirming the presence and type of hemoglobin variants. Additional methods include complete blood count (CBC) and peripheral blood smear examination. Recent advancements have introduced computational approaches such as image processing, machine learning, and deep learning for automated detection.

Although there is no universal cure for SCD, several treatments help manage symptoms and reduce complications. Hydroxyurea is commonly used to reduce the frequency of pain crises by preventing sickling. Blood transfusions are used to treat anaemia and prevent complications such as stroke. Bone marrow transplantation offers a potential cure for some patients with a compatible donor. Emerging gene therapy approaches aim to correct the defective gene and enable the production of normal haemoglobin

LITERATURE REVIEW

Over time, several studies have been conducted for the precise and early detection of Sickle Cell Disease (SCD). Diagnostic approaches have evolved from traditional methods such as complete blood count (CBC), peripheral blood smear examination, and screening tests like sickling and solubility tests to advanced machine learning and deep learning techniques that enable automated detection and improved analytical accuracy. Some of the key approaches are discussed below.

Traditional Approach

Screening tests such as sickling tests and solubility tests are commonly used for preliminary diagnosis, whereas confirmatory techniques such as hemoglobin electrophoresis and High-Performance Liquid Chromatography (HPLC) play a crucial role in identifying hemoglobin variants. Hemoglobin electrophoresis is widely regarded

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue V May 2026

as a standard confirmatory method due to its ability to separate hemoglobin types based on their electrical charge and accurately identify variants such as HbA, HbS, HbF, and HbC.

Comparative studies have shown that when screening methods are evaluated against hemoglobin electrophoresis, they are less accurate. The sickling test demonstrates relatively higher sensitivity, whereas the solubility test may produce false-negative results, particularly in cases of sickle cell trait (SCT), thereby limiting their diagnostic reliability.

Image Processing Approach

There have been numerous image processing techniques explored to automate the detection of sickle cell disease (SCD) using microscopic blood smear images. These methods typically involve preprocessing steps such as grayscale conversion, noise filtering, and image enhancement, followed by segmentation and feature extraction to distinguish between normal and abnormal red blood cells. Techniques such as threshold segmentation combined with morphological operations and Support Vector Machine (SVM) classification have been shown to be fast and accurate. However, these approaches are often evaluated on small datasets and rely on predefined geometric and statistical features for classification, which may limit their performance in complex or noisy conditions. Additionally, they are sensitive to variations in image quality and threshold selection.

Similarly, shape-based detection methods using algorithms such as Circular Hough Transform (CHT) and Watershed Segmentation have been used for identifying and counting red blood cells. Among these, CHT has been found to be faster and more accurate for red blood cell detection, whereas Watershed Segmentation is more effective in separating overlapping cells but is comparatively slower. However, despite achieving reasonable accuracy and reliability, these methods often struggle with large datasets and are sensitive to noise, variations in cell morphology, and scalability issues. This highlights the lack of automation, robustness, and generalization capability.

Machine Learning Approach

Recently, there have been some machine learning techniques that have been implemented for the detection of sickle cell through images and clinical data. The use of Random Forest, Support Vector Machines (SVM), Naïve Bayes, and Logistic Regression is widespread in performing classification tasks. In terms of accuracy, Random Forest is considered the better technique compared to Naïve Bayes which, on the other hand, tends to perform comparatively less well. While these techniques offer good prospects in terms of success, they are dependent on the quality of features being extracted, and their process involves several stages which may pose difficulties during implementation.

Similarly, machine learning models have been employed using clinical data, particularly haematological parameters, for the detection of sickle cell disease. Algorithms such as Random Forest, XGBoost, SVM, and Logistic Regression have demonstrated high accuracy, indicating that SCD can be effectively predicted using clinical features alone. However, these approaches depend strongly on accurate laboratory measurements, and any inconsistencies in data collection can significantly impact model performance. Furthermore, the effectiveness of these models is highly dependent on data quality and feature distribution, which may limit their ability to generalize across varying datasets. Consequently, these methods often lack robustness and generalization in real-world applications.

Deep Learning Approach

In recent years, deep learning approaches have significantly improved the efficiency of sickle cell disease (SCD) detection. Models such as GoogLeNet and ResNet have achieved high accuracy, supported by the use of Grad-CAM for visualizing model decisions, thereby demonstrating the capability of deep learning to identify sickle cells from microscopic images. Furthermore, hybrid approaches have been explored, where a ResNet-50 model is combined with a Random Forest classifier to enhance classification performance, achieving improved accuracy compared to standalone deep learning models.

In addition to classification-based methods, segmentation techniques have also been widely investigated. Deformable U-Net models have been employed for red blood cell (RBC) segmentation, proving highly effective in handling irregular shapes and overlapping cells, and demonstrating greater robustness compared to traditional image processing and classification techniques. Moreover, lightweight architectures such as MobileNetV1 have been introduced to reduce computational complexity while maintaining reasonable accuracy, making them suitable for deployment in resource-constrained environments.

However, despite these advancements, several limitations remain. Many approaches rely on relatively small datasets, resulting in poor generalization to real-world scenarios. Hybrid models often depend on pretrained feature extraction followed by machine learning classifiers, leading to increased system complexity and computational cost. Additionally, segmentation-based methods such as U-Net are sensitive to noise and blur and require pixel-level annotated data, which is expensive and time-consuming to obtain. Although lightweight models improve efficiency, they are generally limited to binary classification tasks and lack deeper clinical insights, thereby restricting their applicability in comprehensive disease analysis.

CONCLUSION AND FUTURE SCOPE

Sickle cell disease (SCD) is a hereditary blood disorder caused by a mutation in the haemoglobin gene, which changes normal red blood cells into rigid, sickle-shaped cells. These abnormal cells can obstruct blood flow and lead to severe complications such as stroke, organ damage, anaemia, and severe pain. Since early diagnosis is essential for effective disease management, several diagnostic methods have been developed for accurate detection and monitoring of the disease.

Traditional diagnostic techniques such as sickling tests, blood smear analysis, and haemoglobin electrophoresis are widely used for SCD detection. However, these methods often require skilled professionals, laboratory infrastructure, and considerable processing time. Recent advancements in artificial intelligence and medical image analysis have encouraged the development of automated machine learning and deep learning-based approaches for sickle cell detection. Convolutional neural network (CNN) models, including lightweight architectures such as MobileNet, have shown promising performance in identifying sickle-shaped cells from blood smear images with improved efficiency and lower computational complexity.

This paper presents a survey of existing machine learning and deep learning approaches for sickle cell disease detection, focusing on widely used datasets, preprocessing techniques, feature extraction methods, and classification models. The survey highlights the potential of automated diagnostic systems in improving detection efficiency and supporting healthcare delivery in resource-constrained environments. In addition, it discusses current challenges and future research directions aimed at achieving greater clinical depth in diagnosis using advanced techniques and enabling the classification of different types and subcategories of sickle cell disease.

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our project guide, Mrs. Kodur Srividya, assistant professor, KSIT for their valuable guidance and support throughout the project. We also extend our heartfelt thanks to the esteemed professors of our college for their continuous encouragement and unwavering support.

REFERENCES

1. A. Author(s). (2026). Title of the article. PMC. https://pmc.ncbi.nlm.nih.gov/articles/PMC6858853/
2. A. Author(s). (2026). Title of the article. ScienceDirect. https://www.sciencedirect.com/science/article/pii/S3050565826000120
3. B. Sen, A. Ganesh, A. Bhan, S. Dixit and A. Goyal. (2021). Machine learning based Diagnosis and Classification of Sickle Cell Anemia in Human RBC. 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), pp. 753–758. https://doi.org/10.1109/ICICV50876.2021.9388610
4. D. C. Rees, T. N. Williams and M. T. Gladwin. (2015). Sickle-cell disease. Blood Cells, Molecules, and

Diseases, 55(1), 9–17. https://doi.org/10.1016/j.bcmd.2015.04.001
5. Goswami, N. G., Goswami, A., Sampathila, N., Bairy, M. G., Chadaga, K., & Belurkar, S. (2024). Detection of sickle cell disease using deep neural networks and explainable artificial intelligence. Journal of Intelligent Systems, 33(1), Article 20230179. https://doi.org/10.1515/jisys-2023-01
6. H. S, S. S. S and R. Arumuga Arun. (2024). A Computationally efficient CNN-based Deep Learning Technique for Sickle Cell Detection. 2024 International Conference on Electronic Systems and Intelligent Computing (ICESIC), pp. 18–22. https://doi.org/10.1109/ICESIC61777.2024.10846167
7. J. Gómez, C. M. Rosales and S. I. Montaño. (2026). Sickle Cell Anemia Prediction System in Machine Learning Based on Clinical Data. IEEE Access, 14, 46048–46062. https://doi.org/10.1109/ACCESS.2026.3674938
8. Johns Hopkins Medicine. (2026). Sickle Cell Disease. https://www.hopkinsmedicine.org/health/conditions-and-diseases/sickle-cell-disease
9. M. Abdulraheem Fadhel, A. J. Humaidi and S. R. Oleiwi. (2017). Image processing-based diagnosis of sickle cell anemia in erythrocytes. 2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT), pp. 203–207. https://doi.org/10.1109/NTICT.2017.7976124
10. M. Rees, T. Williams and M. Gladwin. (2010). Sickle Cell Disease: Current Challenges. Blood Reviews, 24(6), 247–256. https://doi.org/10.1016/j.blre.2010.07.001
11. M. Zhang, X. Li, M. Xu and Q. Li. (2020). Automated Semantic Segmentation of Red Blood Cells for Sickle Cell Disease. IEEE Journal of Biomedical and Health Informatics, 24(11), 3095–3102. https://doi.org/10.1109/JBHI.2020.3000484
12. National Heart, Lung, and Blood Institute. (2026). Sickle Cell Disease. https://www.nhlbi.nih.gov/health/sickle-cell-disease
13. Nigam, R., Sharda, B., & Varma, A. (2024). Comparative study of sickling test, solubility test, and hemoglobin electrophoresis in sickle cell anemia. MGM Journal of Medical Sciences, 11, 31–37. https://doi.org/10.4103/mgmj.mgmj_31_24
14. T. S. Chy and M. A. Rahaman. (2018). Automatic Sickle Cell Anemia Detection Using Image Processing Technique. 2018 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE), pp. 1–4. https://doi.org/10.1109/ICAEEE.2018.8642984
15. University College London. (2020). Digitized Thin Blood Films for Sickle Cell Disease Detection. UCL Research Data Repository. https://rdr.ucl.ac.uk/articles/dataset/Digitized Thin_Blood_Films_for_Sickle_Cell_Disease_Detection/12407567
16. V. Jain, A. K. Dubey and A. Jain. (2024). Hybrid Deep Learning and Machine Learning Approach for Sickle Cell Disease Diagnosis Using ResNet and Random Forest Classifier. 2024 13th International Conference on System Modeling & Advancement in Research Trends (SMART), pp. 606–609. https://doi.org/10.1109/SMART63812.2024.10882544
17. World Health Organization. (2026). Sickle-cell disease. https://www.who.int/news-room/fact-sheets/detail/sickle-cell-disease
18. Y. Kim. (2015). Laboratory diagnosis of hemoglobinopathies. Annals of Laboratory Medicine, 35(5), 409–417. https://doi.org/10.3343/alm.2015.35.5.409