A review of microscopic analysis of blood cells for disease detection with AI perspective

Nilkanth Mukund Deshpande[1], **Dr. Shilpa Shailesh Gite**[2], **and Dr. Rajanikanth Aluvalu**[3]

[1]**Department of Electronics and Telecommunication Symbiosis Institute of Technology, Lavale, Pune, Symbiosis International (Deemed University), Pune-412115, and Dr. Vithalrao Vikhe Patil College of Engineering, Ahmednagar**
[2]**Department of Computer Science, Symbiosis Institute of Technology, Lavale, Pune Symbiosis International (Deemed University), Pune-412115 and Faculty, Symbiosis Center for Applied Artificial Intelligence(SCAAI), Symbiosis International(Deemed University), Pune, India**
[3]**Department of CSE, Vardhaman College of Engineering, Hyderabad**

Corresponding author:
Dr. Shilpa Shailesh Gite[2]

Email address: shilpa.gite@sitpune.edu.in

ABSTRACT

Background Any contamination in the human body can prompt changes in blood cell morphology and various parameters of cells. The minuscule images of blood cells are examined for recognizing the contamination inside the body with an expectation of maladies and variations from the norm. Appropriate segmentation of these cells makes the detection of a disease progressively exact and vigorous. Microscopic blood cell analysis is a critical activity in the pathological examination. It highlights the investigation of appropriate malady after exact location followed by order of abnormalities, which assumes an essential job in analyzing various disorders, treatment arranging, and assessment of results of treatment.
Methodology A survey of different areas where microscopic imaging of blood cells is used for disease detection is done in this paper. Research papers from this area are obtained from a popular search engine, google scholar. The articles are searched considering the basics of blood, such as its composition followed by staining of blood, which is the most essential and mandatory before microscopic analysis. Various methods for classification, segmentation of blood cells are reviewed. Microscopic analysis using image processing, computer vision, and machine learning is the main focus of the research and review. Methodologies employed by various researchers for blood cells analysis in terms of these mentioned algorithms is the critical point of review considered in the study.
Results Different methodologies used for microscopic analysis of blood cells are analyzed and are compared according to various performance measures. From the extensive review, the conclusion is made.
Conclusion Researchers have employed various machine learning and deep learning algorithms for the segmentation of blood cell components and disease detection considering microscopic analysis. There is a scope of improvement in terms of various performance evaluation parameters. Researchers can use different bio-inspired optimization algorithms for performance improvement. Explainable AI can analyze AI system features and make the system more trusted and commercially suitable.

INTRODUCTION

Blood, the most integral part of the body, is constituted of white blood cells (WBC), red blood cells (RBC), platelets, and plasma. This can be further categorized as; cells and platelets are about 45 % of human blood, whereas the remaining 55% is filled by plasma (the yellow fluid in the blood). These components and their physical properties like size, shape, color, and count in the whole blood change due to ingress of any foreign object or micro-organism can lead to any sort of infections.

47 There are different pathological procedures for the detection of diseases. In most cases, microscopic
48 imaging plays a vital role in predicting and detecting abnormalities and occurrences of diseases within the
49 body. Typically health of any person is judged by analyzing different features of blood cells and their
50 counts.

51

Why the study needed

53 Previously manual methods of blood cells analysis were used by pathologists. This might cause error in
54 disease prediction since manual methods are dependent on experience and skills of pathologists. Hence,
55 it is proposed that an automated system of image processing be developed using different algorithms.
56 A simplified, automated and cost effective system is required for detection of diseases. Thus the above
57 components explained are analyzed for knowing health indication of human being and thereby detecting
58 abnormalities, if any. Though many researchers contributed in the study, there is a need to explore the
59 research in many perspectives.

60 1. Segmentation of different blood components is still having some shortcomings, such as overlapping
61 cells during the staining.

62 2. There are different parasitic components in blood cells those need to be identified. So that an existence
63 of a particular malady could be highlighted.

64 3. There are many challenging diseases like leukemia that have many sub-types depending upon the cell
65 morphology. To detect the exact type of leukemia is still challenging.

66 4. In medical imaging, the use of artificial intelligence will have a critical issue that it is used as a
67 black box. Hence it could not be considered full proof and trusted at all times. The technique known as
68 explainable artificial intelligence is the need of study in relation to these analysis concepts.

Who it is intended for

70 It is always a critical and crucial job for diagnosing diseases in the medical field Since these decisions are
71 further related to a patient's life. To provide significant contributions in the current diagnostic system is to
72 be intended in many ways. This area is to be studied by a variety of disciplinary and inter-disciplinary
73 researchers. Following are details that will show to whom the study is intended for:

Academic Researchers

75 Artificial intelligence, machine learning, and deep learning are prime research areas in academics. A
76 disease detection system's performance utilizing microscopic imaging could be improved by applying
77 these algorithms to the current system.

Doctors

79 For diagnosis of diseases, doctors rely on the analysis of blood in many cases, in addition to the symptoms
80 of a patient. Blood cell analysis proves to be the most accurate method of diagnosis in the medical field in
81 most disorders. This study provides diagnostic assistance to doctors for further treatment of the patients.

Pathologists

83 Blood cells analysis and diagnosis is the leading job of a pathologist. He is responsible for statistical
84 analysis of blood, urine, etc. In some instances, the morphology of blood is important, which is analyzed by
85 microscopic methods. The predictions are dependent upon the experience and skill set of the pathologist
86 in some critical diagnostic conditions. The automated and sophisticated methods of diagnosis via
87 microscopic analysis will prove an assisted diagnostic system for them.

Commercial Use

89 Pathological analysis is either equipment-based, chemical-based, or morphology-based. Currently,
90 equipment-based systems are costlier, and there is a need to develop a cost-effective automated system for
91 the same. Morphology-based systems can be studied and employed for commercial use that will prove to
92 be cost-effective.

Bioinformations

94 A bioinformatician is a person with research in biology, medicine, and health-related studies. These fields
95 are studied with information technology. It involves the collection and interpretation of data covering a
96 range of fields, including genetics or pharmaceutics.

⁹⁷ *Haematologics*

⁹⁸ Hematology is the science that deals with the study of blood and its related disorders. There are some
⁹⁹ methods related to blood disorders that contribute a fair amount of suggestive diagnosis in hematology.

¹⁰⁰ *machine learning experts*

¹⁰¹ A machine and deep learning are the branches that prove to be the future of technology in the medical
¹⁰² field. There are different methods of machine and deep learning those could improve the decisions in
¹⁰³ medical imaging. This article provides the guidelines and challenges towards the utilization of machine
¹⁰⁴ learning in macroscopic imaging.

¹⁰⁵ *Technicians in laboratories*

¹⁰⁶ The use of AI in laboratories can guide the technician with less experience.

SURVEY METHODOLOGY

¹⁰⁸ Different approaches are used for searching the articles.

¹⁰⁹ 1. Articles are searched through keywords on one of the popular platforms, google scholar. The popular
¹¹⁰ keyword considered are white blood cell, red blood cell, disease, machine learning, deep learning, image
¹¹¹ processing, and explainability

¹¹² 2. The search is refined by re-arranging the keywords to make the search article specific in-lined with the
¹¹³ area. Papers are considered only from the English language.

¹¹⁴ 3. After getting a large number of articles, their abstract is read to finalize them for the reviewing process.

¹¹⁵ 4. Some papers are finalized for the review; those proved to be significantly contributed to the research
¹¹⁶ subject.

¹¹⁷ 5. Cross-references are also searched by scrutinizing some papers having good research contributions.
¹¹⁸ The following figure 1 shows the co-occurrences of different keywords, considering 02 keywords per
¹¹⁹ article as a threshold. The co-occurrence is analyzed by VOSviewer 1.65.
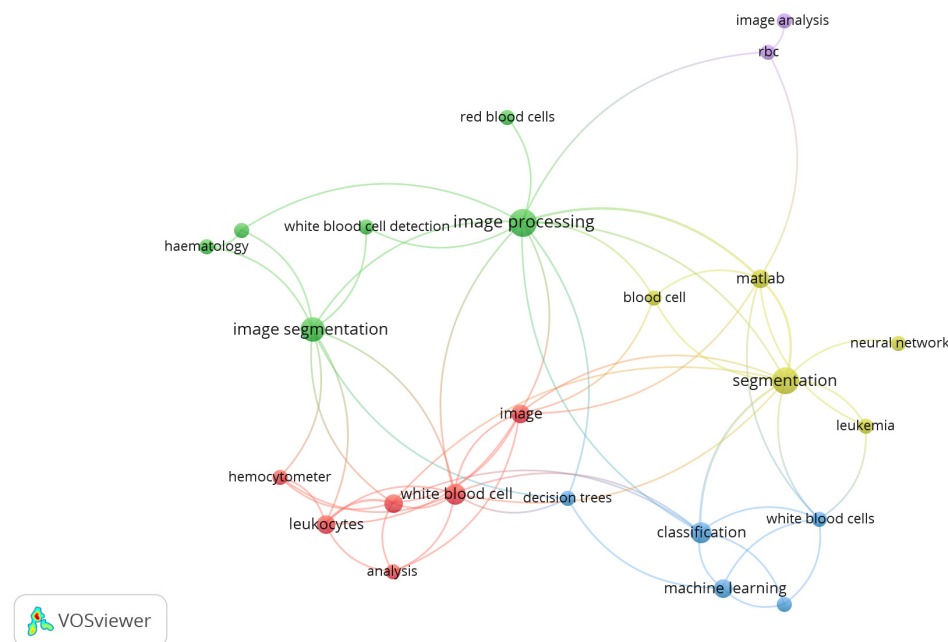
Figure 1. Co-occurrence of keyword
Source: VOSviewer 1.65

¹²⁰

REVIEW OVERVIEW

This overview is divided into different sections. The first section includes some terminologies related to blood analysis- blood composition and staining process. The second section comprises the microscopic imaging applied to various diseases and their analysis. This section is followed by the generalized methodology for the detection of disease by utilizing image processing and computer vision algorithms. The methods and algorithms adopted by different researchers are discussed in this section. A note on publicly available databases is presented further.

Some basic terminology related to blood
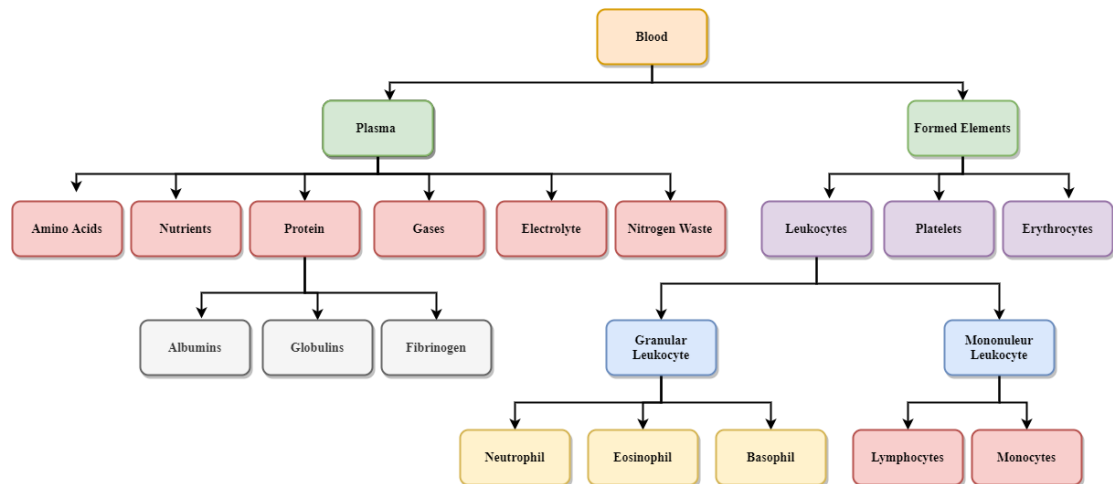
Figure 2. Components of Blood

Composition of blood

Figure 2 shows the details of different blood components. Blood is made up of the following elements-erythrocytes, known as red blood cells (RBC), leukocytes, known as white blood cells (WBC), and platelets. These are combinedly present within the plasma membrane. Leukocytes are further classified into two subcategories called granulocytes which consist of neutrophils, eosinophil and basophils, and agranulocytes, consisting of lymphocytes and monocytes. Blood plasma is a mixture of proteins, enzymes, nutrients, wastes, hormones, and gases. Platelets are small fragments of bone marrow cells. The primary function of red blood cells is to carry oxygen from the lungs to different body organs. Carbon dioxide is carried back to the lungs, which will be exhaled afterward. RBC count is the measure of different diseases in the human body. A low RBC count means anemia and high means polycythemia. White blood cells protect the body against infection. The different components of blood are identified to know about the health of a human being. Microscopic images of blood smear are analyzed for different disease detection.

Staining of Blood Smear For analysis of microscopic blood images, the blood film needs to be prepared. Glass slide is used for making the blood film. For examination and analysis of this film under the microscope, staining is required. Preparation of blood film requires a slide, a tube, and a blood spreader. Generally, weldge method is used for this purpose. On a base slide, a drop of blood is placed. A spreader slide is moved over this blood drop backward to touch the blood to get the blood spread over the slide uniformly. To get perfection and accuracy in the smear, the spreader slide should be inclined at an angle of 30 degrees to 45 degrees to the blood base slide. The prepared blood smear is dried using an air dryer, and then staining is performed. The dried smear is fixed by absolute methanol or ethyl alcohol. Afterward, it is stained using any of the staining methods – rewmanosky stain, leishmon stain, may-grawald giema or wright-giemsa stain, which differs in the liquid used for staining purpose. These stained slides are then used for analysis under the microscope.(26)(42)(60). Generally, laboratories are used to their respective homebrew technics, and therefore, as peripheral blood differential is aphenotypical method, technicians and doctors but also machine learning tools may have problems translating their experience for other laboratories.
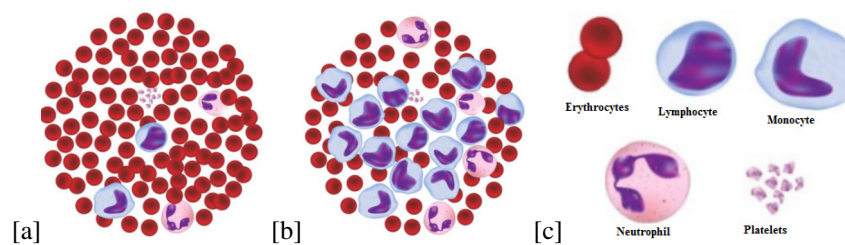
Figure 3. a)Normal Cells, b)Leukemia Cells, c)Different Components

Different areas where microscopic blood analysis is used

Blood smear images are observed under a good quality microscope and are investigated to identify any contamination inside the body(46)(8). Several abnormalities could be recognized from these microscopic images. Though there are different laboratory analysis techniques of blood examination available, image processing and computer vision could play a sound and satisfactory role in identifying maladies, preferably it analyzes the morphological abnormality of different components of blood(49)(34). The following are various areas where image processing and computer vision could be utilized for blood cell analysis(63)(9)(Lavanya et al.)(61).

Blood Cell Counts

RBC and WBC counts are characteristics of a patient's well-being. It is seen that in a large portion of the cases, the absolute blood cell count is influenced because of an infection within the body. Typically machines are present to count the blood cells and other components in the blood. Nevertheless, when required to get certain particular kinds of observations and abnormalities, there is a need for microscopic analysis. Also, the counting of RBC and WBC is possible by automated computer vision methods considering the cell morphologies. The blood smear is formed after staining that outcomes the film of blood. These films are observed under the microscope, and the photographs of these images are used for counting. The cells are preprocessed after the microscopic photograph and then segmented to get the required region of interest for counting. Image processing and computer vision methods are utilized for counting purposes. Isolated blood cells are counted via automated algorithms rather than manual, which enhances accuracy(1)(15)(37).

Detection of viral diseases such as Malaria, dengue, chikunguniya, hepatitis

The decrease in RBC and platelets are observed during viral infections. Moreover, the infections' parasites are also identified to realize the viral ailment like intestinal sickness, dengue, chikungunya, or hepatitis. Pathologist distinguishes these contaminations by microscopic blood cell analysis. This process of identification of different parasites is done by automated techniques involving computer vision, image processing, and machine learning algorithms(46).

Leukemia Detection

Harmful reasons, such as leukemia, seriously influence the body's blood-forming tissues and lymphatic framework. In leukemia, the white blood cells created by bone marrow are anomalous. Figure 3 shows the distinction between normal and leukemia cells. It has two significant subtypes, acute leukemia, and chronic leukemia. Leukemia can further be classified into other following types, namely, acute lymphocytic (ALL), acute myelogenous (AML), chronic lymphocytic (CLL), and chronic myelogenous (CML). Recognition of these malignant growth cells is done manually by microscopic image analysis and requires a competent pathologist. An improved automated system of leukemia detection is designed based on image processing and machine learning techniques, which ends up being proficient when contrasted with manual detection. (41)(17)(56)(54)(48)(29)(40).

Anemia and Sickle Cell Detection

Abatement in hemoglobin or absence of solid RBC prompts anemia. It can cause contamination of viral sicknesses and issues identified with relaxing. Anemia detection is mostly done by identifying sickle cells in the blood. These sickle cells have a typical crescent moon shape. These cells are recognized and categorized as ordinary cells and sickle cells via automated algorithms involving computer vision and machine learning(11)(22)(10)(28).

197 Figure 4 indicates the different blood cells' components through microscopic examination and Wright
198 stained smear of normal blood. For certain neurological disorders diagnoses such as Alzheimer's and
199 Parkinson's diseases, there are no exact criteria related to clinical ways. For improvements in these
200 kinds of diagnoses emerging metabolomics serves as a powerful technique. This includes a study of
201 novel biomarkers and biochemical pathways(64) There is one more challenge to identify proteins from
202 the complex biological network that interact with each other as well as with the cell's environment.
203 Pseudo-Guided Multi-Objective Genetic Algorithm (PGMOGA) proposed that reconstitutes pathways
204 by assigning orientation weighted network edges(27). A gene encoder presented that incorporates two-
205 stage feature selection with an unsupervised type for the classification of cancer samples(7) There is
206 a requirement of finding the correct DNA sequence to get the desired information about the genetic
207 makeup of an organism. A hybrid approach presented utilized Restarting and Recentering Genetic
208 Algorithm (RRGA) with integrated PALS(24). For working on the different datasets, there is a need to
209 get a set of visualization metrics. These are used to quantify visualization techniques—an approach of
210 visualization metrics based on effectiveness, expressiveness, readability, and interactivity. Evolutionalry
211 algorithm (EA) is used here as a case study. This methodology can also be utilized further for other
212 visualization techniques(25). There is also a requirement for the extraction of information from larger
213 datasets. A popular Frequent Itemsets (FIs) mining is a task to find itemsets in a transactional database.
214 The graph-based approach is used for the representation of a complete transactional database(23).

215 **Paragraph**   Out of these different diseases, leukemia is one of the most dangerous in its later stages(19).
216 It develops blasts cells in the blood, which later affect the generation of normal white blood cells. As
217 the number of these blast cells increases, the body gets shortened of healthy cells, leading to frequent
218 infections. Different types of leukemia define the way to treat it. So it is always necessary that the type of
219 leukemia be detected with great accuracy. The morphological differences in the normal and leukemia
220 cells are shown in figure 3.

221 **Generalized Methodology**
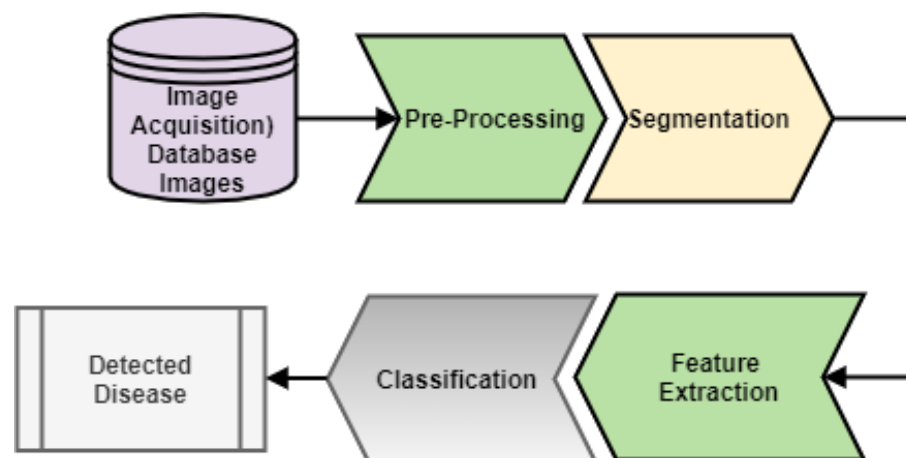
Figure 4. Generalized Methodology

222 A generalized methodology for microscopic blood cell analysis is shown in figure 4. It consists of
223 different stages like image acquisition, image segmentation, feature extraction, and disease detection. The
224 blood sample is taken from the patient by a trained pathologist. After that, a slide is prepared to form a
225 blood smear. The same slide is observed under the excellent quality microscope that will give an image.
226 This image is taken either by camera directly or from an adaptor connected to a microscope. This image
227 is considered for further analysis. Acquired images may have some unwanted regions and overlapping
228 of different components of blood. This image is enhanced by applying a suitable image enhancement
229 technique. So that a good quality image is now available for the analysis. After pre-processing, separation
230 of different blood components is done, including separation of RBC, WBC, plasma, and platelets. By
231 considering the generalized characteristics of blood components, segmentation is done. This process
232 will separate the region of interest for further classification. RBC, WBC and, other components are

233 further classified into their respective sub-classes. This will help specify a particular sub-class image
234 for extracting features for analysis of blood cells, and depending upon the analysis, detection of disease
235 is done. After the segmentation, different features are extracted by considering different components
236 of blood. Features include size, shape, color, count of different blood components like WBC and RBC
237 counts. Analysis of these features will further detect the disease or count the cells. Depending upon
238 different features extracted, the decision about the disease could be taken. To make the decisions, different
239 classifiers could be designed.

Image Pre-Processing

241 Following are the different methods used for pre-processing. Self dual multi-scale morphological
242 toggle (SMTT) block(12), wiener filter(43), median filtering (22)(14)(39)(53)(58), Gaussian filtering
243 (Yildirim and Çinar), gray-scale transformation(43)(22)(14)(39)(53) (13)(16)(58) which has 3 types viz
244 Linear, Logarithmic and Power–law, histogram stretching(22)(14)(39)(53), green color component from
245 the RGB image(41), morphological operations(22), edge detection (16)

Image Segmentation

247 The following are the different segmentation methods employed by the researchers.
248 Watershed transform(12) (11) granulometric analysis and mathematical morphology (MM) operation,
249 fuzzy logic approach(13), zack algorithm (44), K-means clustering(41), marker-controlled watershed
250 segmentation (17), stimulating discriminant measures (SDM) based clustering (17), Hough transform(31),
251 iterative thresholding followed by watershed transform (16), edge thresholding (53)(38), otsu's algorithm
252 (58)(47), a conventional neural network chen prepared laplacian of Gaussian (LOG) and coupled edge
253 profile active contours(C-EPAC) algorithm(47), triangular thresholding DOST algorithm(38), SMACC,
254 Iterative ISODATA clustering algorithm along with rotation and invariant LBP(Duan et al.).

Feature Extraction

256 There are number of features that could be considered for feature extraction purpose. Some of them are
257 given below

1. **Color Features** Color of the cell can be one of the features which can separate a cell from other types. For example is the color of plasma is very different (yellow) than other blood components. In many cases, the color of the cell talks much about the abnormalities.
2. **Geometric Features** These are the features based on the geometry or shape of the cell. These include following,and like this.

$$Elongation = 1 - \frac{majoraxis}{minoraxis} \tag{1}$$

$$Eccentricity = \frac{\sqrt{major_axis^2 - minor_axis^2}}{minor_axis} \tag{2}$$

$$Rectangularity = \frac{area}{major_axis \times minor_axis} \tag{3}$$

$$Convexity = \frac{Perimeter_convex}{Perimeter} \tag{4}$$

$$Compactness = \frac{4 \times pi \times area}{Perimeter^2} \tag{5}$$

3. **Statistical Features** Statistical moments such as mean and standard deviation gives information about the appearance of distribution. Skewness and kurtosis shape the distribution along with the area and perimeter of the shape. The following are the different statistical features.

$$Mean, \bar{x} = \frac{1}{N} \sum_{i=1}^{N} (x_i) \tag{6}$$

$$StandardDeviation, \sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (X_i - \bar{x})^2} \tag{7}$$

$$Skewness, SK = \frac{1}{N} \sum_{i=1}^{N} \frac{(x_i - \bar{x})^3}{\sigma^3} \tag{8}$$

$$Kurtosis, K = \frac{1}{N} \sum_{i=1}^{N} \frac{(x_i - \bar{x})^4}{\sigma^4} \tag{9}$$

4. **Texture Features** There are different texture features that are defined such as entropy, correlation, energy, contrast, homogeneity, and so on.

   Entropy generally defines randomness in the characterization of texture of an image. When co-occurrence elements are same, entropy leads to its maximum value. The equation of entropy as follows.

$$Entropy, Ent = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M(i,j)(-\ln(M(i,j))) \tag{10}$$

   Contrast is the intensity variations in the neighboring pixels in an image.

$$Con = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i-j)^2 (M(i,j)) \tag{11}$$

   Energy (E) is the measure of the extent of repetitions of pixel pairs. It gives an uniformity of the image. It gives a larger value for similar pixels.

$$Energy, E = \sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M^2(i,j)} \tag{12}$$

5. **Correlation Features** The repetitive nature of the texture elements position in the image is an important. An auto-correlation function gives the coarseness in an image.

$$Auto-correlation, P(x,y) = \frac{\sum_{u=0}^{N} \sum_{v=0}^{N} I(u,v)I(u+x,v+y)}{\sum_{u=0}^{N} \sum_{v=0}^{N} I^2(u,v)} \tag{13}$$

6. **Inverse Difference Moment or Homogeneity** gauges the local homogeneity of a picture. IDM features acquire the proportions of the closeness of the distribution of the GLCM components to the diagonal of GLCM. IDM has a scope of determining the image and classify it as textured or non-textured.

$$IDM = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1 + (i-j)^2} M(i,j) \tag{14}$$

7. **Directional Moment** In this, the image alignment is considered with respect to the angle.

$$DM = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M(i,j) |i = j| \tag{15}$$

263 *Classifier for Disease Detection*

264 There are different classifiers for the classification of images which are employed for microscopic imaging
265 of blood cells. These include machine learning algorithms as below.
266 **Decision Tree Classifier**
267
268   It falls under the supervised learning type. It is employed for regression as well as classification. It
269 has roots, branches, nodes, and leaves.
270 Figure 5 shows the different components of the decision tree. The bottom and the upper part are termed
271 roots. A node represents a feature in the dataset. A branch connects two nodes. Different decision tree
272 learning algorithms are there. These include ID3 (Iterative Dicotomizer3), C4.5, C5.0, CART. These
273 algorithms have different characteristics that decide their uses in a particular application.
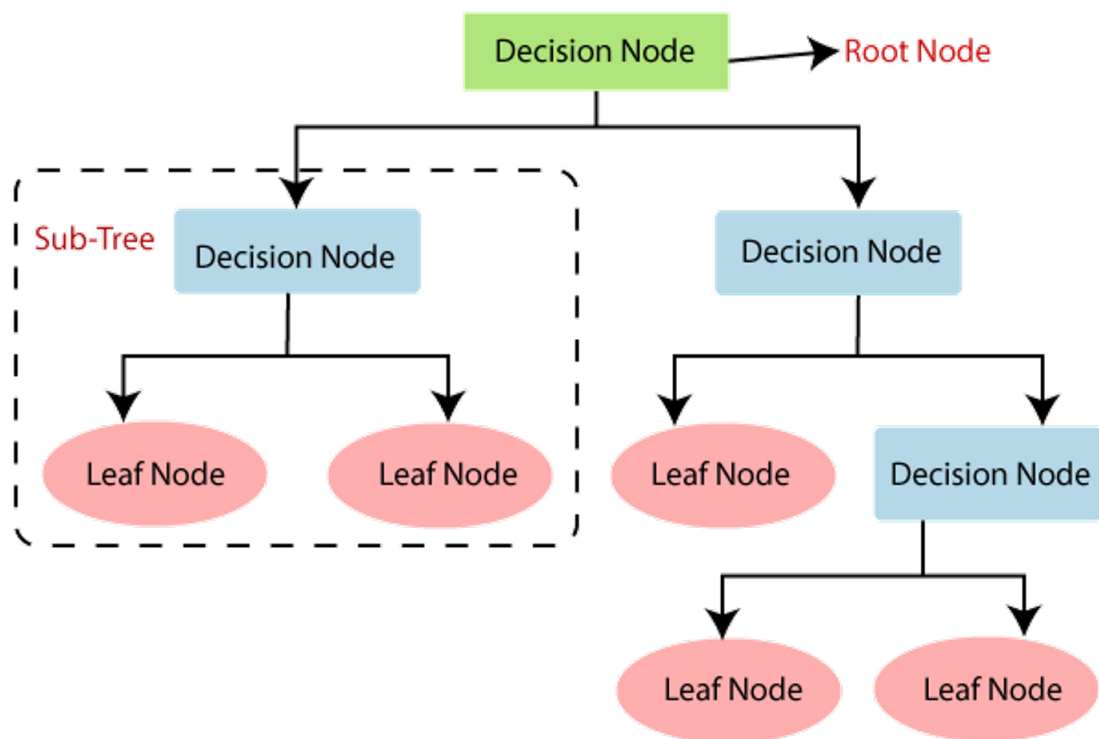
Figure 5. Decision tree
Source:https://pianalytix.com/decision-tree-algorithm/

274 In the ID3 algorithm, before pruning, trees are extended to maximum size. C4.5 does not require cate-
275 gorical features. It is useful in defining distinct attributes in the case of numerical features. C5.0 has a
276 comparatively smaller rule set and also takes less space than C4.5. The CART classification and regression
277 tree are similar to C4.5; the difference is, it does not calculate the ruleset, and it braces numerical target
278 variables. This type generates a binary tree.
279 Decision trees may be affected by the problem of overfitting, which could be further analyzed and is
280 taken care of. Regularization is the process to be adapted during the generation of decision trees for
281 compensating for the overfitting issue.
282 **Random Forest**
283
284 It has many decision trees, with each tree having different training sets. Hence this algorithm is more
285 effective in solving classification problems. An important issue in this algorithm is the selection of a
286 pruning method and branching criteria selection. Popular gain measurement techniques here are the Gain
287 ratio and the Gini index. This algorithm is based on the number of trees to be developed and samples used
288 by the particular node. Figure 6 shows the random forest algorithm.
289 **K-Nearest Neighbours (KNN)**
290 This classifier employs the nearest instances of training in the attribute space. Here according to the k
291 value of neighbor, the new sample value of the class is decided. For getting the class of the new vector, the
292 closest k samples from training data are selected. There are specific methods for calculating the distances
293 according to the classified samples. These are illustrated as follows.
294

$$Euclidean, \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{16}$$
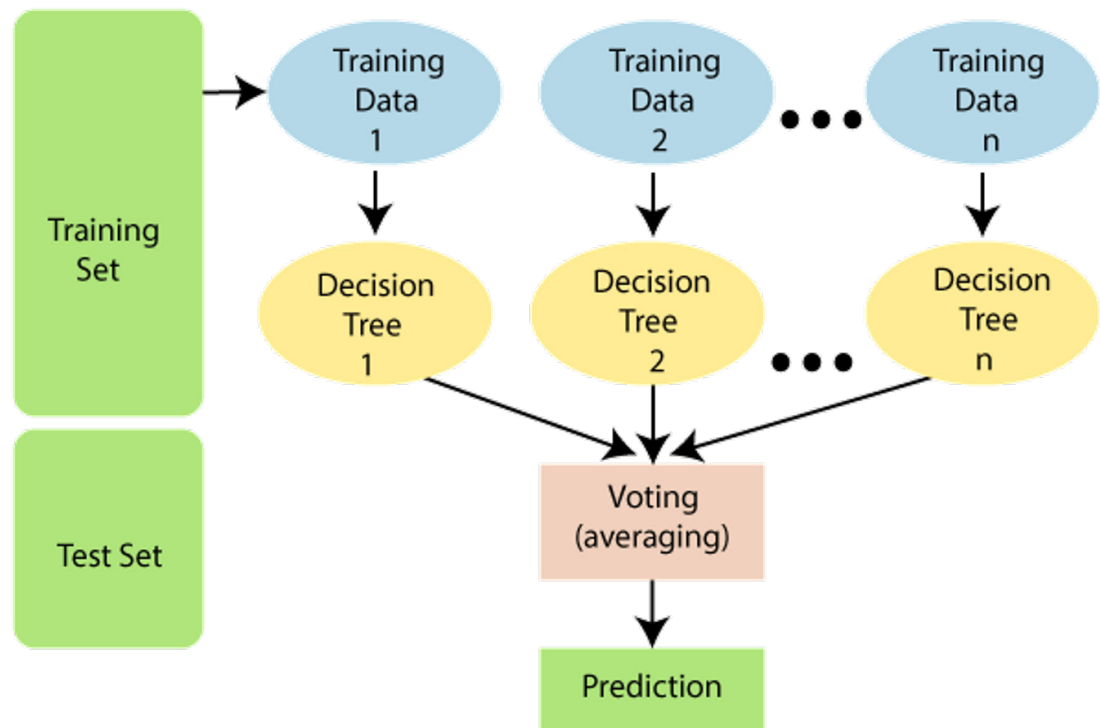
Figure 6. Random Forest Algorithm.
Source https://www.javatpoint.com/machine-learning-random-forest-algorithm

$$Manhattan, \sum_{i=1}^{k} |x_i - y_i| \tag{17}$$

$$Minkowski, \left(\sqrt{\sum_{i=1}^{k} (|x_i - y_i|)^q} \right)^{\frac{1}{q}} \tag{18}$$

Logistic Regression

Regression can be of either linear regression and logistic regressions. Linear regression is a supervised regression algorithm while logistic regression is a supervised classification algorithm. It is categorized into different types. These are binary, multinomial, and the last one is the ordinal. The first type of regression is the,

Binary Logistic Regression model

This is the simplest type where the dependent variable can have either 0 or 1 showing only two possible types. So here we have predictor variables multiple but the target variable is binary or binomial.

$$h_\theta(x) = g(\theta^T x), where 0 \le h_\theta \le 1 \tag{19}$$

g is the logistic or sigmoid function, given as below.

$$g(z) = \frac{1}{1 + e^{-z}}, where z = \theta^x \tag{20}$$

A loss function is defined to know the performance of the algorithm using the weights where,

$$h = g(X\theta) \qquad J(\theta) = \frac{1}{m}.(-y^T \log(h) - (1-y)^T \log(1-h)) \tag{21}$$

After this, loss function is minimized by decreasing and increasing the weights. The gradient descent tells us how the loss changes with the modification of parameters.

$$\frac{\delta J(\theta)}{\delta \theta_j} = \frac{1}{m} X^T (g(X\theta) - y) \tag{22}$$

Multinomial logistic regression Here 3 or more possible un-ordered types of dependent variables are present with no quantitative significance, such as different types A, B or C. The MLR model is an extension of LR (logistic regression) model. LR Model:

$$\pi(x) = \frac{\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots + \beta_p x_p}{1 + e^{\alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots + \beta_p x_p}} \tag{23}$$

$\pi(x)$ is an event probability, α represents the dependent variable,
β_1, β_2 are independent variables, $x_1, x_2 \ldots$ are regression coefficients,
p is the number of independent variables and e is the term representing
the error.

From this LR model, MLR can be obtained as an extended version as below. MLR Model:

$$\pi_j(x_i) = \frac{e^{\alpha_i + \beta_1 jx_1 i + \beta_2 jx_2 i + \ldots\ldots + \beta_p jx_p}}{1 + \sum_{j=1}^{k-1} e^{\alpha_i + \beta_1 jx_1 i + \beta_2 jx_2 i + \ldots + \beta_p jx_p}} \tag{24}$$

here $j_1, j_2 .. j_k$ are the k categories and $n(i_1, i_2 ... i_n)$
are the possible independent levels.

Ordinal regression

It deals with a quantitative significance having 3 or more ordered types for the dependent variable. As an example, variables can be "poor", "good", "very good " or "excellent" categories, can be represented as a score of 0 to 3.

Naïve Bayes Algorithm

It is based on Bayes' theorem and is a probabilistic algorithm. It assumes the occurrence of a feature independent of the other features. Bayes theorem forms its basis,

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} \tag{25}$$

It gives the relation between an arbitrary event X with y causing due to some random process. Here P(X) is the input probability, and P(Y) is considered as output probability, while P(Y—X) defines the probability of Y states versus the X input probability.

The major disadvantage of this algorithm is, it assumes all features as independent features. Hence not possible to get the relationship between the features.

It has three types viz Gaussian, Multinomial, and Bernoulli.

Gaussian Naive Bayes In this type, features have values that are continuous in nature and are considered to have Gaussian distribution.

Multinomial Naive Bayes: In this case, events are considered to occur in multinoial distribution type. The feature vector is formed by the frequencies of occurrence of some events. Typical use includes the document classification.

Bernoulli Naive Bayes: Here, input describes the features with a binary variable independent of each other. In this model, features used are binary term occurrence instead of frequencies.

Support Vector Machine (SVM)

Support vector machine (SVM) is a machine learning algorithm that has a supervised type. It finds its application in classification as well as regression. It has two types. Linear SVM, used for linearly separable data and non-linear SVM, is used for the data that is non-linearly separable; a non-linear type of SVM is used. Figure 7 explains the SVM and its different concepts. The followings are important terms in SVMs
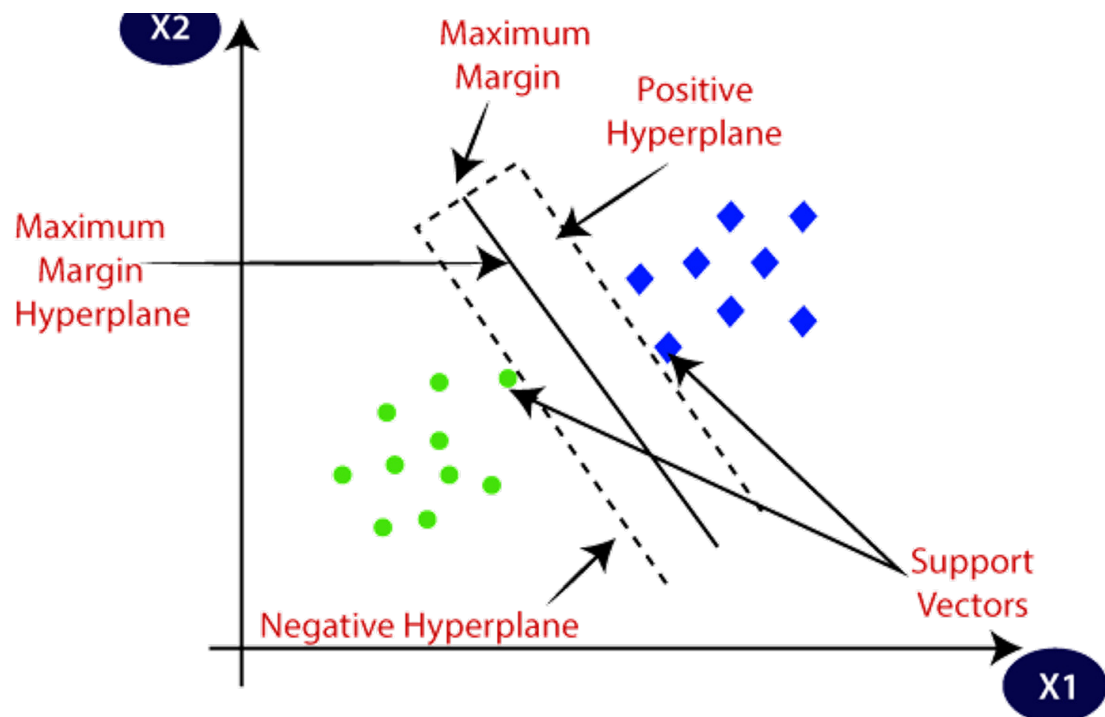
Figure 7. SVM and its Different Concepts
Source: https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm

Hyperplane is the space or plane for decision. It has a division in a set of different objects having separate classes.

Support Vectors These are the data points which are closest to the hyperplane. These data points define the separating line.

Margin The gap between two lines on Closest data points is considered to define the margin. It is given by the gap in the lines of different classes. The perpendicular distance between the line and the support vectors is used to find the margin. Larger margins are good, and smaller ones prove to be bad.

Convolutional Neural Networks

Figure 8 shows the generalized architecture of convolutional neural network(45). The CNN has two basic building blocks:

The Convolution Block — Consists of two layers, namely, the Convolution and the Pooling Layer. For feature extraction, these layers prove to be the essential components.

The Fully Connected Block — has a simple neural network that is fully connected architecture. This layer's function is the classification based on the input coming from the convolutional block.

Pooling Layer - It has the process where the extraction of a value is done from a set of values. It generally uses the maximum or the average value. This reduces the output matrix size. There are some popular architectures of CNN given as, Classic network architectures viz LeNet-5, AlexNet, VGG 16, and Modern architectures namely Inception, ResNeXt, ResNet, DenseNet.

Publicly Available Databases

There are different databases publicly available for the analysis of microscopic blood images. This Dataset contain microscopic images of different parts of blood cells, including white blood cells (WBC) and red blood cells (RBC), along with their sub-classes.

BCCD Database

Total images in this data are 12,500. Each subtype, Eosinophil, Lymphocyte, Monocyte, and Neutrophil, has approximately 3000 images. This Dataset also includes additional 410 images of WBC and RBC with JPEG and xml metadata format. This database is under MIT license(Yildirim and Çinar).

ALL-IDB (Acute Lymphoblastic Leukemia Database) This database is for the analysis and research
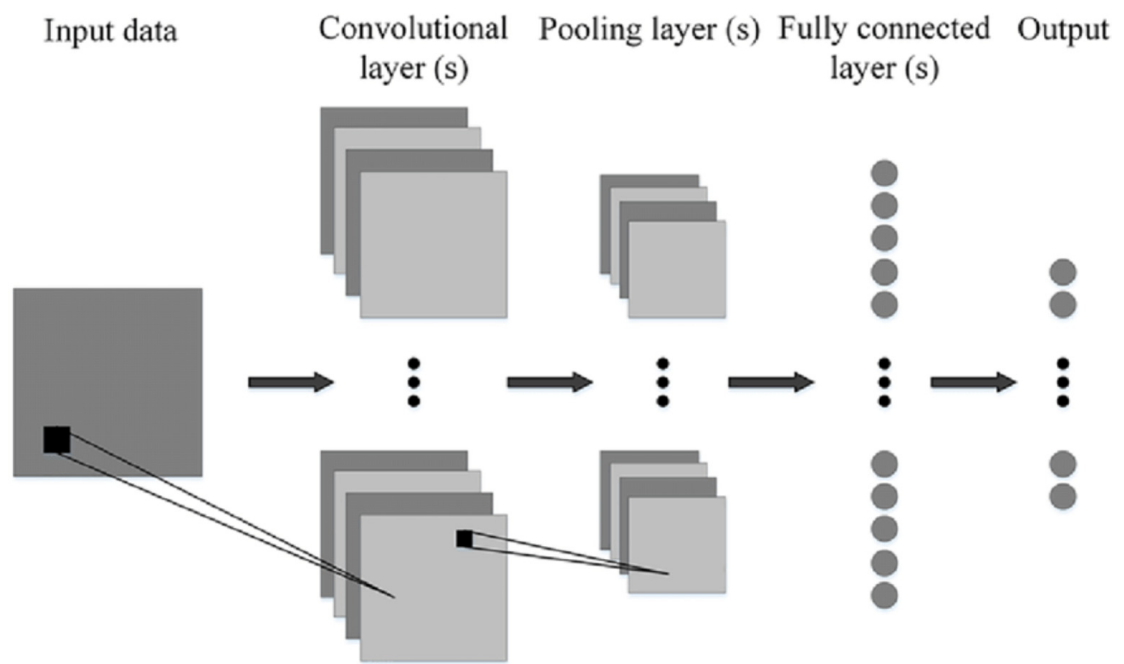
Figure 8. Generalized CNN Architecture (45)

341 on the disease of acute leukemia. This dataset has two subtypes ALL-IDB1 and ALL-IDB2. All images
342 are taken with Canon Powershot G5camera. The microscope has 300 to 500 magnification. Images follow
343 the JPEG format with a color depth of 24-bit. ALL-IDB1 dataset has 109 images with a resolution of
344 2592X1944 and a total element of 3900. It has a total of 510 lymphoblasts. ALL-IDB2 dataset has
345 260 images with a resolution of 257X257 and a total element of 257. It has a total lymphoblast of 130
346 (32)(18).

347 **Atlas of Hematology by Nivaldo Mediros**
348 This database contains different types of microscopic blood cell images. Different image types include
349 maturation sequence, alteration of erythrocytes, anemia images, leukemia images, parasites, fungus
350 images, and some miscellaneous images. The photos of images were taken with the magnification of 200,
351 400, 360, and 1000 by the photo-microscopes of Zeiss and Nicon(3).

352 **ASH Image Bank**
353 This database is from the American Society of Hematology (ASH). It is basically an online image bank of
354 having leukemia cells images. This image library web-based. This offers a wide collection of hematology
355 categories. They provide images with variations in resolutions(4).

356 **Leukocyte Images for Segmentation and Classification (LISC)**
357 This dataset contains images of healthy subjects. The total number of images in the dataset is 400 from
358 100 different blood slides of healthy eight subjects. The size of the images is 720X576. These images
359 are from Hematology-Oncology and BMT Research Center of Imam Khomeini Hospital in Tehran, Iran.
360 A total of 250 images are available for experimentation purposes. Images have different elements of
361 leukocyte like eosinophil, basophil, lymphocyte, monocyte, and neutrophil (50)(21).

362 **C-NMC Dataset**
363 This database is owned by The Cancer Imaging Archive (TCIA) public access. This database has a total
364 of 118 participants, with a total of 15135 images. All images are in bit map image (BMP) format. This
365 data is divided into train sets, preliminary test sets, and final test sets with different cancerous and normal
366 images. Table 1 shows the comparison of different publicly available datasets, and table 2 gives the
367 information of datasets availability along with urls.
368 Table 3 explores the different methods used by researchers along-with the performance measures, datasets,
369 and a number of images used for disease detection.

Name	Image Formats	Number of images	Color depth	Remark
BCCD Database(Yildirim and Çinar)	JPEG, xml, meta-data	12500	Not mentioned	Different sub-types of blood cells
ALL-IDB (Acute Lymphoblastic Leukemia Database ALL-IDB-1	JPEG	109 (510 lymphoblast)	24–bit 2592×1944	Cancerous
ALL-IDB-2(32)(18)	JPEG	260 (130 lymphoblast)	24–bit 257×257	Cancerous
Atlas of Hematology by Nivaldo Mediros(3)	JPEG	300	Not mentioned	Visceral leishmaniasis, cellular simlilarity, morphologic similarities
ASH Image Bank(4)	JPEG	5084	Not mentioned	Cancerous and other different types of images
Leukocyte Images for Segmentation and Classification.(50) (21)	(LISC) BMP	400 and 720×576		Healthy subjects with different sub-types of blood cells.
C-NMC Dataset	BMP	15135	Not mentioned	Normal and cancerous images of blood cells

Table 1. Different databases available

GAP ANALYSIS

After having reviewed the related literature, the following research gaps are obtained.

1. **Gaps in Segmentation of cells**

Overlapping cells are not considered at the segmentation stage by many researchers. As many practical cases have the overlapping of cells, during the staining procedure(Duan et al.)(51).

2. **Gaps in algorithms and methodology**

There are different bio-inspired optimization algorithms that are not used by most researchers for detection purposes. The hybrid system of different algorithms of the machine and deep learning also used by limited researchers that may give improved results(30).

3. **Gaps in system performance**

Performance measurement parameters in most of the cases are limited to accuracy only(44).

4. **Gaps in detected diseases**

Many diseases are not assured to be detected and diagnosed with their sub-types(63) This includes leukemia that has many sub-types. Its sub-types such as L1, L2, L3 are not considered in the case of acute lymphoblastic leukemia in most of the cases. In the case of acute myelogenous leukemia, its different subtypes T1 to T7, are not clearly detected in most of the cases(37)(54). There is less investigation in terms of stages of diseases that gives a threshold for determining the severity of the disease for anemia, thalassemia (14).

5.**Databases**

Accuracy of different stages of blood cell analysis is tested on a limited database.

6. **Gaps in critical decisions**

In medical imaging, artificial intelligence and deep learning are employed by most researchers. It is less acceptable in the case of critical decisions because the implementations of these algorithms have a black box. Hence its features can not be analyzed, and there is always a problem of getting the right detection by wrong reasons.

Name of Dataset	Available at url
BCCD Database(Yildirim and Çinar)	https://github.com/Shenggan/$BCCD_Dataset$, $https://www.kaggle.com/paultimothymooney/blood-cells$
ALL-IDB-1 and 2(32) (18)	https://homes.di.unimi.it/scotti/all/
Atlas of Hematology by Nivaldo Mediros(3)	http://www.hematologyatlas.com/morphologicsimilarities/drmedeiros.htm
ASH Image Bank(4)	https://imagebank.hematology.org/
Leukocyte Images for Segmentation and Classification.(50) (21)	http://users.cecs.anu.edu.au/ hrezatofighi/Data/Leukocyte%20Data.htm
C-NMC Dataset	https://wiki.cancerimagingarchive.net/display/Public/ $C_NMC_2019 + Dataset\%3AALLChallengedatasetof ISBI2019$

Table 2. Different databases available

395

Discussion of explainable AI for critical analysis

In medical imaging, artificial intelligence is mostly used for detection and diagnosis(6)(59). It is generally not been preferred as the final decision about the diagnosis. The main reason for this is, it works as a black box with only input and outputs are known. There might be the occurrence of the right decisions related to the diagnosis, but it might be due to wrong reasons. The algorithms such as decision trees elaborate their implementation to a good extent. However, it limits the parameters such as accuracy of diagnosis, although advanced AI algorithms and deep learning algorithms assure good accuracy in terms of diagnosis but unable to explain the insides of implementation-black box(35). Hence, explainable AI came into the picture to justify the trust of diagnosis in medical imaging(55). This will analyze the black box's features and characteristics in the implementation of AI or deep learning algorithms. In order to make the system to be used with more trust for commercial purposes for the general public, the explainability will prove to be the most suitable and appropriate.

CONCLUSION

Blood cell analysis assumes a crucial job in location and expectation of various issues and maladies identified with the person. There are distinctive neurotic strategies for the equivalent by and by, which ends up being exorbitant and furthermore requires a long understanding of the location. Image processing and computer vision strategies are produced for the investigation of blood cells and the discovery of maladies. The microscopic blood cell analysis framework has various stages of being specific, pre-processing, segmentation, feature extraction, classifier, and illness identification. Pre-processing comprises improving the gained picture quality and commotion expulsion. This incorporates gray-scale conversion, thresholding, filtering, histogram stretching, morphological operations. Pre-processed image is portioned to get the locale of interest for further processing. Here WBC and RBC, and platelets are isolated. Distinctive computer vision techniques utilized for segmentation are edge detection, watershed transformation, mathematical morphology, zack algorithm, k-means clustering, SDM, HSV thresholding, otsu's algorithm. There are overlapping cells at the time of staining of blood smear. Expulsion of these overlapping cells at the time of segmentation is a difficult undertaking. Hough transform evacuates certain overlapping; however, it makes the framework slower. Segmented images are classified by algorithms like SVM, ANN classifier, ELM classifier, circular hough transform. There are various databases accessible for experimentation and investigation of microscopic blood cells, such as BCCD (Kaggle) Database, ALL-IDB1, ALL-IDB2, Atlas of Hematology by Nivaldo Meridos, Leukocyte pictures for division and characterization (LISC), Ash image bank, and C-NMC dataset. There are different application territories where microscopic blood cell examination assumes a crucial job. RBC, WBC count, blood group identification, leukemia detection, sickle cells detection, the partition of various WBC sub-classes, malaria parasite detection, could be performed utilizing complex image processing and computer vision

Author	Year	Methodology	Performance Measure	Database	No. of images
Nimesh Patel and Ashutosh Mishra (44)	2015	K-means clustering for detection of WBC. Histogram and Zack algorithm for grouping WBCs, SVM for classification	Efficiency: 93.57 -%	ALL-IDB	7
Siew Chin Neoh et al.(17)	2015	Multilayer perceptron, Support Vector Machine (SVM) and Dempster Shafer	Accuracy :Dempster-Shafer method: 96.72% SVM model: 96.67%	ALL-IDB2	180
Ahmed S. Negm et al. (41)	2018	Panel selection for segmentation, K-means clustering for features extraction, and image refinement. Classification by morphological features of leukemia cells detection	Accuracy: 99.517% ,Sensitivity: 99.348%,Specificity: 99.529%.	private datasets	757
Sarmad Shafique et al.(54)	2019	Histogram Equalization, Zack Algorithm, Watershed Segmentation, Support Vector Machine (SVM) classification	Accuracy: 93.70% Sensitivity: 92% Specificity: 91%	ALL-IDB	108
Mohamadreza Abbasi et.al.(2)	2019	K-means and watershed algorithm, SVM, PCA	Accuracy, specificity, sensitivity, FNR, precision all are above 97%	private	Not mentioned
Sonali Mishra et.al.(38)	2019	Triangle thresholding, discrete orthogonal S-transform(DOST), adaboost algorithm with random forest(ADBRF) classifier	Accuracy: 99.66%	ALL-IDB1	108
Krishna Kumar et.al.[15]	2019	MI based model, local directional pattern (LDP)chronological sine cosine algorithm (SCA)	Accuracy: 98.7%, TPR:987%, TNR:98%	AA-IDB2	not mentioned
Mohamadreza Abbasi et.al.(2)	2019	K-means and watershed algorithm, SVM, PCA	Accuracy, specificity, sensitivity, FNR, precision all are above 97%	private	Not mentioned
Naima Iltaf et.al.(57)	2019	expectation maximization algorithm, PCA, sparse represenataion	Accuracy, Specificity, Senstivity all more than 92%	ALL-IDB2	260
Nizar Ahmed et.al. (5)	2019	CNN	Accuracy: 88% leukemia cells and 81% for subtypes classification	ALL-IDB, ASH Image Bank	not mentioned
Christian Matek et.al.(36)	2019	ResNeXt CNN	Accuracy, Sensitivity and precision above 90%	private	18365
Ahmed T. Sahlol et.al.(52)	2020	VGGNet, statistically enhanced Salp Swarm Algorithm (SESSA)	Accuracy:96%dataset1 and 87.9% for dataset2	ALL-IDB, C-NMC	not mentioned

Table 3. Comparison of different methods for disease detection

techniques. There is always a need to get proper and trusted diagnoses in the medical science field. Machine learning and computer vision can prove the system of suggestive diagnosis to achieve better accuracy than existing approaches. In a major case, when the morphology is concerned with microscopic imaging, pathologists' experience and skillset always matter. There is a need to have software frameworks utilizing machine learning and artificial intelligence to conquer this problem. Also, there is a problem of explainability of AI algorithms towards the particular diagnostic decisions. For this to take care, AI explainable frameworks could be utilized in the future. For the treatment of different diseases, AI can also put its role in a very trusted manner.

ACTIVE RESEARCH QUESTIONS AND DIRECTIONS TO FUTURE RESEARCH

After an extensive literature survey, the following research questions were raised, motivating further research in this field.

- RQ.1 How to increase the performance of different stages of microscopic analysis for disease detection?

- RQ.2 How to increase the accuracy of detection of critical diseases such as leukemia and its subtypes by using the morphological diagnosis?

- RQ.3 What is the societal impact of analyzing critical diseases at their early stages and defining the stage of diseases?

- RQ.4 Why is there a need to apply hybridized algorithms for the classification of microscopic cells for disease detection?

- RQ5 What are the different performance measures and validation techniques for analyzing the designed framework's performance?

- RQ6 Which validation techniques are to be employed for system validation of AI implemented systems?

- RQ7 How could AI be trusted and used commercially for medical diagnosis by analyzing the features of the black box in AI implementation?

Following are the future perspectives of the work to improve the contributions in this area.
A powerful division of white and red cells in minuscule blood smear pictures to meet better precision could be actualized. To conquer overlapping cells issue at the hour of division will likewise end up being a significant extension. Viable feature extraction by utilizing distinctive image transforms will likewise demonstrate to a significant degree. Different optimization algorithms could be utilized efficiently for the classification of blood cells. Different deep learning algorithms may demonstrate productivity and might give high accuracy to various phases of examining blood cells. The designed algorithms must be tasted with various publicly accessible databases for precision. The precision of the calculation should be similar enough with all the databases. Another parameter like vigor can be presented for this reason. The relative accuracy of various databases can be determined. To gauge the exhibition of a framework with various measures such as true positive, true negative, faults, sensitivity, specificity, precision, FI score, J-score in addition with accuracy. The contribution is still needed for various ailments location, such as diabetes, viral diseases such as chikungunya and dengue, anemia diseases such as pancytopenia, thalassemia, and leukemia.

REFERENCES

[1] Abbas, S. (2015). Microscopic images dataset for automation of RBCs counting. *Data in Brief*, 5:35–40.

[2] Abbasi, M., Kermani, S., Tajebib, A., Amin, M. M., Abbasi, M., et al. (2019). Automatic detection of acute lymphoblastic leukaemia based on extending the multifractal features. *IET Image Processing*, 14(1):132–137.

[3] Acharya, V. and Kumar, P. (2019). Detection of acute lymphoblastic leukemia using image segmentation and data mining algorithms. *Medical and Biological Engineering and Computing*, 57(8):1783–1811.

[4] Agaian, S., Madhukar, M., Chronopoulos, A. T., et al. (2014). Automated screening system for acute myelogenous leukemia detection in blood microscopic images. *IEEE Systems journal*, 8(3):995–1004.

[5] Ahmed, N., Yigit, A., Isik, Z., Alpkocak, A., et al. (2019). Identification of leukemia subtypes from microscopic images using convolutional neural network. *Diagnostics*, 9(3):104.

[6] Ahuja, A. S. (2019). The impact of artificial intelligence in medicine on the future role of the physician. *PeerJ*, 7:e7702.

[7] Al-Obeidat, F., Tubaishat, A., Shah, B., Halim, Z., et al. (2020). Gene encoder: a feature selection technique through unsupervised deep learning-based clustering for large gene expression data. *Neural Computing and Applications*, pages 1–23.

[8] Al-Tahhan, F., Sakr, A. A., Aladle, D. A., Fares, M., et al. (2019). Improved image segmentation algorithms for detecting types of acute lymphatic leukaemia. *IET Image Processing*, 13(13):2595–2603.

[9] Alom, M. Z., Yakopcic, C., Taha, T. M., Asari, V. K., et al. (2018). Microscopic blood cell classification using inception recurrent residual convolutional neural networks. In *NAECON 2018-IEEE National Aerospace and Electronics Conference*, pages 222–227. IEEE.

[10] Alotaibi, K. (2016). Sickle Blood Cell Detection Based on Image Segmentation.

[11] Bala, S. and Doegar, A. (2015). Automatic Detection of Sickle cell in Red Blood cell using Watershed Segmentation. 4(6):488–491.

[12] Belekar, S. J. and Chougule, S. R. (2015). WBC Segmentation Using Morphological Operation and SMMT Operator - A Review. pages 434–440.

[13] Bhagavathi, S. L. and Thomas Niba, S. (2016). An automatic system for detecting and counting rbc and wbc using fuzzy logic. *ARPN Journal of Engineering and Applied Sciences*, 11(11):6891–6894.

[14] Bhanushali, A., Katale, A., Bandal, K., Barsopiya, V., Potey, M., et al. (2016). Automated Disease Diagnosis Using Image Microscopy. (02):2–6.

[15] Bhavnani, L. A., K. Jaliya, U., J. Joshi, M., et al. (2016). Segmentation and Counting of WBCs and RBCs from Microscopic Blood Sample Images. *International Journal of Image, Graphics and Signal Processing*, 8(11):32–40.

[16] Biswas, S. and Ghoshal, D. (2016). Blood Cell Detection Using Thresholding Estimation Based Watershed Transformation with Sobel Filter in Frequency Domain. *Procedia Computer Science*, 89:651–657.

[17] Chin Neoh, S., Srisukkham, W., Zhang, L., Todryk, S., Greystoke, B., Lim, C. P., Hossain, M. A., Aslam, N., et al. (2015). An Intelligent Decision Support System for Leukaemia Diagnosis using Microscopic Blood Images. *Scientific Reports*, 5:1–14.

[18] Conference, I. I. and Processing, I. (2011). ALL-IDB : The Acute Lymphoblastic Leukemia Image Database for Image Processing, Ruggero Donida Labati, Vincenzo Piuri, Fabio Scotti Università degli Studi di Milano , Department of Information Technology ,. *Ieee International Conference On Image Processing*, pages 2089–2092.

[19] Deshpande, N. M., Gite, S. S., Aluvalu, R., et al. (2020). A brief bibliometric survey of leukemia detection by machine learning and deep learning approaches.

[Duan et al.] Duan, Y., Wang, J., Hu, M., Zhou, M., Li, Q., Sun, L., Qiu, S., Wang, Y., et al. Leukocyte classification based on spatial and spectral features of microscopic hyperspectral images. *Optics & Laser Technology*, 112.

[21] E Livieris, I. (2018). Identification of Blood Cell Subtypes from Images Using an Improved SSL Algorithm. *Biomedical Journal of Scientific & Technical Research*, 9(1).

[22] Elsalamony, H. A. (2016). Healthy and unhealthy red blood cell detection in human blood smears using neural networks. *Micron*, 83:32–41.

[23] Halim, Z., Ali, O., Khan, G., et al. (2019). On the efficient representation of datasets as graphs to mine maximal frequent itemsets. *IEEE transactions on knowledge and data engineering*.

[24] Halim, Z. et al. (2020). Optimizing the dna fragment assembly using metaheuristic-based overlap layout consensus approach. *Applied Soft Computing*, 92:106256.

[25] Halim, Z., Muhammad, T., et al. (2017). Quantifying and optimizing visualization: An evolutionary computing-based approach. *Information Sciences*, 385:284–313.

[26] Houwen, B. (2002). Blood Film Preparation and Staining Procedures, B. Houwen (2000) Laboratory

530   Hematology 6:1-7. 22(1):1–14.

[27]   Iqbal, S. and Halim, Z. (2020). Orienting conflicted graph edges using genetic algorithms to discover pathways in protein-protein interaction networks. *IEEE/ACM transactions on computational biology and bioinformatics*.

[28]   Javidi, B., Markman, A., Rawat, S., O'Connor, T., Anand, A., Andemariam, B., et al. (2018). Sickle cell disease diagnosis based on spatio-temporal cell dynamics analysis using 3D printed shearing digital holographic microscopy. *Optics Express*, 26(10):13614.

[29]   Jha, K. K., Dutta, H. S., et al. (2019a). Mutual Information based hybrid model and deep learning for Acute Lymphocytic Leukemia detection in single cell blood smear images. *Computer Methods and Programs in Biomedicine*, 179:104987.

[30]   Jha, K. K., Dutta, H. S., et al. (2019b). Mutual information based hybrid model and deep learning for acute lymphocytic leukemia detection in single cell blood smear images. *Computer methods and programs in biomedicine*, 179:104987.

[31]   Kaur, M. P. and Cells, A. N. B. (2015). Significant Analysis of Leukemic Cells Extraction and Detection Using KNN and Hough Transform Algorithm. 3(1):27–33.

[32]   Labati, R. D., Piuri, V., and Scotti, F. (2011). All-idb: The acute lymphoblastic leukemia image database for image processing. In *2011 18th IEEE International Conference on Image Processing*, pages 2045–2048. IEEE.

[Lavanya et al.]   Lavanya, T., Gubbi, T., Sushritha, S., et al. Detection of sickle cell anemia and thalassemia using image processing techniques.

[34]   Loddo, A., Ruberto, C. D., Putzu, L., et al. (2016). Peripheral blood image analysis Peripheral Blood Image Analysis Preparation of Doctoral Consortium Contributions. (November).

[35]   Lundervold, A. S., Lundervold, A., et al. (2019). An overview of deep learning in medical imaging focusing on mri. *Zeitschrift für Medizinische Physik*, 29(2):102–127.

[36]   Matek, C., Schwarz, S., Spiekermann, K., Marr, C., et al. (2019). Human-level recognition of blast cells in acute myeloid leukaemia with convolutional neural networks. *Nature Machine Intelligence*, 1(11):538–544.

[37]   Miao, H. and Xiao, C. (2018). Simultaneous Segmentation of Leukocyte and Erythrocyte in Microscopic Images Using a Marker-Controlled Watershed Algorithm. *Computational and Mathematical Methods in Medicine*, 2018.

[38]   Mishra, S., Majhi, B., Sa, P. K., et al. (2019). Texture feature based classification on microscopic blood smear for acute lymphoblastic leukemia detection. *Biomedical Signal Processing and Control*, 47:303–311.

[39]   Mohite-patil, M. B. C. T. B. (2016). Automated Red Blood Cells Counting using Image Processing Techniques. 3(12):748–750.

[40]   Moshavash, Z., Danyali, H., Helfroush, M. S., et al. (2018). An Automatic and Robust Decision Support System for Accurate Acute Leukemia Diagnosis from Blood Microscopic Images. *Journal of Digital Imaging*, 31(5):702–717.

[41]   Negm, A. S., Hassan, O. A., Kandil, A. H., et al. (2018). A decision support system for Acute Leukaemia classification based on digital microscopic images. *Alexandria Engineering Journal*, 57(4):2319–2332.

[42]   Nwogoh, B. et al. (2014). Peripheral Blood Film - A Review. 12(2):71–79.

[43]   Patel, N. and Mishra, A. (2015a). Automated leukaemia detection using microscopic images. *Procedia Computer Science*, 58:635–642.

[44]   Patel, N. and Mishra, A. (2015b). Automated Leukaemia Detection Using Microscopic Images. *Procedia Computer Science*, 58:635–642.

[45]   Phung, V. H., Rhee, E. J., et al. (2019). A high-accuracy model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets. *Applied Sciences*, 9(21):4500.

[46]   Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S., Thoma, G., et al. (2018a). Image analysis and machine learning for detecting malaria. *Translational Research*, 194:36–55.

[47]   Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S., Thoma, G., et al. (2018b). Image analysis and machine learning for detecting malaria. *Translational Research*, 194:36–55.

[48]   Putzu, L., Ruberto, C. D., et al. (2013). White Blood Cells Identification and Counting from Microscopic Blood Image. *World Academy of Science, Engineering and Technology*, 7(1):363–370.

[49] Razzak, M. I. (2017). Microscopic Blood Smear Segmentation and Classification using Deep Contour Aware CNN and Extreme Machine Learning King Saud bin Abdulaziz University for Health Sciences. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[50] Rezatofighi, S. H. and Soltanian-Zadeh, H. (2011). Automatic recognition of five types of white blood cells in peripheral blood. *Computerized Medical Imaging and Graphics*, 35(4):333–343.

[51] Safca, N., Popescu, D., Ichim, L., Elkhatib, H., Chenaru, O., et al. (2018). Image processing techniques to identify red blood cells. *2018 22nd International Conference on System Theory, Control and Computing, ICSTCC 2018 - Proceedings*, pages 93–98.

[52] Sahlol, A. T., Kollmannsberger, P., Ewees, A. A., et al. (2020). Efficient classification of white blood cell leukemia with improved swarm optimization of deep features. *Scientific Reports*, 10(1):1–11.

[53] Service, A. N. P. and Wildlife (1986). Special issue. Special issue. *Australian ranger bulletin*, 4(1):9–10.

[54] Shafique, S., Tehsin, S., Anas, S., Masud, F., et al. (2019). Computer-assisted Acute Lymphoblastic Leukemia detection and diagnosis. *2019 2nd International Conference on Communication, Computing and Digital Systems, C-CODE 2019*, pages 184–189.

[55] Singh, A., Sengupta, S., Lakshminarayanan, V., et al. (2020). Explainable deep learning models in medical image analysis. *arXiv preprint arXiv:2005.13799*.

[56] Singh, H. and Kaur, G. (2017). Automatic Detection of Blood Cancer in Microscopic Images: A Review. *Dr. Balkrishan International Journal of Innovations & Advancement in Computer Science*, 6(4):40–43.

[57] Sukhia, K. N., Ghafoor, A., Riaz, M. M., Iltaf, N., et al. (2019). Automated acute lymphoblastic leukaemia detection system using microscopic images. *IET Image Processing*, 13(13):2548–2553.

[58] Thiruvinal, V. J. and Ram, S. P. (2017). Automated Blood Cell Counting and Classification Using Image Processing. pages 74–82.

[59] Tizhoosh, H. R. and Pantanowitz, L. (2018). Artificial intelligence and digital pathology: challenges and opportunities. *Journal of pathology informatics*, 9.

[60] Vives Corrons, J.-L., Albarède, S., Flandrin, G., Heller, S., Horvath, K., Houwen, B., Nordin, G., Sarkani, E., Skitek, M., Van Blerk, M., Libeer, J.-C., Haematology Working Group of the European External Committee for External Quality Assurance Programmes in Laboratory Medicine, et al. (2004). Guidelines for blood smear preparation and staining procedure for setting up an external quality assessment scheme for blood smear interpretation. part i: Control material. *Clinical chemistry and laboratory medicine*, 42(8):922—926.

[61] Xia, T., Jiang, R., Fu, Y. Q., Jin, N., et al. (2019). Automated blood cell detection and counting via deep learning for microfluidic point-of-care medical devices. In *IOP Conference Series: Materials Science and Engineering*, volume 646, page 012048. IOP Publishing.

[Yildirim and Çinar] Yildirim, M. and Çinar, A. Classification of white blood cells by deep learning methods for diagnosing disease classification of white blood cells by deep learning methods for diagnosing disease.

[63] Z., M., S., T., B., A., et al. (2017). Neural Network Classification of White Blood Cell using Microscopic Images. *International Journal of Advanced Computer Science and Applications*, 8(5):99–104.

[64] Zhang, A.-h., Sun, H., Wang, X.-j., et al. (2013). Recent advances in metabolomics in neurological disease, and future perspectives. *Analytical and bioanalytical chemistry*, 405(25):8143–8150.