Innov Biosyst Bioeng, 2024, vol. 8, no. 2, 53–68
doi: 10.20535/ibb.2024.8.2.298201

Original article

RECOGNITION AND CATEGORIZATION OF BLOOD GROUPS BY MACHINE LEARNING AND IMAGE PROCESSING METHOD

Mustafa F. Mahmood

Department of Medical Instrumentation Techniques Engineering, Electrical Engineering Technical College, Middle Technical University, Baghdad, Iraq

Corresponding author: mustafa.falah@mtu.edu.iq

Received 7 February 2024; Accepted 20 May 2024

Background. Red blood cells are one of the components of blood. Blood is an important fluid in the human body. Knowing the blood groups is essential in blood transfusion operations, which depend on fixed conditions to avoid fatal errors. The method that is used to determine the blood groups is a traditional method that relies on medical laboratory technicians, as it is subject to human errors.
Objective. This paper aims to design and implement a prototype to detect and classify blood groups to avoid human error in blood group detection. The proposed system employs image processing and machine learning algorithms for blood group detection and classification.
Methods. The system consists of three stages. First, samples were collected from volunteers. Second, images of the samples were captured using a camera. Third, the images were analyzed using two methods: image processing via MATLAB and machine learning algorithms via Orange, for blood group detection and classification.
Results. The accuracy in processing images using the MATLAB program reached 100%, with processing time ranged from 1.5 to 1.6 seconds. Additionally, using machine learning with neural networks in the Orange program, the accuracy was 99.7%, with training times of 13.7 seconds and testing times of 1.2 seconds. Neural networks outperformed other models, as shown in the experimental results. The study concluded that automated blood type detection using image processing and machine learning methods is effective and feasible compared to manual methods. The proposed system outperformed previous studies in terms of accuracy, processing time, training time, and testing time using both methods.
Conclusions. This study underscores the urgent need for precise blood type determination before emergency blood transfusions, which currently relies on manual inspection and is susceptible to human errors. These errors have the potential to endanger lives during blood transfusions. The main goal of the research was to develop an approach that combines image processing and machine learning to accurately classify blood groups.
Keywords: blood group classification; MATLAB image processing; Orange machine learning; processing time.

Introduction

Blood is an effective liquid substance in the human body, as it transports oxygen and necessary materials to the body through the heart in two blood circulations. It consists of red and white blood cells, plasma, and platelets [1]. It constitutes about 8% of body weight [2]. If the body is exposed to an accident or chronic blood disease, it needs to replace the lost blood by transfusing blood from one person to another. Before the transfer process, confirm the blood type by examining the blood type. Red blood cells consist of four basic groups: A, B, AB, and O, classified according to the substances contained in them. They also contain a protein on the surface of red cells called the rhesus factor (Rh factor). If the cells contain this protein, the factor is Rh positive; Otherwise, the factor is Rh negative [3]. In addition, it contains ABO antigens on the surface of red cells to classify blood groups. The blood group contains an antibody. If the blood drops are granulated using a kit, the kit is called a kit for this type. The test uses three blood type groups (A, B, D). A and B antibodies are responsible for the four blood types A, B, AB, and O; The other is responsible for the Rh factor. For example, the first drop is granular and the rest is granular, it is called type A-. In addition, the first and third drops are granular, and the second drop is non-granular, and is called type A+. As well as the rest of the blood types.

Recently, there has been significant development in computer assistance in classifying blood groups and in various clinical applications such as measuring blood pressure, temperature, etc. due to speed and reduction of human errors [4, 5]. Computer programs were used to detect or determine blood types, including MATLAB, Python, Orange, etc.

In the literature, many theories have been proposed to explain the identification and classification of blood groups and Rh factors using several methods such as image processing or analysis and machine learning using several algorithms such as SVM, KNN, Tree, and neural networks, summarized based on the program used in terms of image processing or analysis and counting of granules in blood drops, as some of the works poor some parameters such as processing time and accuracy [6–16]. In addition, some work classified blood groups using machine learning from the standard method, as some work lacked the time for analysis and learning and the accuracy of classification [17–20]. Moreover, some work classified blood regarding image processing and machine learning, as it was poor learning and test times [21–23]. Ferraz *et al.* [6] studied the classification of the red blood group via image processing. The system involves blood samples, a camera, a PC, and IMAQ-Vision software. The experimental results showed that the system enables the classification of blood groups according to the granules present in each drop of blood. In addition, it takes around 2 minutes to process. Regardless, there are a lot of benefits to the system, including its cheap price and easy design process. Nevertheless, the system was a prototype, limited to a sample and low had accuracy. Dhande *et al.* [7], Rahman *et al.* [8], Vatshav *et al.* [9], and Ravindran *et al.* [10] classified of blood groups according to image processing using MATLAB. Where the system used (i.e. samples, camera, and processing methods). The experimental results show that systems able to detect blood group using image processors in MATLAB. In addition, human errors decreased compared with the traditional method. The systems shared several benefits: they were cheap and easy to design. However, it is limited to samples and accuracy and processing time are not calculated. Shaban *et al.* [11] used image manipulations to display blood aggregates. Blood samples and taking pictures of the samples using a camera and MATLAB are the components of this system. Experiments have shown that the accuracy of the system is 98%. Pavithra *et al.* [12] implements image processing for blood group detection using MATLAB graphical interfaces. Samples, camera, and processing represent the components of the paper. Experiments have proven that the system is capable of detecting blood type with 90% accuracy. Yamin *et al.* [13] presented a mechanism to detect blood type using image processing in MATLAB. Photographs of blood samples and image processing are components of this study. Experiments have achieved that it is able to detect blood type and Rh factor with an accuracy of 98%. The systems are shared several advantages such as simple and low-cost. Although the systems did not calculate the time to detect the blood groups. Jamil *et al.* [14] designed a system capable of using image processors in MATLAB and capable of finding blood type and Rh factor. Images of blood samples and image processing are the tools of this paper. The experimental results of the study showed that the image processing time reached 16 seconds, and the accuracy and sensitivity of the system were 96% and 95%, respectively. Sahastrabuddhe *et al.* [15] explained the method of detecting blood types and Rh factor using the image processing unit in MATLAB. Sample images and image processing are considered components of this study. Experimental results showed that the accuracy of blood type detection is 90%, with a time ranging between 4–6 seconds. The systems are simple and low-cost. Sathiyan *et al.* [16] implemented a system that uses image processing to detect blood aggregates. Images of samples and processing using a Raspberry Pi 3 microcontroller are considered components of the paper. The system had several benefits, was cheap and easy to design. However, it is limited to samples and accuracy and processing time are not calculated.

Using methods from computer vision and machine learning, Ferraz and colleagues created a prototype for blood type categorization Ferraz *et al.* [17], data processing, the SVM method, and a sample are all part of the system. There was evidence of system compatibility with blood grouping in the experiments. We also used standard deviation and quick Fourier transform for the evaluation. The system has several benefits, such as being easy to develop and affordable. Be that as it may, the system did not meet expectations in regards to precision, testing time, or sample size. A technique for categorising red blood types using a Support Vector Machine was presented by Gurav *et al.* [18], a camera, a computer, preprocessing procedures, feature extraction techniques, and the Support Vector Machine (SVM) algorithm are all part of the system. It also requires blood samples. Results from experiments demonstrated that the method works by examining the granulation in blood droplets, which allows for the classification of blood types. A few of the benefits of the system include its low price and easy design. It must be emphasised that the system was still in its early phases of development, therefore its accuracy was not ideal and it collected minimal data. Using image processing techniques and artificial intelligence, the authors were able to correctly identify blood types inside a certain system Bhagat *et al.* [19], using the program's graphical user interface, the system gathers image samples and analyses them. Results from the trials showed that the suggested technology successfully used a certain approach to categorise blood types. Reducing the room for human mistake while interpreting fractions is another way to improve the standard method. The system had several benefits was cheap and easy to design. However, it is limited to samples and accuracy and processing time are not calculated.

Odeh *et al.* [20] designed a system using computer vision technology to determine blood types using a machine learning method. Blood group samples, a reference image, and a computer camera are the components of this study. The accuracy of the system in detecting blood types and the Rh factor was 95.3%, as the test results showed. Rosales *et al.* [21] demonstrate the use of image processing and machine learning in determining and classifying blood types. Pictures of samples and the method of detection or classification are the components of the paper. Experimental results showed that the accuracy of the system in detecting blood type was 100%, in addition to the accuracy of the system using the learning mechanism, which was 97% correct using SVM and Tree. The systems have several benefits was cheap and easy to design. However, it is limited to samples and training time are not calculated. Vadgave *et al.* [22] detected and classified quantum and factorial species using a system based on learning and image processing. Images of blood samples, image processing, learning machine (SVM and KNN) algorithms are the components of this study. Dada *et al.* [23] presented a system for detecting blood types using a learning mechanism. The system relied on images of samples and the processing method using Ortho-AutoVue, which are the components of this study. Experimental results showed that the system reduced classification time to 17%. The systems offers a number of advantages, such as low cost and easy design. Nevertheless, the prototype gadget displayed less-than-ideal accuracy and training times.

The contributions in this paper can be summarized as follows:

- Design and implementation of two applications for detecting and classifying blood types using two separate programs MATLAB and Orange.
- Using MATLAB to process images to detect blood types. In addition, the Orange program classifies blood types using machine learning for several methods such as logistic regression, neural network, tree, KNN, and SVM.
- Optimization of parameters such as accuracy, learning time, testing time and processing time based on previous studies.
- Using the proposed system in medical laboratories, hospitals and health centers without any intervention by laboratory technicians.

System Model

The suggested system has multiple steps. The initial step involves extracting blood samples from an individual. This is done by disinfecting the finger area with alcohol, puncturing the finger with a lancet, and utilizing a transparent or white surface for collecting the blood. Blood group kits are also utilized in this process. During the second step, capturing an image using a mobile phone or a standalone camera is necessary. The image must have a satisfactory resolution and be legible when processed by the program. The ultimate phase employs two techniques to categorize a blood group through a graphical user interface (GUI) in MATLAB and an artificial intelligence in Orange (AI_O) software. In order to identify a blood type and determine the duration of processes, we have devised novel techniques illustrated as a schematic figure in Fig. 1.

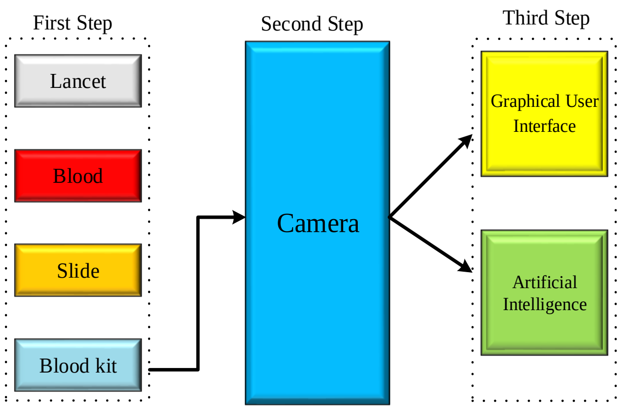


Figure 1: Block diagram for the proposed system

Materials and Methods

Participants and Sampling

The current study obtained ethical approvals from the Iraqi Ministry of Health and Environment – Anbar Health Directorate – the Research Committee at the Training and Human Development Center according to research protocol (2303) on 18/04/2024, relying on the principles stipulated in the Declaration of Helsinki. Participants' information and written consent forms were collected electronically by them. It was agreed to collect blood group samples from patients in Heet General Hospital in accordance with the aforementioned protocol. This study used samples collected in two different ways. One of these methods was collected from volunteer individuals, including various genders and age groups, with approximately 700 samples for a long period of time. The other method was using the web such as Kaggle, Google, and various sites. They were packaged into eight different packages according to Blood type and Rh factor. The total number of samples in this study was 1585. The aim was to enhance the system's accuracy by employing two technologies, namely GUI and AI_O, to categorize blood types based on group and Rh factor. Specimens were obtained via a medical lancet for sampling purposes. Furthermore, three droplets of blood extracted from the donor were carefully deposited onto a pristine white surface to get a distinct and detailed image. In addition, include blood group kits in the droplets based on the presence of antibodies A, B, and D, from right to left, with A on the right and D on the left. Special sticks were used to mix the blood with the drops, and a little time was allocated for the results to manifest. Next, capture a snapshot of three droplets using a mobile phone or a camera. This photograph is a valuable data source, as it is saved on the computer and will be analyzed later.

Experimental Setup

The experimental setup consists of a portable camera or camera, MATLAB program version 2022b, and Orange program version 3.36.1. Implemented a GUI in a MATLAB program to provide a visual interface that includes analytical tools, eleven axes for picture display, and two text fields for displaying results. The analysis tools comprise a push-button feature allowing image insertion, grey and binary conversion, morphological operations, splitting, detection, and blanking out. There are a total of eleven axes, each serving a specific purpose. One axis is dedicated to inserting images, two axes are used for converting photos to grey and binary, five axes are allocated for morphological operations on images, and the remaining three axes are utilized for splitting up images. Two texts are provided for blood classification and processing time, as illustrated in Fig. 2a. The software version is 3.36.1. The flow chart depicted in Fig. 2b comprises two distinct steps. The initial phase encompasses importing photos, embedding images, building models, conducting tests and scoring, and generating a confusion matrix. The suggested system employs multiple models, including Logistic Regression, Neural Network, Tree, KNN, and SVM. The second phase entails the importation of photos, the embedding of images, and the generation of forecasts.

System Overview

The proposed system in the MATLAB program analysis tools comprises seven functions: picture insertion, grey image conversion, binary image conversion, morphological operations, image splitting, object detection, and image blanking. Transferring digital images captured by a camera to a computer, these snapshots were saved to analyze and categorize them based on blood groups and the Rh factor. The images can be in many formats, such as jpg, png, bmp. Upon analysis, the images are converted into a matrix based on the image's color dimensions [24] and the number of rows and columns [25]. This transformation can be mathematically represented by Equations below accordingly:

$$I_{(x,y)} = \begin{vmatrix} \mathrm{I}_{(0,0)}\mathrm{I}_{(0,1)}\mathrm{I}_{(0,2)} \cdots\cdots.\mathrm{I}_{(0,\mathrm{R}-1)} \\ ::::: \\ ::::: \\ \mathrm{I}_{(\mathrm{C}-1,0)}\mathrm{I}_{(\mathrm{C}-1,1)}\mathrm{I}_{(\mathrm{C}-1,2)} \cdots\cdots\mathrm{I}_{(\mathrm{C}-1,\mathrm{R}-2)} \end{vmatrix},$$

$$\mathrm{RGB} = 3 \times \mathrm{R} \times \mathrm{C},$$

where RGB represents the color elements (red, green, and blue) in color images, with a fixed value of 3. C represents a column in the image matrix, whereas R represents a row in the image matrix.

Next, the color image is converted to grayscale and binary, respectively. The purpose of this conversion is that some instructions cannot deal with color images, so they are converted to these gradations to avoid errors in classifying the images, as several instructions were used to convert them to these transformations, such as "imadjust and imbinarize." The morphological unit includes several operations, the first of which is a color inversion, where the image changes from white to black and vice versa. Some granules in the image are combined with their neighboring areas according to a threshold value determined based on several experiments [26, 27]. Furthermore, it is essential to ascertain the boundary of the blood droplet by examining the presence and distribution of granules, whether they originate from a single location or multiple locations. Canny was selected to detect the edges following extensive testing of various methods such as "Sobel, Prewitt, Roberts, log, Zerocross, and Approxcanny". The technique parameter

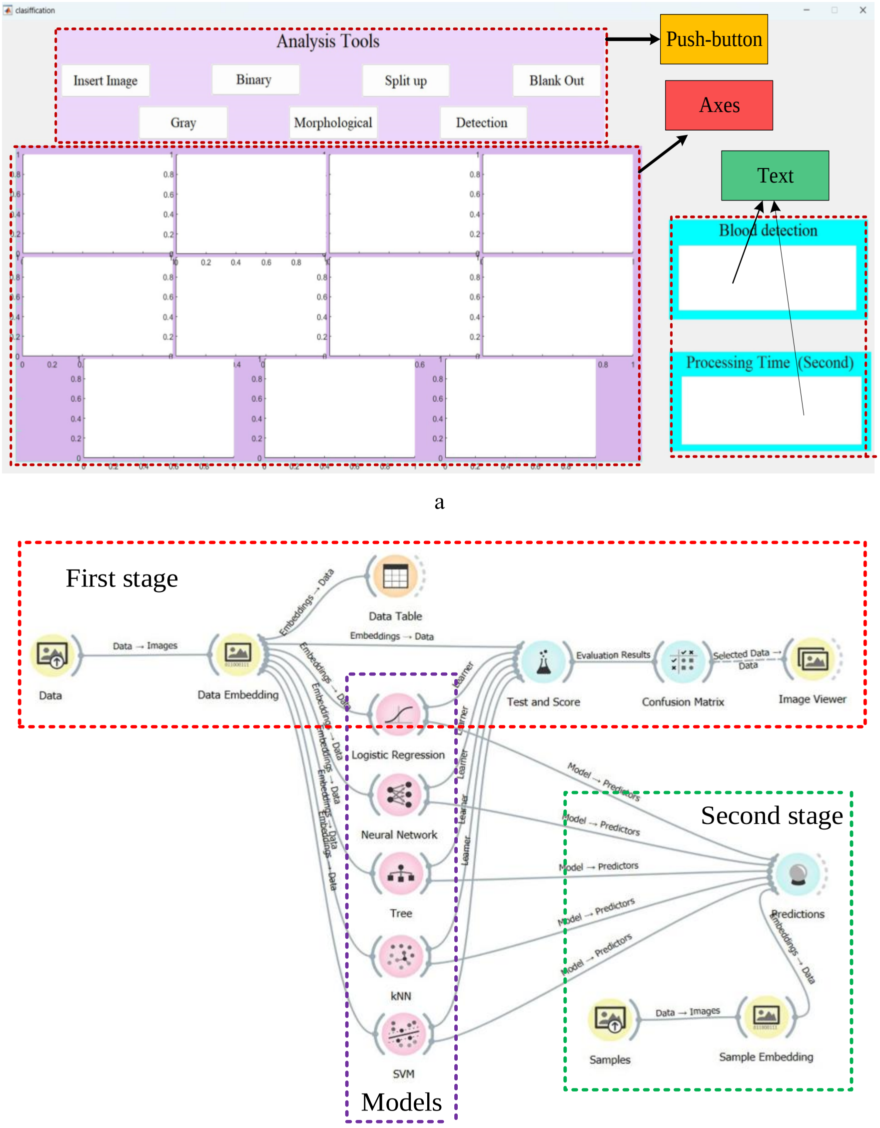


Figure 2: Snapshot of (a) graphical interface in MATLAB program and (b) design flow chart in Orange program

"canny" is shown to be superior to other types of edge detection [28, 29].

Repairing the missing sections in the binary image by identifying the boundaries of the blood droplets was based on the designated regions. This installation employed a specific methodology to fill the holes, as certain regions were inaccessible using conventional methods. Therefore, this alternative approach was utilized to follow the given instructions "imfill" [30]. To support the image and give it more clarity, noise or shapes have been removed far from the area of blood drops or the edges of the frame on which the blood drops are placed, spots have been removed far from these drops and not included in the calculations, and the reflected light have beeb removed from the place where these drops are placed, so this technique was used in terms of using the instruction "mclearborder" [31, 32]. In order to analyze antibodies A, B, and D, the image was partitioned into three sections using the dimensions specified in Equation above. This division allows for identifying blood groups (A, B, AB, and O) in the first two images. The third image is used to determine the group's Rh factor (negative or positive). The segmentation of the images was performed using the "bwlabel" method. The utilization of octal counting has been implemented to minimize the duration of data processing and has gained extensive prevalence [33, 34]. Data is cleared in each GUI to choose a blank-out button.

The proposed system in the Orange program involves two steps. The first step includes importing, image embedding, models, test and score, and confusion matrix. The importing image includes data or images collected from the volunteers and divided according to blood groups and the Rh factor (A+, A-, B+, B-, AB+, AB-, O+, and O-) because they are considered the basis for learning and comparing the results with them, as they were divided into eight groups. This section utilizes image embedding to extract data from images and capture the fundamental characteristics of the image. It is a digital representation of images in a vector with fewer dimensions. The image is compressed to enhance the complexity of its visual content. To optimize the machine learning process and assess visual information's visual and semantic components. The procedure often involves working with fixed-dimensional vectors and is essential in machine learning. Each pixel represents an embedded image, enabling machine learning to comprehend the image by combining pixels, which carry distinct data [35, 36]. Model units in the proposed system used Logistic Regression, Neural Network, Tree, KNN, and SVM models from AI_O to compare them based on train time, test time, classification accuracy (CA), and precision; the mentioned techniques were used because they are considered one of the most critical methods in the classification process. A supervised machine learning approach called logistic regression is often used to classify objects because it is considered a prediction method. This process examines the type and finds a relationship with a group of transactions, dependent or fixed binary variables, and is considered one of the practical tools in decision-making.

The Logistic regression, a form of supervised machine learning, is commonly employed for object classification due to its reputation as a predictive technique. This procedure analyses the category and establishes a correlation with a set of transactions, either dependent or fixed binary variables. It is widely regarded as one of the efficient tools in the decision-making process [37, 38]. The sigmoid function is the function used in logistic regression. It can be represented using Equation below. Which converts any actual values from independent variables into a value between 0-1 to convert the outputs of continuous values of the linear regression function into categorical value outputs. The sigmoid function, utilized in logistic regression, can be mathematically expressed by Equation below. This process aims to transform the actual values of independent variables into a range of 0-1; this is done to convert the continuous output values of the linear regression function into categorical values [39, 40]. It can be represented using Equations below:

$$\sigma(s) = \frac{1}{1+e^{-s}} ,$$

$$s = \left(\sum_{i=1}^{n} wixi \right) + b,$$

$$s = w.x + b,$$

where $\sigma(s)$ acts as the sigmoid function, s acting values, w acting weight, b acting bias, and xi acting input value.

The neural network is one of the artificial intelligence methods for computers to learn how to process information like the human brain. Neural nodes or interconnected neurons are used in a layered structure similar to the human brain, also called machine learning. It can create an adaptive system that computers use to learn from their mistakes and continuously improve the processing process, which is used to solve complex problems in classification [41, 42]. It involves three layers: input, hidden, and output layers. It can be simple or complex regarding the number of hidden layers, called multilayers. Information from the outside world is entered into the network from the input layer, which processes and analyzes the data and passes it to the subsequent layer. The input layer and other hidden layers are considered components of these layers. There can be more than one layer to analyze the outputs of the previous layers. As the hidden layers increase, the processing becomes more accurate and requires significant time to be passed to subsequent layers. The output layer displays the outcome of the artificial neural network's complete data processing [43], as demonstrated in Fig. 3a. Although decision trees are a supervised learning technique, they are primarily used to solve classification problems. It was given this name because the information is organized similarly to a tree. The internal nodes are information attributes, branches of decision rules, and leaf nodes for each outcome [44]. It involves four parts: root node, internal node, branches, and leaves. The decision-making process usually begins at the top nodes of the Tree. These nodes, which typically include an input attribute, represent an option or test condition. Based on the test result, the procedure moves to one of the child nodes. For an internal node, the branches indicate what can happen to a test

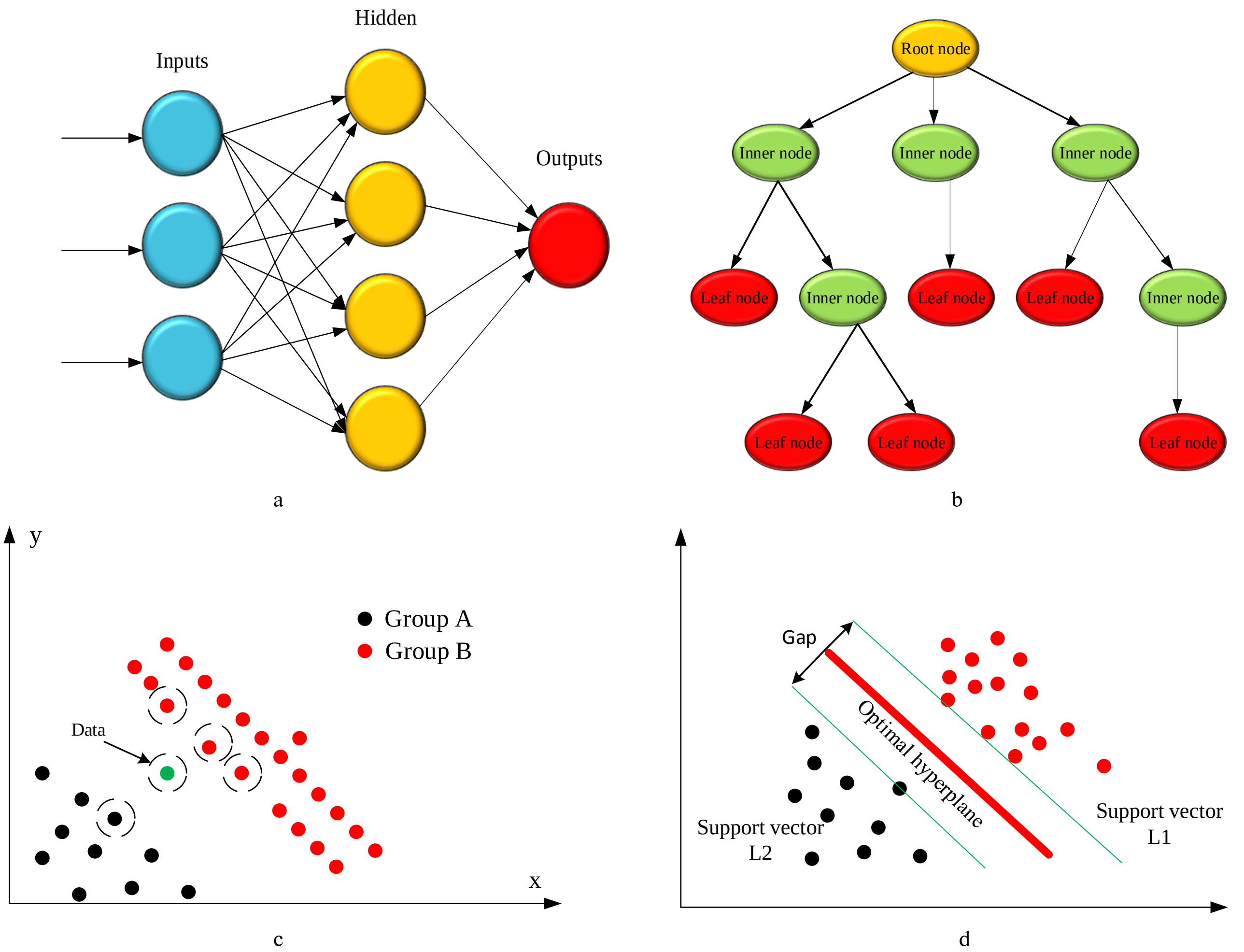


Figure 3: Design using: (a) Neural Network, (b) tree, (c) K-Nearest Neighbors algorithm, and (d) support vector machine algorithm

condition. Each branch is a parent of a child node or another internal node. Leaves stand for the verdict or classification. These represent the decision-making process's conclusions [45], as demonstrated in Fig. 4. KNN is a type of supervised learning method that is used in both regression and classification. The prediction method depends on calculating the distance between the sampled data and summing the training points. The points closest to the sample data are selected through this distance, meaning the category with the highest probability will be selected [46]. The distance for the data in Fig. 5 is closer to group B with the highest probability, so it is classified as group B. The distance is calculated according to Equation below [47]:

$$D_{\text{data}} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\ ,$$

where D_{data} acts as a distance for data, $x_{1,2}$ and $y_{1,2}$ represent the coordinates of the first and second points on an x- and y-axis, respectively.

The definition of a SVM is a machine learning algorithm that determines the boundaries between data points based on predefined classes, labels, or outputs. It uses supervised learning models to solve complex classification, regression, and outlier detection problems [48, 49]. As it depends on the margin or Optimal hyperplane subject to Equation below, which is the distance between the data points or the sports vector and the decision line, the critical thing in this method is to increase the margin distance, as demonstrated in Fig. 6. The upper and lower Optimal hyperplanes are support vectors in L1 and L2 for calculating the margin using Equations below within this framework:

$$w.x + b = 0\ ,$$
$$w.x + b = 1\ ,$$
$$w.x + b = -1\ .$$

where w signifies the weight vector associated with the hyperplane, b represents the bias or offset, and x corresponds to the input vector.

The test and score tests learning algorithms on data-based data input, models, and test data. The output for the classification algorithm. Finally, the confusion matrix displays the ratios of the actual class to the projected class-based test and score, the output chosen data subset from the confusion matrix, and data together with further details on whether a particular data instance was chosen. The second step involves sampling images that were obtained as previously described. The sample embedding is embedding images of samples as shown above, where it analyzes the image into a vector containing a thousand elements to obtain all the information from each image to classify the image according to blood groups and Rh factor, and predictions in this area classify blood groups according to types and Rh factor. It depends on the type of model used in the classification process; the sample images entered through the sample images, and the step of embedding the images.

Experimental Results

The analysis and simulation indicate that the results obtained from the MATLAB and Orange programs are identical. However, the difference is that the MATLAB program classifies one image in each process. In contrast, the Orange program classifies an indefinite number of images in each process in terms of speed in classification. Details will be noted in this section. Simulation results were obtained from computer properties. The device type is MSI. The processor is Intel(R) Core(TM) i7-10750H CPU @ 2.60, 2.59 GHz. The installed RAM is 16.0 GB. The System type is the 64-bit operating system.

MATLAB Results

This method relies on counting the granules in antibodies A and B to determine the blood type among types A, B, O, and AB. To differentiate between Rh factor negative and positive, consult Fig. 4 to quantify granules determined by antibody D.

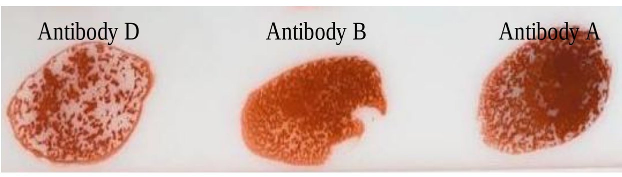


Figure 4: Snapshot of a blood drop-based Antibody

The granules were counted using the MATLAB program for more than one sample. A statistical table was made to calculate all antibodies A, B, and D and create an algorithm for them according to the antibodies A and B for blood types and antibodies D for the Rh factor. The final processing time was calculated from the start of inserting the image to the process of determining the blood type using the tic-toc prompt in the MATLAB program. The experiment was conducted on approximately 100 samples to arrive at the algorithm applied in the program, and the accuracy was approximately 100%, expressing Equation [50]

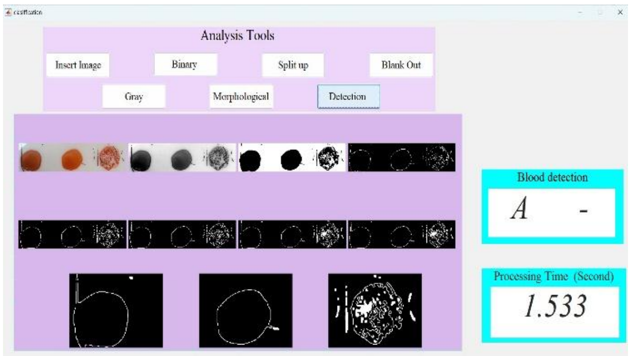
$$\text{Accuracy}\ \% = \frac{\frac{1}{S}\sum_{i=1}^{s} x_{\text{h\_e}}}{\frac{1}{S}\sum_{i=1}^{s} y_{\text{m\_s}}} \times 100\%\,,$$

where s is the total number of samples, $x_{\text{h-e}}$, and $y_{\text{m-s}}$ acting reading sampling by human eyes and reading sampling by MATLAB-system, respectively. In addition, the time for processing one sample is approximately 1.5–1.6 seconds.

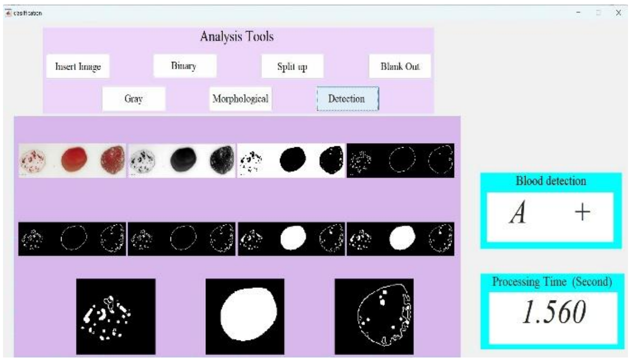
Fig. 5 shows that blood group detection and processing time in GUI were as follows: The blood group and processing time for A negative and positive are 1.53 and 1.56 seconds, as shown in Fig. 5 a and b, respectively. The blood group and processing time for AB negative and positive was about 1.54 seconds, as shown in Fig. 5 c and d, respectively. The blood group and processing time for B negative and positive are 1.53 and 1.56 seconds, as shown in Fig. 5 e and f, respectively. The blood group and processing time for O negative and positive are 1.6 and 1.53 seconds, as shown in Fig. 5 g and h, respectively.

Orange Results

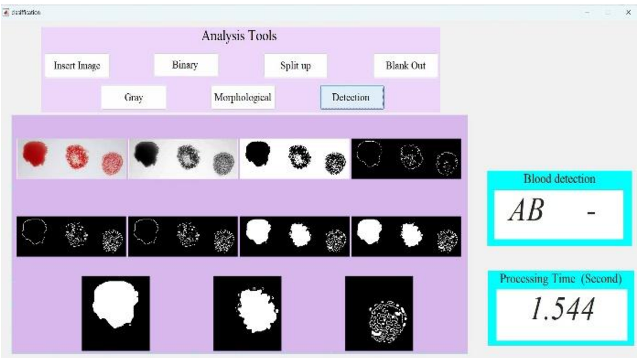
The proposed system was tested on 30 samples of different species. Some samples were manipulated to see if the proposed system could classify species. One of the methods of manipulation was to change the name of the image in the sample, expose it to a temperature to dry the blood drops, and change the light intensity in the samples, which ranges from light to dark in some samples. The proposed system was trained on 1585 data sets using artificial intelligence for five methods: Logistic Regression, Neural Network, Tree, KNN, and SVM. The data was captured using a phone and regular camera and divided into eight groups classified according to blood type and rheumatic factor. The samples were identified according to the type of Rh factor using a name for the image, for example, blood type A+. The image was named with this name, and the sequence of similar images was followed by A+1, etc.



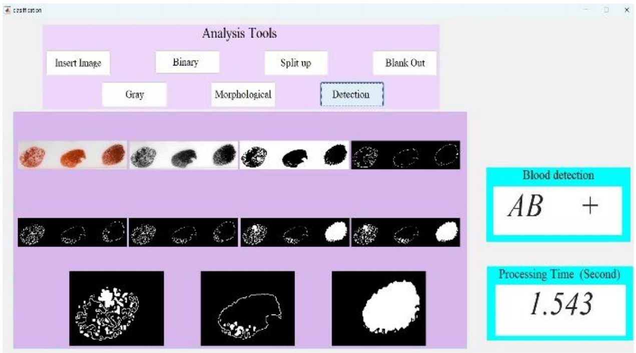
a



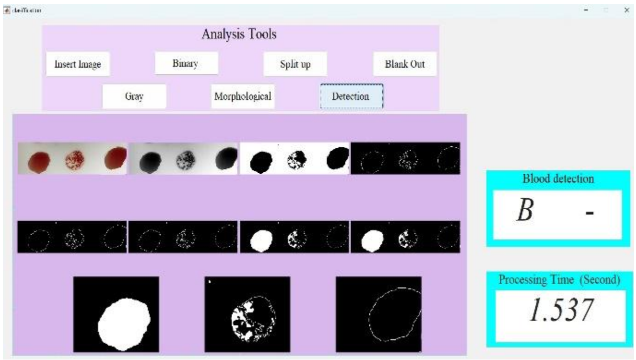
b



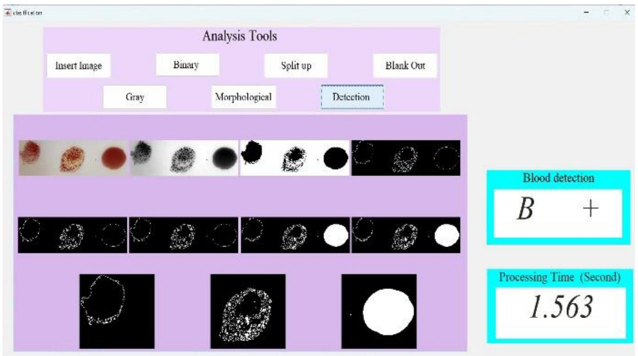
c



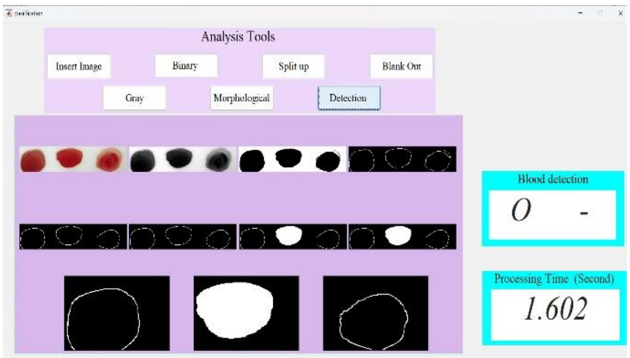
d



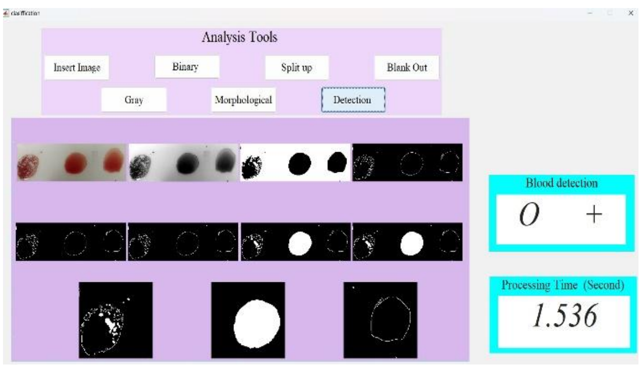
e



f



g



h

Figure 5: Snapshot of results via GUI: (a) A-, (b) A+, (c) AB-, (d) AB+, (e) B-, (f) B+, (g) O-, and (h) O+

The data extracted from the confusion matrix is based on prediction and truth which in turn shows each of the following points: true positive point (*TP*), false negative point (*FN*), false positive point (*FP*) and true negative point (*TN*) shown in Fig. 6, from which the accuracy is calculated according to Equation

$$\text{Accuracy}\,(\%) = \frac{TN + TP}{TP + FP + TN + FN}\,.$$

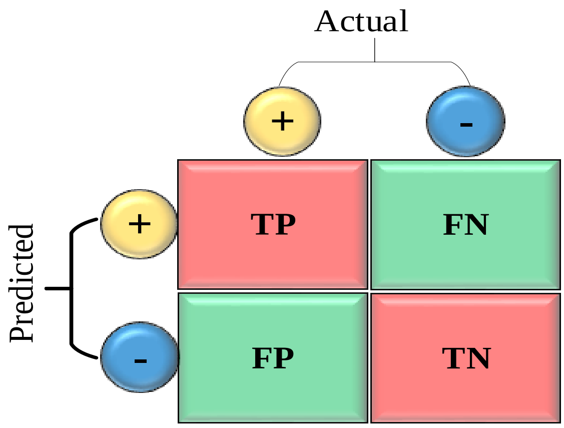


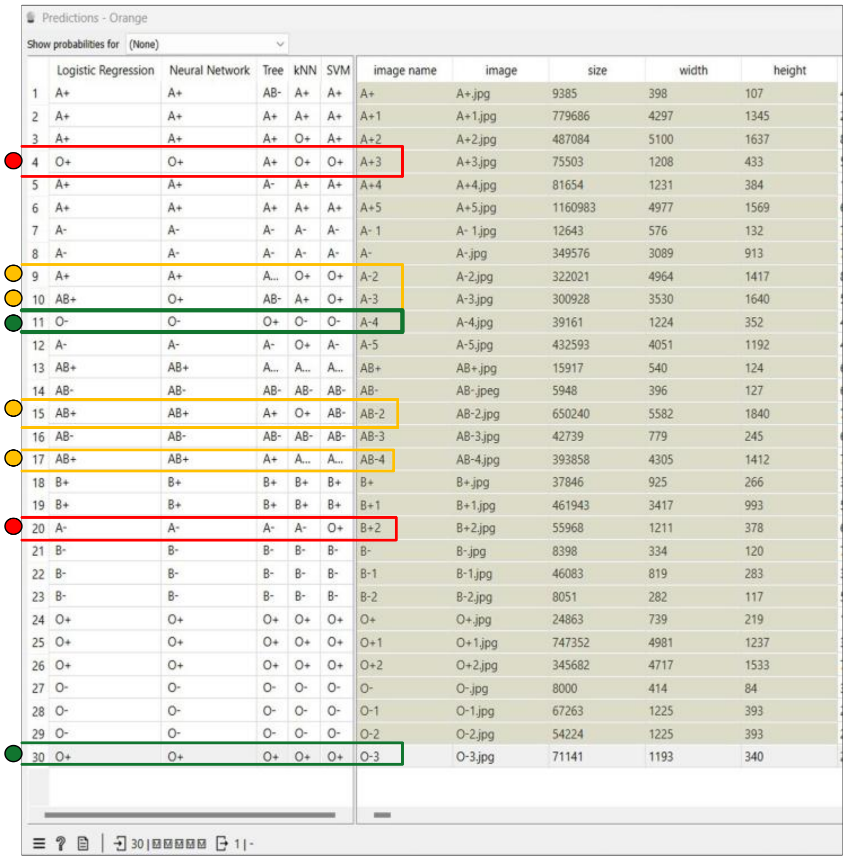
Figure 6: Clarification confusion matrix based-prediction and actual

The training data obtained from the confusion matrix in the software used for and the match rate of each method are shown in Fig. 7. Fig. 7 a and b observed a merging in the training data between AB+ and O+ and vice versa due to some data being close to each other in Logistic Regression and Neural Networks models, respectively. In addition, Fig. 7 c and d observed a merging in training data for different blood groups and Rh factors due to the closeness of some data to each other in the Tree and KNN models. Fig. 7 e observed a merging in the training data between AB+, O+, and A+ and vice versa due to some data being close to each other in the SVM model. The Machine learning methods used in the proposed system are Logistic Regression, Neural Network, Tree, KNN, and SVM to classify a blood group and Rh-factors A+, A-, B+, B-, AB+, AB-, O+ and O-. The setting for logistic regression is parameters as regularization in ridge type and strength at 1. In addition, the setting for neural network is parameters hidden layers are 100, activation functions is ReLu, solver function was Adam, and regularization about 0.0001. More, the setting for KNN is parameters NO. of neighbors is 5, metric is Euclidean, and weight is uniform. Furthermore, the setting for SVM is parameters SVM type was SVM with cost was 1 and regression loss epsilion is 0.1, with the kernel function is RBF. Moreover, the setting for Tree is Parameters Min. number of instances in leaves is 2, Do not split subsets smaller than is 5, limit the maximal tree depth to 100, and Classification Stop when majority reaches is 95%.

Fig. 8 displays the results obtained from the prediction in the software used. Where the lighting intensity changed, the samples marked with red could not be classified correctly for several reasons, the most important of which was that they were not trained on similar models. Light affects the pixel value of each location in the image. In addition, the samples marked in yellow were changed from the shape of drops using heat on the blood and were not recognized by the software or the human eye. Moreover, in the green samples, the name of the image was changed, as the result was identical to the classification in terms of species and rhizome, except for the tree method, which could not classify one of the samples. Furthermore, compare the results based on train time, test time, classification accuracy, and precision for each model, as shown in Fig. 9. For the time train in seconds, Logistic Regression achieved 43 seconds, the highest, and KNN achieved 6 seconds, the lowest. In addition, in test time in seconds, SVM achieved 2 seconds, the highest, and Tree was the lowest. Moreover, the Neural Network achieved 99.7 %, the highest in classification average accuracy and precision, and Tree achieved 98.2 %, the lowest. Table 1 explains the accuracy value for each blood group based on the machine learning method.

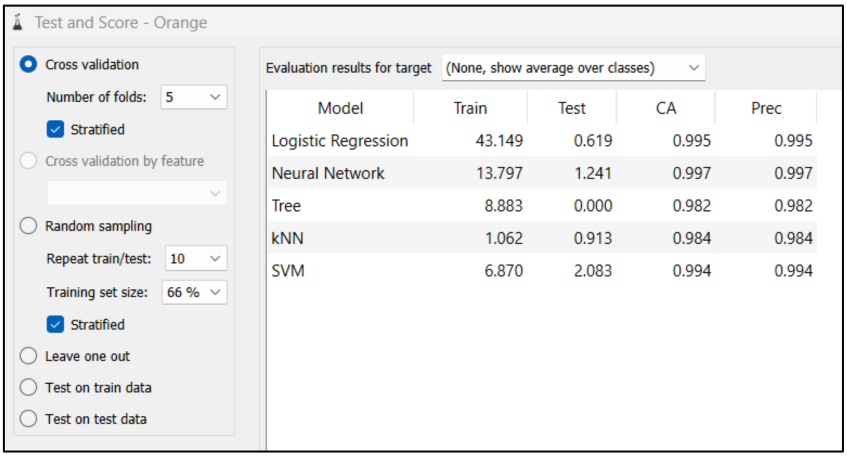
Table 1: Calculating the accuracy of the blood groups for machine learning

Blood groups	Accuracy in (%)				
	Logistic Regression	Neural Network	Tree	KNN	SVM
A+	100	100	99.6	99.7	99.7
A-	100	100	99.6	99.7	100
B+	100	100	100	99.7	100
B-	100	100	100	99.5	100
AB+	99.5	99.7	98.7	99.4	99.4
AB-	100	100	99.6	99.7	100
O+	99.5	99.7	99.2	99.1	99.6
O-	100	100	99.8	99.7	100



	Logistic Regression	Neural Network	Tree	kNN	SVM	image name	image	size	width	height
1	A+	A+	AB-	A+	A+	A+	A+.jpg	9385	398	107
2	A+	A+	A+	A+	A+	A+1	A+1.jpg	779686	4297	1345
3	A+	A+	A+	O+	A+	A+2	A+2.jpg	487084	5100	1637
4	O+	O+	A+	O+	O+	A+3	A+3.jpg	75503	1208	433
5	A+	A+	A-	A+	A+	A+4	A+4.jpg	81654	1231	384
6	A+	A+	A+	A+	A+	A+5	A+5.jpg	1160983	4977	1569
7	A-	A-	A-	A-	A-	A- 1	A- 1.jpg	12643	576	132
8	A-	A-	A-	A-	A-	A-	A-.jpg	349576	3089	913
9	A+	A+	A...	O+	O+	A-2	A-2.jpg	322021	4964	1417
10	AB+	O+	AB-	A+	O+	A-3	A-3.jpg	300928	3530	1640
11	O-	O-	O+	O-	O-	A-4	A-4.jpg	39161	1224	352
12	A-	A-	A-	O+	A-	A-5	A-5.jpg	432593	4051	1192
13	AB+	AB+	A...	A...	A...	AB+	AB+.jpg	15917	540	124
14	AB-	AB-	AB-	AB-	AB-	AB-	AB-.jpeg	5948	396	127
15	AB+	AB+	A+	O+	AB-	AB-2	AB-2.jpg	650240	5582	1840
16	AB-	AB-	AB-	AB-	AB-	AB-3	AB-3.jpg	42739	779	245
17	AB+	AB+	A+	A...	A...	AB-4	AB-4.jpg	393858	4305	1412
18	B+	B+	B+	B+	B+	B+	B+.jpg	37846	925	266
19	B+	B+	B+	B+	B+	B+1	B+1.jpg	461943	3417	993
20	A-	A-	A-	A-	O+	B+2	B+2.jpg	55968	1211	378
21	B-	B-	B-	B-	B-	B-	B-.jpg	8398	334	120
22	B-	B-	B-	B-	B-	B-1	B-1.jpg	46083	819	283
23	B-	B-	B-	B-	B-	B-2	B-2.jpg	8051	282	117
24	O+	O+	O+	O+	O+	O+	O+.jpg	24863	739	219
25	O+	O+	O+	O+	O+	O+1	O+1.jpg	747352	4981	1237
26	O+	O+	O+	O+	O+	O+2	O+2.jpg	345682	4717	1533
27	O-	O-	O-	O-	O-	O-	O-.jpg	8000	414	84
28	O-	O-	O-	O-	O-	O-1	O-1.jpg	67263	1225	393
29	O-	O-	O-	O-	O-	O-2	O-2.jpg	54224	1225	393
30	O+	O+	O+	O+	O+	O-3	O-3.jpg	71141	1193	340

Figure 8: Snapshot of the results obtained from the prediction



Model	Train	Test	CA	Prec
Logistic Regression	43.149	0.619	0.995	0.995
Neural Network	13.797	1.241	0.997	0.997
Tree	8.883	0.000	0.982	0.982
kNN	1.062	0.913	0.984	0.984
SVM	6.870	2.083	0.994	0.994

Figure 9: Snapshot of the results obtained from the test and score

Discussion

Knowing your blood type is important for the purpose of giving or receiving blood. Blood type O is considered a universal donor, as a person with this type can donate blood to all types, taking into account the Rh factor, and only receives blood from the same type as his. AB is a universal recipient, as it can receive blood from all types, taking into account the Rh factor. For one antibody to react with another antigen, the heterologous antigens must be similar or closely related in structure and be found in unrelated plants or animals. Human blood group A and B antigens, which are connected to bacterial polysaccharides, are an illustration of this. It is thought that the initial formation of anti-A antibody occurs upon exposure to pneumococci or other related bacteria. Anti-A antibody is typically detected in people with blood types other than A, such as type B and type O. After coming into contact with a comparable bacterial cell wall component, naturally occurring anti-B antibody is produced. When choosing the appropriate blood type for transfusion, naturally occurring antibodies play a significant role.

There is another method, without the need to use a special kit, to detect blood types using immunochromatography, as this method is similar to a pregnancy test. However, this method is not currently approved in hospitals and centers because it has some limitations in its use, such as the sample volume needs are reduced before reaching the microliter level, it is difficult to duplicate many markers at one time for analysis, and some systems have concerns about sensitivity, and repeat testing is difficult.

The results were discussed for detecting and classifying blood groups and Rh factor using MATLAB and Orange programs, respectively. The current study can be used in medical laboratories, hospitals, and health centers, without any intervention by health laboratory technicians. The mechanism is through an advertising poster or an explanatory video, from the beginning of the finger prick to the display of the result, which is done by the patient or his companion. Image processors and artificial intelligence were used to detect and classify blood types due to the speed of classification and detection of the type with high accuracy, as dozens of samples can be classified at one time compared to the traditional process that requires time and sometimes results in human errors.

Previous studies were summarized using the MATLAB program to detect the blood group and artificial intelligence in classifying the blood group using several methods such as Logistic Regression, Neural Network, Tree, KNN, and SVM. 100 samples were used, which were detected using the MATLAB program. The accuracy was very high, 100%, with a time ranging from 1.5–1.6. The results were superior to previous studies in terms of accuracy and processing time. The system was superior at accuracy 100% wherever the accuracy in [11–15, 21] was 98%, 90%, 98%, 96%, 90%, and 100% in Image Possessing method, respectively. In addition, the system was superior at processing time 1.5-1.6 sec wherever the processing time in [14, 15] was 16 sec and 4–6 sec in Image Possessing method, respectively, as shown in Table 2. 1585 samples were trained using the Orange program to classify blood types and Rh factors within eight features. The system was tested on 30 samples using methods Logistic Regression, Neural Network, Tree, KNN, and SVM and variables were calculated such as accuracy, precision, learning time and test time, as the current system was superior to previous studies in these variables. The proposed system were superior at accuracy 99.4% wherever the accuracy in [21] was 97% in SVM method. In addition, the system were superior at accuracy 98.4% wherever the accuracy in [21] was 97% in KNN method. Moreover, The system were superior at accuracy 99.7% wherever the accuracy in [20] was 95.3% in Neural Network method. Using machine learning, different methods were used to compare the results obtained and arrive at the ideal method among the methods mentioned in classifying blood types based on accuracy and time. In addition, the parameters were compared with previous studies, as shown in Table 3. The MATLAB program was used in the image processing technique, as the processing process was simple as shown. In addition, it processes individual images to determine the blood type and the Rh factor, as it makes only one diagnosis after performing the process of mixing the blood drops with the kit for blood groups, taking a picture and inserting it in the program. The data is saved to be retrieved at a later time, and the mobile phone can be used to perform the processing. Moreover, it avoids human error in reading blood types. Furthermore, anyone can take this test after reading the test instructions.

Table 2: Compared between current study and previous studies in accuracy and possessing time

References / Years	Methods	Accuracy (%)	Processing Time, seconds
[11] / 2022	Image Possessing	98	N/A
[12] / 2019	Image Possessing	90	N/A
[13] / 2017	Image Possessing	98	N/A
[14] / 2019	Image Possessing	96	16
[15] / 2016	Image Possessing	90	4-6
[21] / 2022	Image Possessing	100	N/A
Proposed System	**Image Possessing**	**100**	**1.5-1.6**

N/A: not available

Table 3: Compared between current study and previous studies in techniques, accuracy, training time, and testing time

References / Years	Methods	Techniques	Accuracy (%)	Training Time, seconds	Testing Time, seconds
[20] / 2021	Machine learning	Neural Network	95.3	N/A	N/A
[21] / 2022	Machine learning	SVM and Tree	97	N/A	N/A
[22] / 2018	Machine learning	SVM	N/A	N/A	N/A
Proposed System	**Machine learning**	**Logistic Regression**	**99.5**	**43.149**	**0.619**
		Neural Network	**99.7**	**13.797**	**1.241**
		Tree	**98.2**	**8.883**	**0**
		KNN	**98.4**	**1.062**	**0.913**
		SVM	**99.4**	**6.87**	**2.083**

N/A: not available

The Orange program was used to conduct machine learning because it is easy to understand and quick to apply. It is considered a fairly new technology and the tools are free. In addition, the processing time in this program is faster than MATLAB as shown in previous studies. Installing it on the computer does not take a long time, and it is possible to find more than one blood type at a time. There were several obstacles to using the MATLAB program, as it requires a computer with very good specifications. In addition, some tools require sums of money to activate them, and it requires a modern version of the program. There are some limitations that affect all the parameters of this study that can be avoided or considered in the future, including the quality of images, the amount of lighting, the dimensions of the images, taking a picture at the appropriate time before the blood samples dry, and the type of surface on which the samples were mixed. In addition, place the blood group kit in the correct place as shown in Fig. 4. The method of mixing blood drops must be from the inside out. To determine the outer frame for sampling, it is preferable that the background of the captured images be white so as not to affect any reading.

Conclusions

This study focuses on the urgent need to accurately determine blood type before emergency blood transfusions, which rely on manual examination and are prone to human errors. These errors put people's lives at risk during blood transfusions. The main goal of the research was to develop an approach that combines image processing and machine learning to detect and classify blood types with high accuracy and short time. The proposed system includes two parts: detection and classification of blood types using two programs, MATLAB and Orange. The MATLAB program was used to detect blood types, which performs image analysis. The other program used machine learning techniques such as logistic regression, neural networks, decision trees, KNN, and SVM to classify blood types. The results indicate that using 100 samples in MATLAB, each detection runs reaches a 100% accuracy rate within 1.5-1.6 seconds. In addition, the results conducted on the Orange program, by training 1585 samples to test 30 samples, show that neural networks have exceptional accuracy compared to other techniques, as they achieved an accuracy rate of up to 99.7%. The training lasts 13.7 seconds, while the test takes 1.2 seconds. Improving the proposed method for detecting blood type and treating problems with a very high accuracy of up to 100% and a short time for practical use should be a top priority in future studies. In addition, using video technology instead of photos. Furthermore, use techniques that do not depend on the blood group kit, such as the technique of detecting types using a fingerprint or other methods.

Acknowledgements

The author would like to express their gratitude to the Electrical Engineering Technical College at Middle Technical University in Baghdad, Iraq, for assisting them in carrying out the experiments.

Interest Disclosure

The author has no conflicts of interest to declare.

References

[1] Skull A. Circulation - anatomy, physiology and pathophysiology. In: McGloin S, Skull A. Principles of acute care nursing. SAGE Publications; 2022. p. 86.

[2] Leslie HA, van Velzen MJM, Brandsma SH, Vethaak AD, Garcia-Vallejo JJ, Lamoree MH. Discovery and quantification of plastic particle pollution in human blood. Environ Int. 2022 May;163:107199. DOI: 10.1016/j.envint.2022.107199

[3] Asadpour M, Olsen TL, Boyer O. An updated review on blood supply chain quantitative models: A disaster perspective. Transport Res Part E Logistics Transport Rev. 2022;158:102583. DOI: 10.1016/j.tre.2021.102583

[4] Pei Q, Luo Y, Chen Y, Li J, Xie D, Ye T. Artificial intelligence in clinical applications for lung cancer: diagnosis, treatment and prognosis. Clin Chem Lab Med. 2022 Jun 30;60(12):1974-1983. DOI: 10.1515/cclm-2022-0291

[5] Sameer HA, Mutlag AH, Gharghan SK. CT-scan method-based artificial neural network for diagnosis of COVID-19. J Techniq. 2022;4(4):24-32. DOI: 10.51173/jt.v4i4.701

[6] Ferraz A, Moreira V, Silva D, Carvalho VH, Soares FO, Automatic system for blood type classification using image processing techniques. In: BIODEVICES 2011 - Proceedings of the International Conference on Biomedical Electronics and Devices; 2011 Jan 6-29; Rome. p. 368-73.

[7] Dhande A, Bhoir P, Gade V. Identifying the blood group using Image Processing. Int Res J Eng Technol. 2018;5:2639-42.

[8] Rahman S, Rahman MA, Khan FA, Shahjahan SB, Nahar K. Blood group detection using image processing techniques. BRAC University; 2017. 37 p.

[9] Vatshav S, Karthik K, Ch A, Kumar Patnaik MS. Determination and classification of blood types using image processing techniques. J Maharaja Sayajirao Univer Baroda. 2021;55(2):44-9.

[10] Ravindran G, Joby T, Pravin M, Pandiyan P. Determination and classification of blood types using image processing techniques. Int J Comp Appl. 2017;157(1):12-6. DOI: 10.5120/ijca2017912592

[11] Shaban S, Elsheweikh D. Blood group classification system based on image processing techniques. Intell Automat Soft Comput. 2022;31(2):817-34. DOI: 10.32604/iasc.2022.019500

[12] Pavithra V, Rajeshwari J. Identification of blood group by using image processing techniques. J Eng Res Manag. 2019;3(5):90-5.

[13] Yamin A, Imran F, Akbar U, Tanvir SH. Image processing based detection & classification of blood group using color images. In: Proceedings of 2017 International Conference on Communication, Computing and Digital Systems (C-CODE); 2017 Mar 8-9; Islamabad, Pakistan. Piscataway, NJ: IEEE; 2017. p. 293-8. DOI: 10.1109/C-CODE.2017.7918945

[14] Jamil MMA, Oussama L, Hafizah WM, Abd Wahab MH, Johan MF. Computational automated system for red blood cell detection and segmentation. In: Intelligent data analysis for biomedical applications. Amsterdam: Elsevier; 2019. p. 173-89. DOI: 10.1016/B978-0-12-815553-0.00008-2

[15] Sahastrabuddhe AP, Ajij S. Blood group detection and RBC, WBC counting: an image processing approach. Int J Eng Comput Sci. 2016;5(10):18635-9. DOI: 10.18535/ijecs/v5i10.49

[16] Paul Sathiyan S, Glady Jennifer KS, Swathi S, Mariya Sharmini G. Blood group determination and classification using Raspberry Pi3. In: Proceedings of AIP Conference. 2022 Apr 15. DOI: 10.1063/5.0003932

[17] Ferraz A, Brito JH, Carvalho V, Machado J. Blood type classification using computer vision and machine learning. Neural Comput Appl. 2017;28(8):2029-40. DOI: 10.1007/s00521-015-2151-1

[18] Gurav GV, Patil S. Automatic blood group classification based on SVM. Int J Sci Res Dev. 2017;5(8):317-21.

[19] Raj Bhagat SM, Deshmukh P, Khanvilkar S. Automatic detection of human blood group system using deep learning and image processing. Int Res J Eng Technol. 2021;8(4):350-4.

[20] Odeh N, Toma A, Mohammed F, Dama Y, Oshaibi F, Shaar M. An efficient system for automatic blood type determination based on image matching techniques. Appl Sci. 2021;11(11):5225. DOI: 10.3390/app11115225

[21] Rosales MA, de Luna RG. Computer-based blood type identification using image processing and machine learning algorithm. J Adv Comput Intell Intell Inform. 2022;26(5):698-705.

[22] Vadgave AA, Vadgave AA, Ashtekar AS, Langade P. Determination of blood group using image processing. Int Res J Moderniz Eng Technol Sci. 2023;5(6):3172-8.

[23] Dada A, Beck D, Schmitz G. Automation and data processing in blood banking using the Ortho AutoVue® Innova System. Transfus Med Hemother. 2007;34(5):341-6. DOI: 10.1159/000106558

[24] Sharma S, Sharma H, Sharma JB. A new optimization based color image watermarking using non-negative matrix factorization in discrete cosine transform domain. J Ambient Intell Human Comput. 2022;13:4297-319. DOI: 10.1007/s12652-021-03408-1

[25] Han ZY, Wang J, Fan H, Wang L, Zhang P. Unsupervised generative modeling using matrix product states. Phys Rev X. 2018 Jul;8(3):031012. DOI: 10.1103/PhysRevX.8.031012

[26] Heriansyah R, Utomo WM. Performance evaluation of digital image processing by using Scilab. JUITA Jurnal Informatika. 2021;9(2):239-47. DOI: 10.30595/juita.v9i2.8434.

[27] Abdulazeez AM, Zeebaree DQ, Zebari DA, Hameed TH. Leaf identification based on shape, color, texture and vines using probabilistic neural network. Computaciyn y Sistemas. 2021;25(3):617-31. DOI: 10.13053/cys-25-3-3470

[28] Lynn N, Sourav A, Santoso A. Implementation of real-time edge detection using Canny and Sobel algorithms. In: IOP Conference Series: Materials Science and Engineering, Volume 1096, The 6th International Conference on Industrial, Mechanical, Electrical and Chemical Engineering; 2020 Oct 20; Solo, Indonesia. DOI: 10.1088/1757-899X/1096/1/012079

[29] Makandar A, Kaman S, Biradar R, Javeriya SB. Impact of edge detection algorithms on different types of images using PSNR and MSE. LC Int J STEM. 2022;3(4):1-11. DOI: 10.5281/zenodo.7607059

[30] Krestanova A, Kubicek J, Penhaker M, Timkovic J. Premature infant blood vessel segmentation of retinal images based on hybrid method for the determination of tortuosity. Lékař a technika-Clinician and Technology. 2020;50(2):49-57. DOI: 10.14311/CTJ.2020.2.02

[31] Ullah B, Khan A, Fahad M, Alam M, Noor A, Saleem U, Kamran M. A novel approach to enhance dual-energy X-ray images using region of interest and discrete wavelet transform. J Inf Process Syst. 2022;18(3):319-31. DOI: 10.3745/JIPS.01.0086

[32] Li Y, Liu H, Tian Z, Geng W. Near-infrared vascular image segmentation using improved level set method. Infrared Phys Technol. 2023;131:104678. DOI: 10.1016/j.infrared.2023.104678

[33] Wang L, Gu X, Liu Z, Wu W, Wang D. Automatic detection of asphalt pavement thickness: A method combining GPR images and improved Canny algorithm. Measurement. 2022;196:111248. DOI: 10.1016/j.measurement.2022.111248

[34] Kowalczyk M, Ciarach P, Przewlocka-Rus D, Szolc H, Kryjak T. Real-time FPGA implementation of parallel connected component labelling for a 4K video stream. J Signal Process Syst. 2021;93(5):481-98. DOI: 10.1007/s11265-021-01636-4

[35] Miikkulainen R, Liang J, Meyerson E, Rawal A, Fink D, Francon O, et al. Evolving deep neural networks. In: Kozma R, Alippi C, Choe Y, Morabito FC, editors. Artificial intelligence in the age of neural networks and brain computing. Academic Press; 2019. p. 293-312. DOI: 10.1016/B978-0-12-815480-9.00015-3

[36] Burlina P, Paul W, Liu TA, Bressler NM. Detecting anomalies in retinal diseases using generative, discriminative, and self-supervised deep learning. JAMA Ophthalmol. 2022;140(2):185-9. DOI: 10.1001/jamaophthalmol.2021.5557

[37] Shaikh F, Rao D. Prediction of cancer disease using machine learning approach. Mater Today Proc. 2022;50(Part 1):40-7. DOI: 10.1016/j.matpr.2021.03.625

[38] Choudhary G, Singh SN. Prediction of heart disease using machine learning algorithms. In: Proceedings of 2020 International Conference on Smart Technologies in Computing, Electrical and Electronics; 2020 Mar 19-21; Bengaluru, India. p. 197-202. DOI: 10.1109/ICSTCEE49637.2020.9276802

[39] Singh P. Logistic regression. In: Machine learning with PySpark: With natural language processing and recommender systems. Apress Berkeley; 2022. p. 75-103. DOI: 10.1007/978-1-4842-7777-5_5

[40] Zaidi A, Al Luhayb ASM. Two statistical approaches to justify the use of the logistic function in binary logistic regression. Math Probl Eng. 2023;2023. DOI: 10.1155/2023/5525675

[41] Li X. Artificial intelligence neural network based on intelligent diagnosis. J Ambient Intell Human Comput. 2021;12:923-931. DOI: 10.1007/s12652-020-02108-6

[42] Gholami V, Khaleghi MR, Pirasteh S, Booij MJ. Comparison of self-organizing map, artificial neural network, and co-active neuro-fuzzy inference system methods in simulating groundwater quality: Geospatial artificial intelligence. Water Resour Manag. 2022;36(2):451-69. DOI: 10.1007/s11269-021-02969-2

[43] Sinha K, Uddin Z, Kawsar H, Islam S, Deen M, Howlader M. Analyzing chronic disease biomarkers using electrochemical sensors and artificial neural networks. TrAC Trends Anal Chem. 2023;158:116861. DOI: 10.1016/j.trac.2022.116861

[44] Mustapha MT, Ozsahin DU, Ozsahin I, Uzun B. Breast cancer screening based on supervised learning and multi-criteria decision-making. Diagnostics. 2022;12(6):1326. DOI: 10.3390/diagnostics12061326

[45] Li J, Huang Q, Ren S, Jiang L, Deng B, Qin Y. A novel medical text classification model with Kalman filter for clinical decision making. Biomed Signal Process Control. 2023;82:104503. DOI: 10.1016/j.bspc.2022.104503.

[46] Khotimah BK. Performance of the K-nearest neighbors method on identification of maize plant nutrients. J Infotel. 2022;14(1):8-14. DOI: 10.20895/infotel.v14i1.735

[47] Uddin S, Haque I, Lu H, Moni MA, Gide E. Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. Sci Rep. 2022;12(1):6256. DOI: 10.1038/s41598-022-10358-x

[48] Jo T. Support vector machine. In: Machine learning foundations: supervised, unsupervised, and advanced learning. Springer Cham; 2021. p. 167-88. DOI: 10.1007/978-3-030-65900-4_8

[49] Montesinos López OA, Montesinos López A, Crossa J. Support vector machines and support vector regression. In: Multivariate statistical machine learning methods for genomic prediction. Springer, Cham; 2022. DOI: 10.1007/978-3-030-89010-0_9

[50] Mahmood MF, Mohammed SL, Gharghan SK, Zubaidi SL. Wireless power transfer based on spiral-spider coils for a wireless heart rate sensor. In: Proceedings of 2020 13th International Conference on Developments in eSystems Engineering (DeSE); 2020 Sept 7-9; Liverpool, United Kingdom. p. 183-8. DOI: 10.1109/DeSE51703.2020.9450775

М.Ф. Махмуд

Технічний коледж електротехніки Середнього технічного університету, Багдад, Ірак.

РОЗПІЗНАВАННЯ ТА КЛАСИФІКАЦІЯ ГРУП КРОВІ ЗА ДОПОМОГОЮ МЕТОДІВ МАШИННОГО НАВЧАННЯ ТА ОБРОБКИ ЗОБРАЖЕНЬ

Проблематика. Еритроцити є одним із компонентів крові. Кров є важливою рідиною в організмі людини. Знання груп крові має важливе значення під час переливання крові, яке здійснюється в контрольованих умовах, що забезпечує уникнення фатальних помилок. Для визначення груп крові використовується традиційний метод, коли задіяні медичні лаборанти, тому можуть мати місце помилки, викликані людським фактором.
Мета. Ми маємо на меті розробити й реалізувати прототип для виявлення та класифікації груп крові, що дасть змогу уникнути помилок, викликаних людським фактором, при визначенні групи крові. Запропонована система використовує обробку зображень і алгоритми машинного навчання для виявлення та класифікації груп крові.
Методика реалізації. Система складається з трьох етапів. На першому ми збираємо зразки у волонтерів. На другому використовуємо камеру, щоб зробити зображення для зразка. На третьому аналізуємо зображення двома методами: обробкою зображень за допомогою Matlab і застосовуючи алгоритми машинного навчання за допомогою Orange для виявлення та класифікації груп крові.
Результати. Точність обробки зображень за допомогою програми MATLAB досягала 100 %, а час обробки становив 1,5–1,6 с. Крім того, у програмі Orange з використанням машинного навчання нейронними мережами точність досягла 99,7 %, час навчання становив 13,7 с, а час тестування – 1,2 с. Також, як показали результати експерименту, нейронні мережі були кращими за інші модельні блоки. Наші результати показали, що автоматичне визначення групи крові за допомогою методів обробки зображень і машинного навчання є ефективним і виправданим порівняно з ручним методом. Запропонована система перевершила попередні дослідження щодо точності, часу обробки, часу навчання та часу тестування завдяки використанню двох методів.
Висновки. Це дослідження зосереджено на нагальній потребі в точному визначенні групи крові перед екстреним переливанням крові, яке базується на ручній перевірці та чутливе до помилок, викликаних людським фактором. Ці помилки потенційно можуть загрожувати життю під час переливання крові. Основною метою дослідження було розробити підхід, який поєднує обробку зображень і машинне навчання, для точної класифікації груп крові.

Ключові слова: класифікація груп крові; MATLAB-обробка зображень; машинне навчання Orange; час обробки.