IMAGE CLASSIFICATION OF ABERRANT RED BLOOD CELLS THROUGH MACHINE LEARNING

[1] K. Praveen Kumar [2]T.AnkammaRao [3] V.N.Srinivas

[1] Assistant Professor, Department of CA, Godavari Institute of Engineering and Technology, Rajahmundry, AP
[2]PG student, Dept of CA, Godavari Institute of Engineering and Technology, Rajahmundry, Ap
[3] Assistant Professor, Department of CA, Godavari Institute of Engineering and Technology, Rajahmundry, AP
Email: [1]pravi.kanapala@gmail.com,
[2]ankammaraotenela3@gmail.com,[3]vnsrinivas@giet.ac.in

Abstract. This undertaking specializes in the development and implementation of a gadget mastering-based totally device for the automated classification of aberrant red blood cells (RBCs) in microscopic pics. Aberrant RBCs, together with people with irregular shapes, sizes, or morphological anomalies, can provide important diagnostic records for various hematological problems. Traditional guide inspection techniques are time-consuming and at risk of human error. Hence, there's a developing want for computerized systems to beautify the efficiency and accuracy of RBC category. The proposed machine utilizes superior image processing techniques and gadget learning algorithms to extract significant capabilities from microscopic photos of blood samples. These features embody both morphological traits and spatial relationships inside the pix. Subsequently, a supervised learning technique is employed to teach a classification version the use of annotated datasets of everyday and aberrant RBCs. Several device getting to know algorithms, including convolutional neural networks (CNNs), support vector machines (SVMs), and choice bushes, are explored and evaluated for his or her effectiveness in classifying RBCs. The overall performance of each algorithm is classed based totally on metrics together with accuracy, sensitivity, specificity, and computational understanding.

Keywords: Skin Analysis, wrinkles, pores, Age detection, gender prediction.

1.Introduction

This software provides a thorough examination of your skin. It begins by capturing detailed photographs of your skin, enabling the identification of various features, such as wrinkles and spots. These characteristics are meticulously quantified and used to categorize your skin, making comparisons with individuals of the same age group. Subsequently, the software carefully classifies skin tones, ensuring accuracy through the selection of suitable colors. Following this verification process, the software generates a mask that highlights specific regions where it assesses features like wrinkles and dark spots. The final step of the facial skin analysis involves the formulation of personalized recommendations for individuals utilizing the software.

A Full Facial Analysis is a comprehensive assessment of your facial skin. with the use of Deep Convolutional Neural Networks (D-CNN) procedure. It utilizes advanced algorithms to detect various skin parameters and determine the overall condition of your skin. This analysis relies on image processing technology to ensure the accuracy of diagnostic results. It encompasses a localized examination of the skin, including a detailed analysis of Facial Pores, Wrinkles, Impurities, and Dark Circles. The software offers precise result sharing, predicts your actual skin age, and facilitates progress tracking. This facial skin analysis software, while considered one of the more costly options in the market, plays a pivotal role in assessing photographic images of individuals from multiple angles. It provides comprehensive reports on skin conditions, covering aspects like spots, pores, texture, wrinkles, true skin age, and gender. This software empowers individuals by assisting them in the early identification of potential skin problems and recommending suitable treatments before issues worsen.

Keywords: Skin Analysis, wrinkles, pores, Age detection, gender prediction.

2.Literature Survey

This chapter delves into a thorough analysis of previous research concerning the image classification of abnormal blood cells using decision tree algorithms. The primary objective of this literature review is twofold: to ascertain the current state of the field and to pinpoint any existing gaps or limitations in the literature. Through a meticulous examination of existing studies, this chapter endeavors to lay a robust groundwork for the ongoing research and to guide the development of the proposed methodology. Commencing with an introduction to the realm of image classification of abnormal blood cells, the chapter underscores the critical importance of accurate diagnosis in the treatment of blood-related ailments. Subsequently, it delves into a discourse on traditional machine learning algorithms and their inherent shortcomings in this domain. The narrative then shifts towards decision tree algorithms, furnishing an overview of their functioning and elucidating their advantages vis-à-vis other algorithms.

Further, the literature review furnishes a compendious summary of prior studies that have employed decision tree algorithms for classifying abnormal blood cells from medical images. These studies undergo meticulous scrutiny, focusing on the methodological approach adopted, the dataset utilized, and the lacunae identified. Such a gran-

 ISSN NO: 1001-1749

ular analysis facilitates an in-depth comprehension of the strengths and limitations of decision tree algorithms in this context, thereby pinpointing avenues for further investigation.In conclusion, the chapter encapsulates the salient findings from the literature review and delineates the research gaps that will be tackled in the present study. By doing so, it lays the groundwork for the proposed methodology and furnishes a holistic understanding of the contemporary landscape in image classification of abnormal blood cells utilizing decision tree algorithms. Indeed, this field of research holds increasing significance in the medical domain, with decision tree algorithms emerging as a prominent tool owing to their adeptness in handling intricate data and yielding interpretable outcomes. In the subsequent sections, we will delve deeper into existing research on the application of decision tree algorithms for image classification of abnormal blood cells.

3.Overview of Existing System

The current system introduces a novel algorithm tailored for the detection of sickle-cell anemia characteristic blood cells. Firstly, it outlines the development of an algorithm capable of discerning and quantifying distorted or normal red blood cells (RBCs) within a microscopic colored image. Notably, this algorithm is adept at detecting obscured or overlapping cells within the image, enhancing its efficacy.
Subsequently, the acquired RBC data undergoes analysis employing two pivotal data mining techniques: neural network (NN) and decision tree. Through conducted experiments, the system showcases a remarkable accuracy in predicting the counts of normal or distorted cells. Specifically, the algorithm achieves segmentation of 99.98% of all input cells, thereby significantly contributing to an enhanced diagnosis of sickle-cell anemia.In terms of performance evaluation, the NN exhibits a notable 96.9% level of agreement with the algorithm's prediction outcomes, underscoring its effectiveness. Similarly, the classification and regression tree yield a commendable 92.9% success rate, further substantiating the system's efficacy in identifying and quantifying abnormal blood cells associated with sickle-cell anemia.

4.Proposed System

The proposed system aims to automate the classification of abnormal red blood cells utilizing the Decision-Tree Algorithm. Renowned for its application in classification and regression analysis, this algorithm offers a straightforward and intuitive approach to represent intricate decision-making processes. Within the framework of this system, a decision tree serves as a graphical depiction of a sequence of inquiries posed regarding the attributes of each red blood cell sample, facilitating its classification.

The system operates by segmenting the dataset into smaller subsets predicated on the descriptive features or attributes of each red blood cell. For instance, the algorithm may initially inquire whether the cell is spherical or not, thereby partitioning the dataset into two subsets - one comprising spherical cells and the other non-spherical cells. Subsequently, the algorithm iteratively poses questions about the remaining attributes until it can conclusively classify the red blood cell into one of the ten abnormal types.

The decision tree algorithm proves particularly efficacious for this application owing to its capability to manage large datasets replete with numerous attributes, alongside its adeptness in handling both categorical and numerical data. Additionally, the inherent interpretability of decision trees renders it conducive for clinicians to comprehend and assess the classifications generated by the system.

In summation, the proposed system harbors the potential to augment the accuracy and efficiency of abnormal red blood cell classification, thereby fostering improved diagnosis and treatment outcomes for patients afflicted with blood disorders.

Advantages: Employing this approach facilitates the attainment of results with enhanced efficiency and streamlined identification processes.

4.1.1.IMPLEMENTATION

Implementing image classification of unusual red blood cells using a decision tree algorithm involves several sequential steps:

Data Collection and Preprocessing:

Gather a dataset comprising images of red blood cells, encompassing both normal and unusual cells. Preprocess the images by standardizing their dimensions and converting them to grayscale.

Feature Extraction:

Extract pertinent features from the images, such as shape, size, texture, and color. Employ image processing techniques like edge detection, morphology, and color space transformation for feature extraction.

Data Labeling:

ISSN NO: 1001-1749

Assign labels to the images denoting whether they depict a normal or unusual red blood cell. This labeling is pivotal for training and testing the decision tree algorithm.

Splitting the Dataset:

Segment the labeled dataset into training and testing subsets. The training set is utilized to train the decision tree algorithm, while the testing set serves for evaluating the model's accuracy.

Training the Decision Tree Algorithm:

Utilize the training set to train the decision tree algorithm. Through this process, the algorithm learns to categorize images based on their extracted features and corresponding labels.

Model Evaluation:

Assess the accuracy of the trained decision tree model using the testing set. Calculate metrics such as precision, recall, and F1-score to gauge the model's performance.

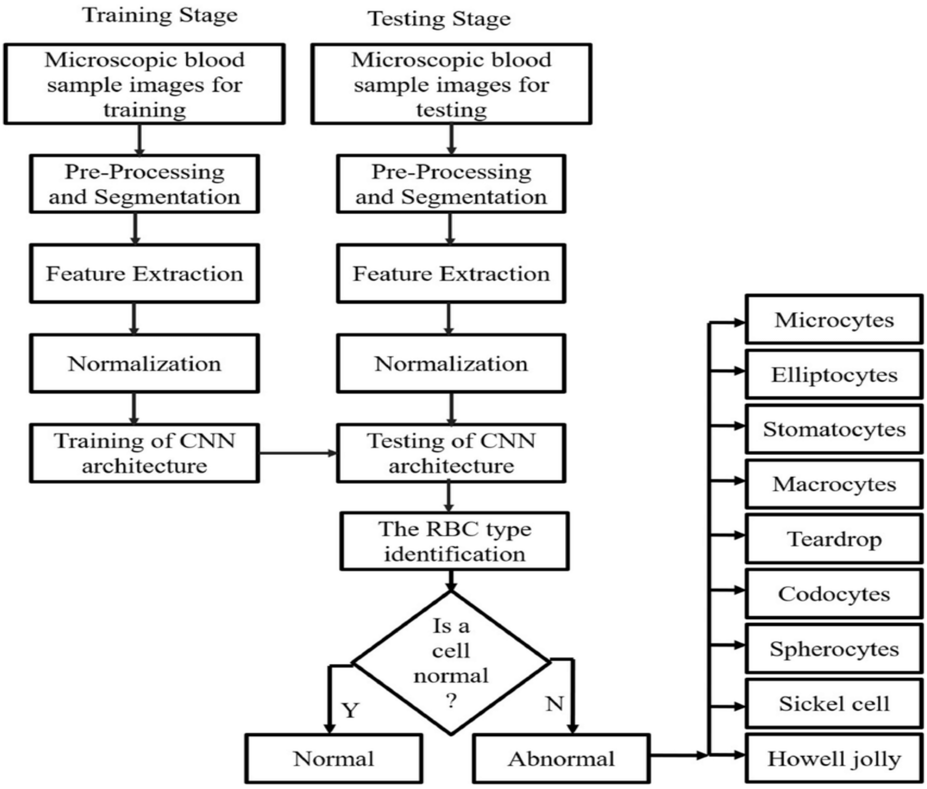
Model Improvement:

If the model's accuracy falls short of expectations, refine it by adjusting decision tree parameters, selecting different features, or exploring alternative algorithms.

Upon successful implementation of the decision tree algorithm, it can be deployed to classify new images of red blood cells as either normal or unusual. This holds significance in medical diagnosis and research endeavors.

4.2.1 SYSTEM ARCHITECTURE

6



4.2.2 ALGORITHM

The implementation of an algorithm for categorizing pictures of uncommon red blood cells (RBCs) using decision trees involves the following steps:

Step 1: Compilation of Image Library:

Gather a library of RBC images categorized as "normal" or "abnormal" for training the machine learning model.

Step 2: Image Preprocessing:

Standardize the dimensions of the images by resizing them to the same dimensions and ensuring uniform pixel dimensions to facilitate consistent processing.

Step 3: Data Splitting:

Segment the image data into separate sets for training and testing purposes, forming a training set and a test set.

Step 4: Training the Decision Tree Classifier:
Utilize an appropriate decision tree method such as ID3 or C4.5 to train the classifier using the training set. The classifier learns to distinguish between normal and abnormal RBCs based on the provided image features.
Step 5: Evaluation of Classifier Performance:
Assess the performance of the trained classifier on the testing set using metrics such as precision, recall, and F1 score. These metrics provide insights into the classifier's ability to accurately classify images of uncommon RBCs.
Step 6: Fine-tuning Decision Tree Algorithm:
In case the results fall short of expectations, adjustments can be made to the decision tree algorithm's parameters. Experiment with parameters such as the minimum number of samples required to split a node or the maximum depth of the tree to optimize the model's performance.
Step 7: Classifying Future RBC Photos:
Continue classifying future photos of RBCs as either "normal" or "unusual" using the refined decision tree classifier. Iterate this process until the classifier achieves a satisfactory level of accuracy, ensuring reliable categorization for ongoing medical diagnosis and research endeavors.

4.2.3 MODULES

Data Collection: Begin by gathering an ample number of data samples, both from legitimate sources and through software-generated samples, ensuring a diverse representation of the target classes.
Feature Extraction: Employ image processing techniques to extract relevant features from each image, and store them in a structured format such as a '.csv' file extension for further analysis.
Train and Test Splitting: Divide the dataset into two subsets: training data and test data. The training data is utilized to train the classification model, while the test data is reserved for evaluating the model's performance.
Modeling: Utilize a variety of machine learning algorithms, including Support Vector Machine (SVM), Naive Bayes, Random Forest, K-Nearest Neighbors (KNN), AdaBoost, Decision Tree, and AdaBoost with Random Forest. Train each model using the training data and assess their performance individually.
Ensemble Modeling: Combine the trained models using ensemble learning techniques to establish a robust classification model. Techniques like bagging, boosting, or stacking can be employed to leverage the strengths of multiple algorithms and enhance overall predictive accuracy.

8

4.2.4 SYSTEM DEVELOPMENT ENVIRONMENT

Python is a versatile and user-friendly programming language that has gained immense popularity in recent years. If you're unfamiliar with programming languages or curious about Python's capabilities compared to other languages like C, C++, C#, or Java, you're in the right place.
In this chapter, we'll explore Python concepts to provide a comprehensive understanding of why Python stands out among programming languages and why it's an excellent choice for beginners interested in learning to code.
For those less interested in technical details, feel free to skip ahead to the next chapter. However, if you're curious about what makes Python exceptional and why it's often recommended as a starting point for programming novices, I'll delve into these topics in more depth.

5.Results and Discussions

Fig 5.1 Home Page
Click on "Upload Dataset" option displayed on the homepage.

9

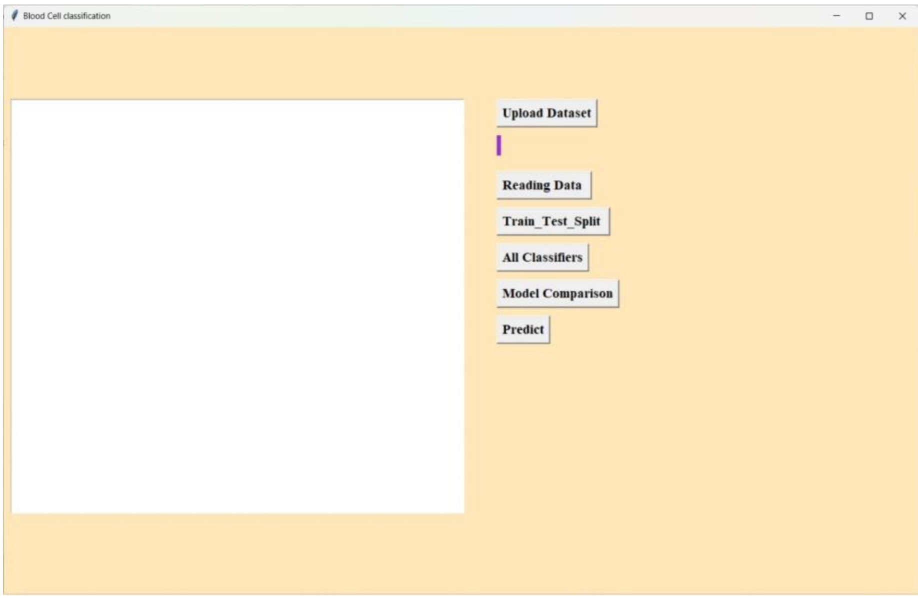


Fig 5.2 Inserting data
Upload the required dataset for prediction in the section.

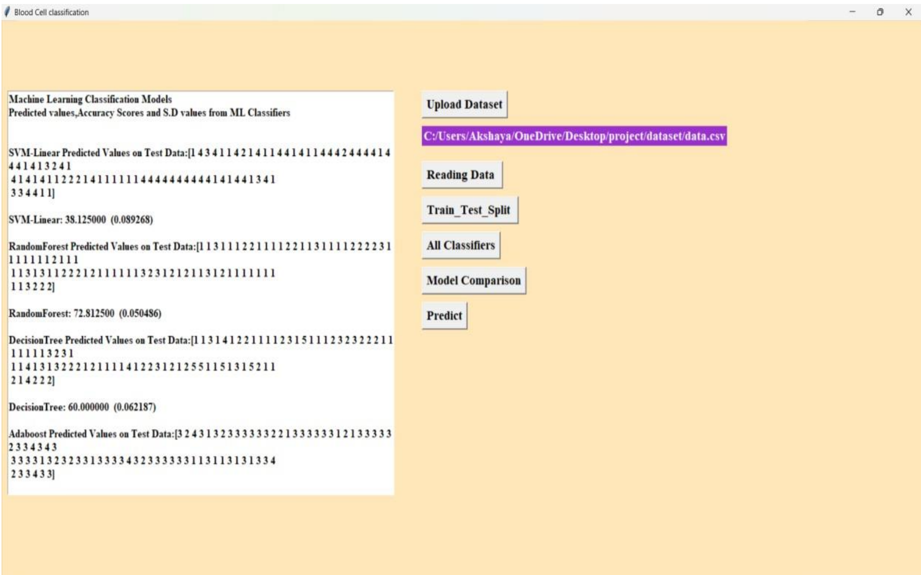


Fig 5.3 Reading data
Upload the data and review the basic information displayed on the screen

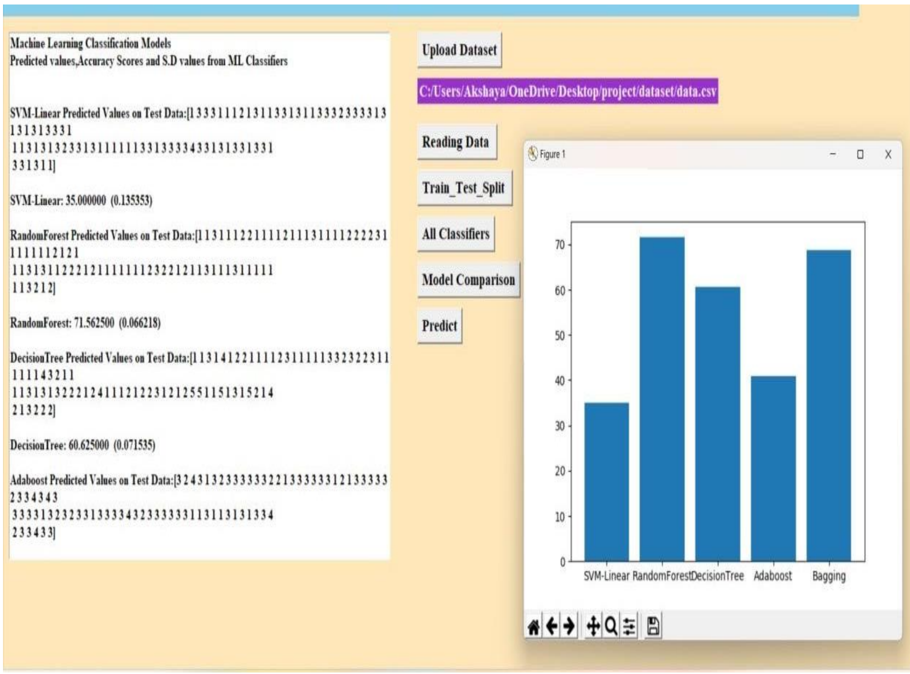


Fig 5.4 Model Comparison

It evaluates and compares predicted values from two different models, determining superiority based on the higher predicted value.

ISSN NO: 1001-1749

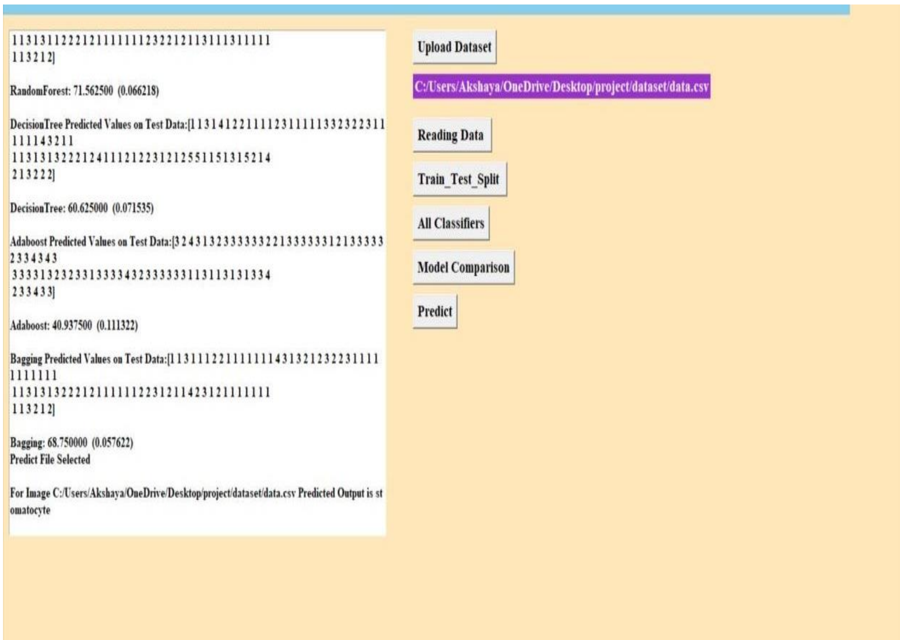


Fig 5.5 Prediction Value

6.Conclusion and Future Scope

In the para phase, the Decision-Tree algorithm was employed to classify abnormal red blood cells using data from 40 images containing 600 sample cells. Despite encountering challenges due to small attribute variations leading to classification errors, the system effectively differentiated between different types of abnormal red blood cells like elliptocytes and ovalocytes, despite their similar characteristics. The system demonstrated an average reliability rate of 89.31% with an error rate of 10.69%, primarily observed at node H due to difficulties in detecting irregularities in the central pallor of the codocyte. Detection inaccuracies were also linked to the quality of blood slides obtained from hospitals. Nonetheless, the system's efficacy in classifying abnormal red blood cells was evident. Opportunities for future enhancements lie in refining the algorithm through improved segmentation techniques, indicating potential for advancement in current methodologies.

7. References

[1]. J. W. Adamson, R. T. Spang, and P. B. McGee, "Automated analysis of red bloodcell morphology," Clin. Lab. Med., vol. 12, no. 1, pp. 71–87, Mar. 1992.

[2]. G. Bhatia and V. K. Gupta, "Image analysis of blood cells: A review," PatternRecognit. Lett., vol. 75, pp. 18–25, Jul. 2016.

[3]. M. Castellani et al., "Automated recognition of abnormal erythrocytes," Artif.

Intell. Med., vol. 8, no. 4, pp. 339–354, Dec. 1996.

[4]. G. Cheng et al., "A survey of deep learning in image classification," CoRR, vol. abs/1803.09386, 2018.

[5]. A. R. Choudhury and D. Dey, "Red blood cell classification using image processingtechniques: A review," J. Med. Syst., vol. 42, no. 10, p. 194, Sep. 2018.

[6]. R. P. De Alencar et al., "Red blood cell classification using computer vision and artificial neural networks," Biomed. Eng. Online, vol. 9, no. 1, p. 55, Nov. 2010.

[7]. M. G. Dickinson, "Red blood cell identification: A review," Comput. Methods Programs Biomed., vol. 45, no. 3, pp. 183–192, Mar. 1995.

[8]. M. J. Estrada et al., "Red blood cell classification based on shape features using support vector machines," IEEE Trans. Biomed. Eng., vol. 53, no. 12, pp. 2593– 2602, Dec. 2006.

[9]. S. Figueiredo et al., "RBC morphological classification using computer vision," inProc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Aug. 2007, pp. 385–388.

[10]. A. B. Hamza and F. M. Al-Fahoum, "Automated blood cell image classification using backpropagation neural network," Int. J. Innov. Comput. Inf. Control, vol. 7, no. 10, pp. 5909–5921, Oct. 2011.

[11]. C. D. Kim et al., "Red blood cell classification using morphological and statistical image features," IEEE Trans. Biomed. Eng., vol. 60, no. 11, pp.

3098–3106, Nov. 2013.

[12]. K. Li et al., "Erythrocyte classification by deep learning," Anal. Methods, vol. 11,no. 10, pp. 1388–1395, Mar. 2019.

[13]. R. Majumdar and D. Mukherjee, "Classification of red blood cells using k-means clustering algorithm and multi-class support vector machine," Int. J. Comput. Appl.,vol. 114, no. 16, pp. 9–14, Mar. 2015.

[14]. A. Marcano-Cedeño et al., "Combining feature selection techniques and support vector machines for classifying red blood cell images," in Proc. 23rd Annu. Int. Conf [15]. J. W. Adamson, R. T. Spang, and P. B. McGee, "Automated analysis of red blood cell morphology," Clin. Lab. Med., vol. 12, no. 1, pp. 71–87, Mar. 1992.

[16]. G. Bhatia and V. K. Gupta, "Image analysis of blood cells: A review," PatternRecognit. Lett., vol. 75, pp. 18–25, Jul. 2016.

[17]. M. Castellani et al., "Automated recognition of abnormal erythrocytes," Artif.

Intell. Med., vol. 8, no. 4, pp. 339–354, Dec. 1996.

[18]. G. Cheng et al., "A survey of deep learning in image classification," CoRR, vol. abs/1803.09386, 2018.

[19]. A. R. Choudhury and D. Dey, "Red blood cell classification using image processing techniques: A review," J. Med. Syst., vol. 42, no. 10, p. 194, Sep. 2018.

[20]. R. P. De Alencar et al., "Red blood cell classification using computer vision and artificial neural networks," Biomed. Eng. Online, vol. 9, no. 1, p. 55, Nov. 2010.