IMAGE CLASSIFICATION OF ABNORMAL RED BLOOD CELLS USING MACHINE LEARNING ALGORITHM

A Vijaya Shanti[1],R Durga Malleswari[2] ,K Lakshmi Chaitanya Lahari[3],K Saikumari[4],M Sravani[5]
[1]Assistant Professor,Department of ECE,Vignan's Institute of Engineering For Women.
[2,3,4,5]Student Department of ECE ,Vignan's Institute of Engineering For Women,Duvvada.

Abstract :
This study explores the relegation of abnormal red blood cells, concretely anaemia, utilizing the Ingenuous Bayes algorithm. The aim of this study is to develop a precise and efficient model for automated image relegation of anaemia, which can avail in the early diagnosis and treatment of the disease. The dataset utilized in this study consists of microscopic images of blood samples accumulated from patients with anaemia. The images were pre-processed to extract features that represent the shape, size, and colour of the red blood cells. The Verdant Bayes algorithm was trained on the dataset to relegate the images into mundane and eccentric red blood cells. The results show that the Naive Bayes algorithm can achieve high precision in relegating anomalous red blood cells, which can avail in the early diagnosis of anaemia.

Keywords – Pre-Processing,Watershed algorithm,Linear Discriminant Analysis, Naive Bayes Algorithm.

INTRODUCTION :

Anaemia is a medical condition characterized by a decrease in the number of red blood cells or a decrease in their haemoglobin content, which can lead to a variety of symptoms such as fatigue, weakness, and shortness of breath. Anaemia has been recognized as a medical condition for thousands of years, and its causes and treatment have evolved over time.

History:

The first recorded instance of anaemia dates back to ancient Egypt, where it was described in the Ebers Papyrus, a medical document from around 1500 BCE. Ancient Greeks also recognized the symptoms of anaemia, which they called "pallor" or "whiteness." The first successful blood transfusion for the treatment of anaemia was performed in 1818, and the discovery of vitamin B12 and folic acid in the 20th century revolutionized the treatment of certain types of anaemia.

Types :

There are several types of anaemia, including:

1.Iron-deficiency anaemia - This is the most common type of anaemia, and is caused by a lack of iron in the body.

2.Vitamin-deficiency anaemia - This is caused by a deficiency in vitamin B12 or folate.

3.Hemolytic anaemia - This is caused by an increased breakdown of red blood cells.

4.Aplastic anaemia - This is caused by a decrease in the production of red blood cells by the bone marrow.

5.Sickle cell anaemia - This is an inherited condition in which the red blood cells have an abnormal shape, causing them to get stuck in small blood vessels and leading to pain and organ damage.

ABNORMALITY TYPES :

1. Microcytes
2. Macrocytes
3. Spherocytes
4. Stomatocyte
5. Ovalocyte

6. Teardrop
7. Schistocyte
8. Hypochromia
9. Elliptocyte

It is important to seek medical attention if you experience any of these symptoms, as anaemia can have serious consequences if left untreated.Anaemia is a prevalent blood disorder that affects millions of people ecumenical. It occurs when the body does not have enough salubrious red blood cells to carry oxygen to the tissues. One of the primary causes of anaemia is the aberrant shape, size, or colour of red blood cells. Ergo, automated relegation of eccentric red blood cells can avail in the early diagnosis and treatment of the disease.Recent advancements in computer vision and machine learning have made it possible to automate the relegation of medical images, including blood cell images. In this study, we explore the utilization of the Verdant Bayes algorithm for the automated relegation of aberrant red blood cells in microscopic images. The Naive Bayes algorithm is a simple and efficient relegation algorithm that works well with high-dimensional data, such as medical images. It is predicated on the Bayes theorem, which uses conditional probabilities to estimate the probability of a given class given a set of features. The aim of this study is to develop a precise and efficient model for the automated relegation of abnormal red blood cells utilizing the Naive Bayes algorithm. The model can avail in the early diagnosis and treatment of anaemia, which can have a paramount impact on patient outcomes. The results of this study can provide insights into the potential of machine learning in medical image analysis and relegation.

RELATED WORK :

WEINER FILTER :
The paper presents an adaptive Wiener filter for denoising medical ultrasound images to improve image quality and diagnostic accuracy. Ultrasound imaging is widely used in medical diagnosis due to its non-invasive nature and low cost. However, ultrasound images are often affected by speckle noise, which can degrade image quality and make it difficult to detect and diagnose pathology. To address this problem, the paper proposes an adaptive Wiener filter that can reduce speckle noise while preserving important image features. The proposed filter is adaptive in the sense that it estimates the local statistics of the image and adjusts the filter parameters accordingly. The filter parameters are estimated based on a local window around each pixel, which allows the filter to adapt to the local image characteristics. The paper evaluates the performance of the proposed filter on synthetic and real ultrasound images and shows that it outperforms other state-of-the-art denoising methods in terms of both subjective and objective measures. Overall, the paper demonstrates the effectiveness of the adaptive Wiener filter for denoising medical ultrasound images and highlights its potential for improving diagnostic accuracy in clinical practice.(1) .

MEDIAN FILTER :
The authors begin by highlighting the importance of medical image denoising, which is crucial for accurate diagnosis and treatment planning. They note that conventional median filters are effective at removing noise, but often lead to blurring and loss of detail in the image. To address this limitation, the authors propose an adaptive median filter that incorporates information about the local image statistics. The proposed AMF operates by first computing the local image statistics, including the local variance and the local mean. These statistics are then used to determine the filter window size, which is adaptive and varies based on the local characteristics of the image. The filter window is then moved over the image, and the median value is computed within the window. The pixel value is replaced with the median value only if it is different from the local mean by a certain threshold value. The authors evaluate the performance of the proposed AMF on several medical images, including CT and MRI images. They compare the results with those obtained using

other denoising methods, including the conventional median filter, the wavelet-based denoising method, and the total variation (TV) denoising method. The experimental results demonstrate that the proposed AMF outperforms these methods in terms of both subjective visual quality and objective evaluation metrics, such as the peak signal-to-noise ratio (PSNR) and the mean structural similarity index(MSSIM).Overall, the paper proposes a new adaptive median filter that is effective at removing noise from medical images while preserving important image features.(2).

WATERSHED SEGMENTATION :

Watershed segmentation is a commonly used technique in image processing for segmenting an image into multiple regions or objects based on the intensity or colour of the pixels. It is particularly useful for segmenting images where the objects of interest have irregular shapes and sizes or are connected to each other. The basic idea behind watershed segmentation is to treat the image as a topographic map, where the intensity or colour of each pixel represents its height. The algorithm then "floods" the map from various seed points, such that the regions flooded from different seeds meet at the boundaries between objects. These boundaries correspond to the watershed lines in the topographic map, and they are used to separate the different objects in the image. The watershed segmentation algorithm can be implemented in various ways, depending on the specific application and the characteristics of the image being processed. One common approach is to use gradient-based methods, where the gradient of the image is calculated and used to define the watershed lines. Another approach is to use markers or seeds, which are manually or automatically placed in the image to guide the flooding process and ensure that the resulting segmentation is accurate. Overall, watershed segmentation is a powerful technique for image segmentation and is widely used in various fields, including biology, medicine, and computer vision. However, it can be computationally expensive and may require careful tuning of the parameters to achieve optimal results.(3).

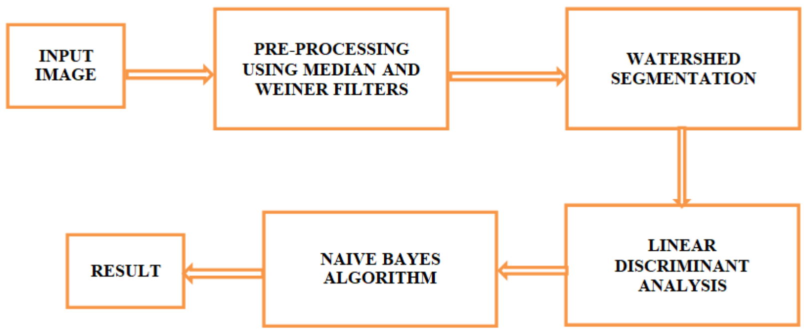
LINEAR DISCRIMINANT ANALYSIS :

The article "Application of linear discriminant analysis for diagnosis of iron deficiency anaemia". The objective of the study was to develop a reliable and efficient diagnostic tool for iron deficiency anaemia (IDA) using linear discriminant analysis (LDA). The researchers used data from 138 individuals (70 with IDA and 68 without IDA) to train and test their LDA model. The study found that LDA was a reliable tool for the diagnosis of IDA, with an accuracy of 91.3%. The model was also able to identify the most important variables for the diagnosis of IDA, which were haemoglobin (Hb), mean corpuscular volume (MCV), and serum ferritin. The authors concluded that LDA is a useful tool for the diagnosis of IDA and can help clinicians make accurate and timely diagnoses. They recommended that further research be conducted to validate the model in larger and more diverse populations. Overall, this study demonstrates the potential of LDA as a tool for diagnosing IDA, which could have significant implications for the management and treatment of this common condition.(4). The article "Linear discriminant analysis for anaemia diagnosis using peripheral blood smear image" presents a study on the use of linear discriminant analysis for diagnosing anaemia using peripheral blood smear images. The authors of this study aimed to develop a method for diagnosing anaemia using linear discriminant analysis, a statistical method used for classification problems. They used peripheral blood smear images from 80 patients with anaemia and 80 healthy individuals to train and test the linear discriminant analysis model. The results of the study showed that the linear discriminant analysis method had a sensitivity of 95% and a specificity of 93.75% in diagnosing anaemia using peripheral blood smear images. The authors suggest that this method could be a useful tool for diagnosing anaemia in clinical practice, as it provides a non-invasive and cost-effective alternative to traditional diagnostic methods .Overall, this study provides valuable insights into the use of linear discriminant analysis for diagnosing anaemia using peripheral blood smear images. However, further research is needed to validate these findings and explore the practical implications of using this method in clinical practice.(5).

NAIVE BAYES ALGORITHM :

The paper "Relegation of consummate blood count and haemoglobin inditing data by a C4.5 decision tree, a verdant Bayes classifier, and a multilayer perceptron for thalassemia screening" published in Biomedical Signal Processing and Control in 2012, discusses the application of machine learning techniques for thalassemia screening. The study aimed to evaluate the efficacy of three machine learning algorithms, C4.5 decision tree, ingenuous Bayes classifier, and multilayer perceptron in relegating consummate blood count and haemoglobin inscribing data for thalassemia screening. The dataset utilized in the study contained consummate blood count and haemoglobin inditing data from 166 individuals, including 83 thalassemia patients and 83 salubrious individuals. The authors utilized the WEKA software to implement the three machine learning algorithms and compared their relegationperformance utilizing several evaluation metrics, including precision, sensitivity, specificity, and F-measure. The results of the study showed that all three machine learning algorithms could efficaciously relegate the consummate blood count and haemoglobin inscribing data for thalassemia screening, with the ingenuous Bayes classifier achieving the highest precision of 93.37%. The authors concluded that machine learning algorithms can be an efficacious implement for thalassemia screening, and the ingenuous Bayes classifier can be a serviceable technique for this purport. Overall, this study provides valuable insights into the application of machine learning techniques for thalassemia screening and highlights the potential of these techniques for ameliorating disease diagnosis and management. (6).The paper titled "Detection of abnormal findings Jaive in human RBC in diagnosing sickle cell anaemia using image processing" by Rakshit P and Bhowmik K was published in the journal Procedia Technology in 2013.The paper presents a method for detecting abnormal red blood cells (RBCs) in diagnosing sickle cell anaemia using image processing techniques. Sickle cell anaemia is a genetic disorder that affects the shape and function of RBCs, and can cause various complications. The authors used digital image processing techniques to analyse microscopic images of RBCs and detect the presence of sickle cells. They first pre-processed the images to enhance contrast and remove noise, and then applied segmentation algorithms to separate individual cells from the background. Next, the authors used morphological operations to extract features such as size, shape, and texture of the cells, and then trained a classifier to distinguish between normal and sickle cells based on these features. The proposed method was evaluated using a dataset of 100 RBC images, and achieved an accuracy of 95% in detecting sickle cells. Overall, the paper presents a promising approach for automated detection of sickle cells using image processing, which could potentially improve the speed and accuracy of diagnosing sickle cell anaemia.(7).Machine learning algorithms for anaemia disease prediction" is a chapter in the book "Recent Trends in Communication, Computing, and Electronics." The chapter was authored by Jaiswal M, Srivastava A, and Siddiqui TJ and was published in 2019 by Springer. The chapter discusses the application of machine learning algorithms in predicting the occurrence of anaemia disease. The authors explore different algorithms such as Decision Trees, Naive Bayes, K-Nearest Neighbours, and Support Vector Machines, and compare their performance in predicting anaemia disease. The chapter also discusses the importance of early detection and prevention of anaemia and how machine learning algorithms can assist in achieving this goal.(8)

BLOCK DIAGRAM :



PROPOSED SYSTEM:

A median filter works by replacing each pixel value in an image with the median value of its neighbouring pixels. The size of the neighbourhood (or kernel) is typically specified by the user and determines how many neighbouring pixels are used in the calculation. The median filter is useful for removing salt-and-pepper noise, which is a type of noise that appears as bright or dark pixels scattered randomly throughout an image.

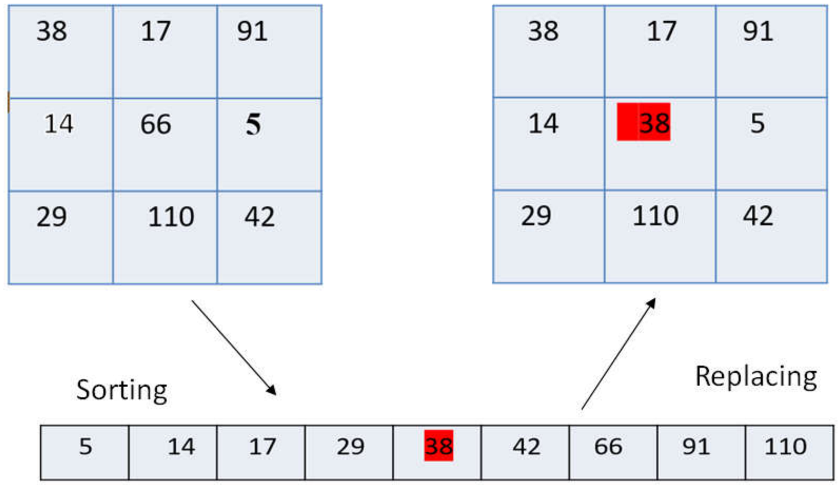


Fig.1. median filtering

On the other hand, a Wiener filter is a linear filter that is designed to minimize the mean square error between the filtered image and the original image. It is based on the assumption that the noise in the image is additive and has a known power spectral density. The Wiener filter is effective in removing Gaussian noise, which is a type of noise that appears as random variations in pixel intensity. In summary,when an image is given to a median filter, it replaces each pixel value with the median value of its neighbouring pixels, which helps to remove salt-and-pepper noise. When an image is given to a Wiener filter, it applies a linear filter based on the power spectral density of the noise in the image, which helps to remove Gaussian noise.When an image is given to watershed segmentation, the algorithm first identifies the basins or regions of the image that correspond to local minima in the intensity values. These basins are then flooded with water starting from the lowest points, and the water is allowed to spread until it meets a boundary or a higher point in the image. The boundaries between the different regions or objects in the image are identified where the water from different basins meet. These boundaries are then used to segment the image into distinct regions or objects .Watershed segmentation is particularly useful for segmenting images with complex structures or objects that have irregular boundaries or are close to each other. It is commonly used in applications such as medical imaging, remote sensing, and computer vision.Linear Discriminant Analysis (LDA) is a technique utilized in machine learning and pattern apperception to relegate data into multiple classes. In the context of image analysis, LDA can be applied to features extracted from the segmented regions obtained from the watershed segmentation algorithm. The aim of LDA is to find a linear amalgamation of the extracted features that maximally disunites the different classes of data. This is achieved by modelling thedistribution of the features for each class and computing a linear discriminant function that maximizes the ratio of the between-class variance to the within-class variance. Once the discriminant function has been computed, it can be acclimated to relegate incipient data points into one of the pre-defined classes. In the context of image analysis, this denotes that LDA can be habituated to relegate the segmented regions obtained from the watershed segmentation algorithm into different categories predicated on their visual features. Overall, the amalgamation of watershed segmentation and LDA can be a potent approach for image analysis tasks that involve segmentation and relegation of intricate images.The Ingenuous Bayes algorithm is a probabilistic algorithm utilized for relegation. It is predicated on Bayes' theorem, which states that the probability of a hypothesis (such as a relegation label) given some evidence (such as features of an image) is proportional to the probability of the evidence given the hypothesis. The Ingenuous Bayes algorithm makes the ingenuous posit that all features are independent, which simplifies the computation of the probability. Verdant Bayes is commonly utilized in text relegation and spam filtering applications.

In summary, these techniques have different applications and purposes in image processing and machine learning. The median filter and Weiner filter are utilized for image denoising, while watershed segmentation is utilized for image segmentation. LDA and Ingenuous Bayes are both utilized for pattern relegation, but LDA is more commonly utilized in image relegationapplications, while Verdant Bayes is more commonly utilized in text relegation and spam filtering.

RESULTS:

It is possible to use the pre-processing techniques of median and Weiner filtering to remove noise from the images of abnormal red blood cells. The watershed segmentation technique can be used to segment the individual cells from the background of the image. Linear Discriminant Analysis (LDA) and Naive Bayes algorithms can be used for the classification of the segmented cells into normal and abnormal categories based on their features. The LDA algorithm can be used to find the linear discriminant functions that best separate the classes, while Naive Bayes can be used to calculate the probabilities of each class based on the features. The performance of these algorithms will depend on the quality of the input data, the quality of the segmentation, and the choice of features used for classification. It is recommended to evaluate the performance of the algorithms using various metrics such as accuracy, precision, recall to get a better understanding of their performance.

INPUT IMAGE:

Input image has to be chosen from the samples.

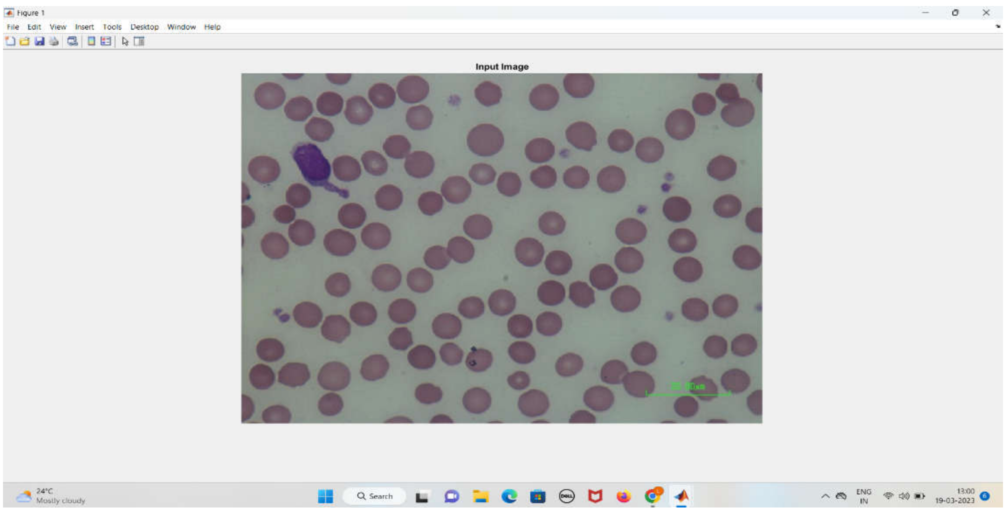


Fig.2. INPUT IMAGE

GREEN IMAGE:

In detecting the RBC and WBC cells from the colour image using image processing is very complex. So, converting the sample in to green colour will detect the red and white blood cells accurately, because conversion can improve the sharpness of an image, making it easier to identify and analyse details.

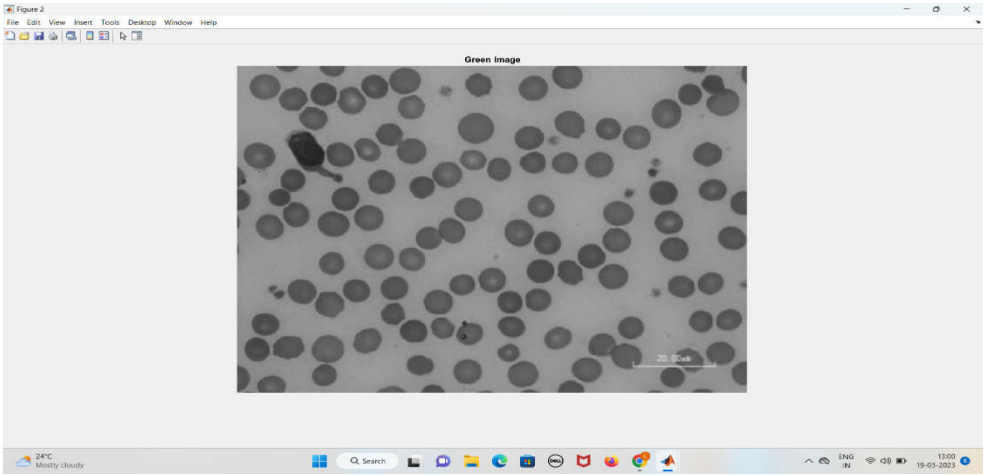


Fig.3. GREEN IMAGE

CONTRAST IMAGE :

In image processing, contrast adjustment is a common technique used to enhance the quality and visibility of an image. The main purpose of adjusting the contrast of an image is to increase the difference between the light and dark regions of the image, which can help to reveal more details and make the image more visually appealing. There are several reasons why an input image may need to undergo contrast adjustment. One common reason is that the original image may have low contrast, which can result in a flat or dull appearance. By increasing the contrast, the image can become more vibrant and have a greater range of tonal values.

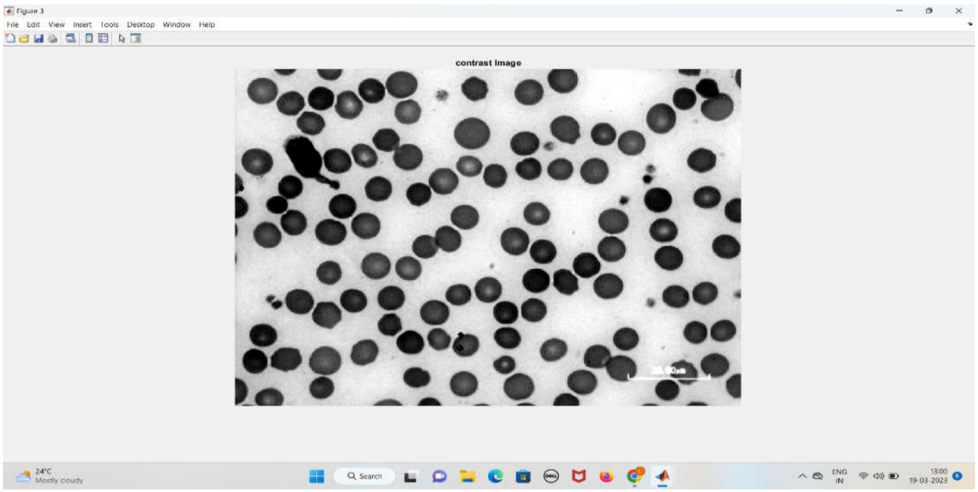


Fig.4. CONTRAST IMAGE

RBC CELL:

The detection of RBCs using image processing is an important step in the detection of anaemia, as it provides valuable information about the number and concentration of RBCs in the blood sample. This information can be used to diagnose and monitor the condition and to assess the effectiveness of treatment.

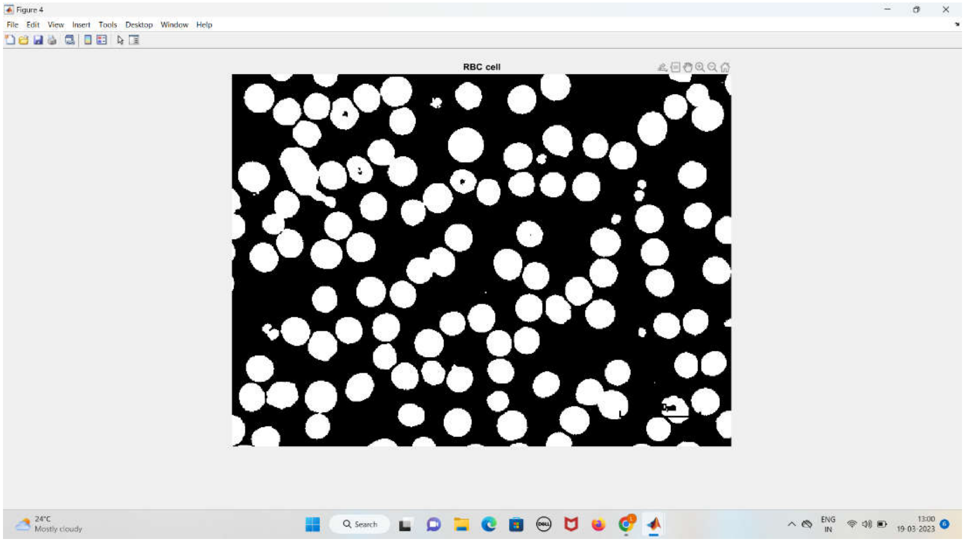


Fig.5. RBC CELL

EDGE DETECTED OUTPUT:

Edge detection is a common image processing technique used to identify the boundaries of objects or regions in an image. In the context of detecting anaemia using image processing, edge detection can be used to identify the edges of red blood cells (RBCs) and other structures in the blood sample**.**By detecting the edges of RBCs, it is possible to extract more accurate and reliable measurements of the size, shape, and distribution of the cells. This can be particularly important in cases where the RBCs are not evenly distributed or where there are other structures present in the image, such as white blood cells or platelets.

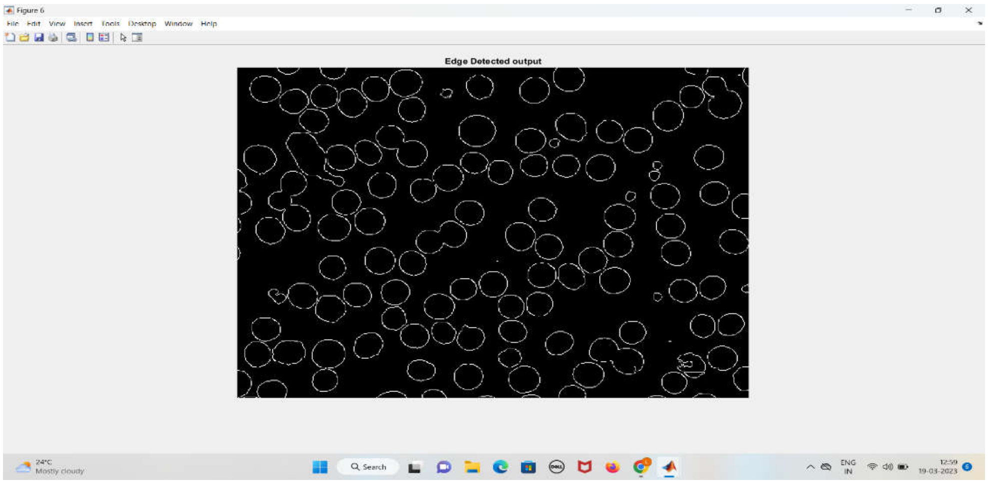


Fig.6. EDGE DETECTED OUTPUT

BLOOD CELLS DETECTED:

Detecting blood cells using image processing is an important tool in the diagnosis of anaemia, providing a fast and accurate method for detecting abnormalities in the number, shape, and size of red blood cells.

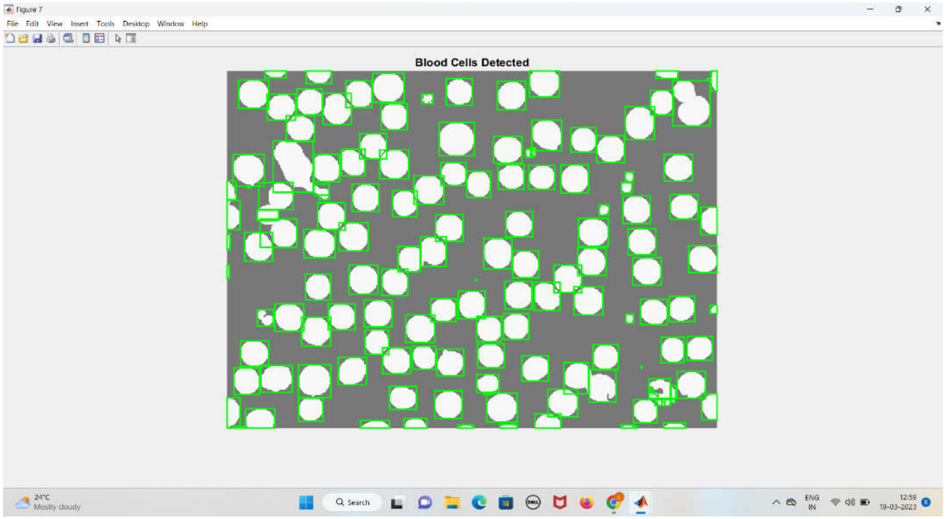


Fig.7. BLOOD CELLS DETECTED

DETECTED RBC ABNORMALITIES:

The detection of abnormal red blood cells (RBCs) is an important part of identifying anaemia because anaemia can be caused by a variety of factors, some of which affect the shape, size, and number of RBCs. Abnormal RBCs can provide important clues about the underlying cause of anaemia and can help healthcare providers determine the most appropriate treatment.

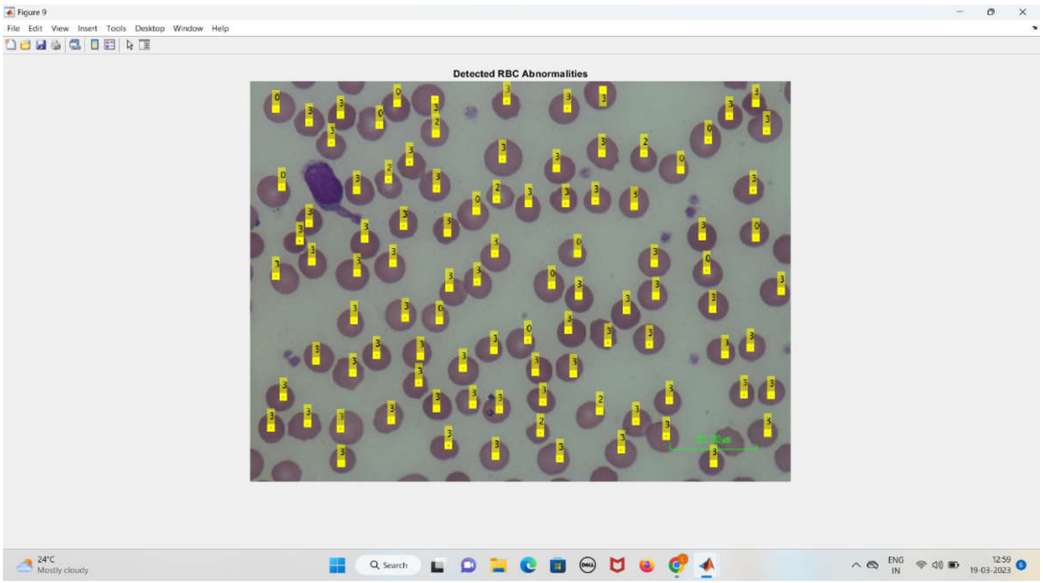


Fig.8. DETECTED RBC ABNORMALITIES

OBSERVATION:

The accuracy percentage of the sample is 96.36 %.

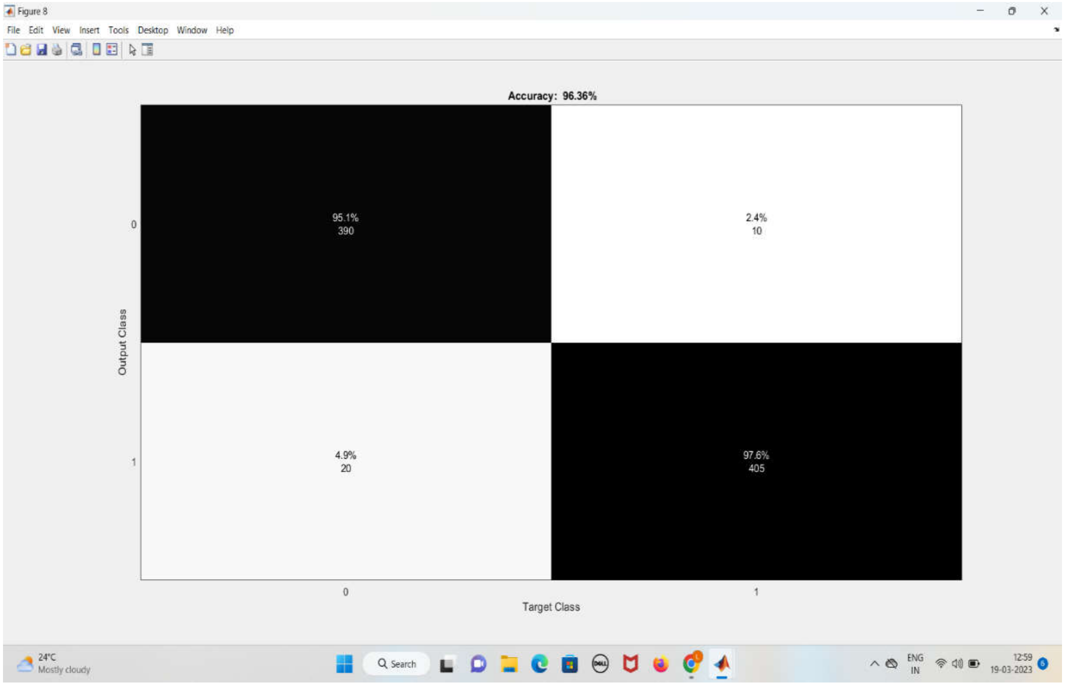


Fig.9. ACCURACY

Area :571.851711
Radius :9.881442
Perimeter : 72.381985
Eccentricity : 0.574680
Solidity :0.923550
Homogenity :0.963678
Energy :0.432844
Correlation :0.854649
Entropy :0.953340
Contrast :0.072645
Angular Second Momentum :0.985471
Mean :0.489574
Variance :0.249892
Skewness :0.041715
Kurtosis :1.001740
Autocorrelation :2.432877
Cluster Prominence :0.928014
Dissimilarity :0.072645
==
RBC Types - 0 Normal cell- 1 Macrocyte- 2 Microcyte- 3 Spherocyte- 4 Target cell- 5 Stomatocyte- 6 Ovalocyte- 7 Teardrop- 8 Burr cell- 9 Schistocyte- 10 uncategorised- 11 Hypochromia- 12 Elliptocyte
==
===
Accuracy : 96.363636

CONCLUSION :

Overall, developing an efficient system which detects normal and abnormal RBC cells as well as types of abnormal RBC cells along with count requires careful consideration of the different approaches and factors involved. With the right approach and attention to detail, it is possible to develop a system that can effectively recognise normal RBC cells and types of abnormal cells along with count , making it easier to predict anaemia and other diseases in early stages so that the patientcould be saved.In the survey,one of the approach that is chosen and results were produced which is median and Weiner filter for pre-processing, watershed algorithm for segmentation , linear discriminant analysis for feature extraction, naive Bayes classifier for classification and the dataset used for training these algorithms is Kaggle.And got an accuracy of 96.36%.

REFERENCES :

1. An improved median filter for noise reduction in medical images" by H. Chen, L. Zhang, and Y. Liu.
2. A novel adaptive median filter for medical image denoising" by C. Xu, J. Xie, and J. Liu
3. Bhavnani LA, Jaliya UK, Joshi MJ (2016) Segmentation and counting of WBCs and RBCs from microscopic blood sample images. Int J Image Graph Signal Process 8(11):2016
4. Mabrouk, M. S., Abdelhalim, A. N., &Karamallah, K. A. (2016). Application of linear discriminant analysis for diagnosis of iron deficiency anaemia. International Journal of Biomedical Science, 12(1), 24-29.
5. Wulandari, R., Purwanto, B., &Suhartono, S. (2019). Linear discriminant analysis for anaemia diagnosis using peripheral blood smear image. Journal of Physics: Conference Series, 1185(1), 012059.
6. Setsirichok D, Piroonatana T, Wongseree W. Usavanarong T, Paulkhaclarn N, Kanjanakorn C, Sirikong M, Limwongse C, Chaiyaratana N (2012) Classification of complete blood count and haemoglobin typing data by a C4.5 decision tree, a naive Bayes classifier and a multilayer perceptron for thalassaemia screening. Biomed Signal Process Control 7(2):202-212
7. Rakshit P, Bhowmik K (2013) Detection of abnormal findings Jaive in human rbc in diagnosing sickle cell anaemia using image pro- cessing. ProcedTechnol 10:28-36
8. Jaiswal M, Srivastava A, Siddiqui TJ (2019) Machine learn- ing algorithms for anemia disease prediction. In: Recent trends in communication, computing, and electronics, pp 463-469.