Süleyman Demirel Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 29, Sayı 2, 268-277, 2025

Süleyman Demirel University
Journal of Natural and Applied Sciences
Volume 29, Issue 2, 268-277, 2025

DOI: 10.19113/sdufenbed.1649624

Vision Transformer-Based Blood Group Classification on Slide Images

Tansel Uyar[*1]

[1]Başkent Üniversitesi, Mühendislik Fakültesi, Biyomedikal Mühendisliği Bölümü, 06790, Ankara, Türkiye

(Alınış / Received: 03.03.2025, Kabul / Accepted: 15.04.2025, Online Yayınlanma / Published Online: 25.08.2025)

Keywords
Blood Group, Image Processing, Classification, Vision Transformer

Abstract: Blood is an essential fluid in the human body, enabling the transport of oxygen and nutrients. In cases of accidents, persistent hematological illnesses, or surgical procedures, blood transfusions are imperative for the restoration of lost blood volume. Therefore, it is imperative to ascertain the patient's blood type prior to any blood transfusion. In contemporary applications of blood group determination, the utilization of test serums (Anti-A, Anti-B, Anti-D) that facilitates the precipitation of antigens within the bloodstream has become a prevailing practice. Blood groups are determined by detecting the presence of antigens according to the precipitation of antigens. The observation of precipitation on slides is conducted by the laboratory specialist. However, this observation process can be arduous and time-consuming, requiring sustained attention for extended periods. Consequently, machine vision has emerged as a prevalent tool in contemporary automated approaches, mitigating the expert load required for these processes. Machine vision is a field that is constantly evolving, and one of the most current methods employed is the use of vision transformers. The objective of the present study was to achieve high-accuracy classification of blood groups by employing vision transformers for the purpose of discrimination. The study examined a classification application and its results using a vision transformer for machine vision, a technique particularly suited to automation systems in laboratory settings. The findings of experimental studies have indicated the notable efficacy of vision transformers in the classification of blood groups. In antigen detection, an average accuracy of 98.08% for stochastic gradient descent moment optimizer and 98.74% for adaptive moment estimation optimizer was obtained. In the detection of Rh protein, an average accuracy of 89.84% for stochastic gradient descent moment optimizer and 97.16% for adaptive moment estimation optimizer was obtained. Furthermore, experimental findings have indicated that the Adaptive Moment Estimation optimizer when employed in conjunction with vision transformers attains better classification performance.

Lam Görüntülerinde Görü Dönüştürücü Tabanlı Kan Grubu Sınıflandırması

Anahtar Kelimeler
Kan Grubu, Görüntü İşleme, Sınıflandırma, Görü Dönüştürücü

Öz: Kan, insan vücudunda oksijen ve besin taşınmasını sağlayan temel bir sıvıdır. Kazalar, kalıcı hematolojik hastalıklar veya cerrahi prosedürler durumunda, kaybedilen kan hacminin geri kazanılması için kan transfüzyonları zorunludur. Bu nedenle, herhangi bir kan transfüzyonundan önce hastanın kan grubunun belirlenmesi zorunludur. Kan grubu tespitinin güncel uygulamalarında, kan dolaşımındaki antijenlerin çökelmesini kolaylaştıran test serumlarının (Anti-A, Anti-B, Anti-D) kullanımı yaygın bir uygulama haline gelmiştir. Kan grupları, antijenlerin çökelmesine göre antijenlerin varlığının tespit edilmesiyle belirlenir. Slaytlardaki çökelmenin gözlemlenmesi laboratuvar uzmanı tarafından gerçekleştirilir. Ancak, bu gözlem süreci zorlu ve zaman alıcı olabilir ve uzun süreler boyunca sürekli dikkat gerektirebilir. Özetle, makine görüsü güncel otomatikleştirilmiş yaklaşımlarda yaygın bir araç olarak ortaya çıkmış ve bu süreçler için uzman yükünü azaltmıştır. Makine görüsü sürekli gelişen bir alan olmakla birlikte kullanılan en güncel yöntemlerden biri de görü dönüştürücüleridir.

*Corresponding author: tuyar@baskent.edu.tr

Sunulan çalışmanın amacı, kan gruplarının ayrımı amacıyla görü dönüştürücülerini kullanarak kan gruplarının yüksek doğrulukta sınıflandırılmasını elde etmektir. Çalışmada, özellikle laboratuvar ortamlarındaki otomasyon sistemleri için uygulanabilecek bir teknik olan makine görüsü yaklaşımında görü dönüştürücü yönteminin kullanımıyla gerçekleştirilen bir sınıflandırma uygulaması ve sonuçları incelenmiştir. Deneysel çalışmaların bulguları, görü dönüştürücülerinin kan gruplarının sınıflandırılmasında dikkate değer bir başarıya sahip olduğunu göstermiştir. Antijen tespitinde stokastik gradyan iniş moment eniyileyici için ortalama %98,08 ve uyarlamalı moment tahmini eniyileyici için ortalama %98,74 oranında bir doğruluk elde edilmiştir. Rh proteinin tespitinde ise stokastik gradyan iniş moment eniyileyici için ortalama %89,84 ve uyarlamalı moment tahmini eniyileyici için ortalama %97,16 oranında bir doğruluk elde edilmiştir. Öte yandan, deneysel bulgular, Uyarlamalı Moment Tahmini eniyileyicinin görü dönüştürücüleriyle birlikte kullanıldığında daha iyi sınıflandırma performansı elde ettiğini göstermiştir.

1. Introduction

Blood is a vital fluid in the human body, facilitating the movement of oxygen and essential nutrients via the heart along circulatory pathways. It comprises erythrocytes, leukocytes, plasma, and thrombocytes [1]. In the event of an accident, chronic hematological disorder or surgical operation, the body requires replenishment of lost blood through transfusion from a donor. Prior to conducting a blood transfusion, it is essential to perform a comprehensive array of compatibility assessments, referred to as pretransfusion tests [2,3]. One such compatibility assessment is the blood group evaluation. The determination of blood group can be fundamentally characterized as the identification of specific antigens present on red blood cells [4]. According to the aforementioned definition, the blood group system is a classification system predicated on the presence or absence of antigens in human blood. The ABO blood group system, initially described by Karl Landsteiner in 1900, underwent a renaming in 1907 by Dr. Ludwik Hirszfeld and Dr. Emil Von Dungern, who classified it as the A, B, AB, and O blood group system [5]. A significant study on the determination of blood groups was conducted by Landsteiner and Wiener in 1940, which is widely regarded as a foundational research in the field [6]. In this study, researchers found that antiserum created by injecting blood samples from the Macaca rhesus monkey into rabbits caused aggregations in the blood of 85% of white Americans [6]. The newly discovered factor was designated "Rh" by the authors, deriving its name from the Rhesus monkey. Individuals carrying the Rh factor in their red blood cells were categorized as positive (+) for this factor, while those lacking the factor were designated as Rh negative (-) [7]. The standard procedure for determining blood group involves the administration of tests for A, B, and D (Rh) antigens. However, tests for antigens other than those mentioned are only conducted in exceptional circumstances [4]. The utilization of test kits is essential in the detection of A, B, and D (Rh) antigens. These kits contain serum fluids that facilitate the precipitation of the specific antigens. Depending on the precipitation or non-precipitation status of these three antigens, eight blood group combinations are formed. Following the application of the test kit, the sedimentation conditions are examined visually, and the blood group is determined.

While these visual inspections are typically conducted manually by laboratory experts, the use of automated methods in automated systems is also gaining popularity [8-13]. The existing literature includes approaches such as image processing, machine learning, and deep learning in the studies carried out for automated determination of blood group. In the context of image processing, methodologies employed in related studies include pixel counting with edge detection subsequent to thresholding [14-19], morphological operations with thresholding [20], and image matching technique [21]. In the study of Ayan et al., after implementing adaptive thresholding, edge detection and morphological operations were sequentially executed [14]. It was reported that 99% accuracy was obtained from the test images of the Gel Test Method. In the study of Atıcı et al., after implementing global thresholding, morphological operations and edge detection operations were applied, respectively, and pixel numbers were used as features [15]. As a result of the classification of the test images, 84.61% accuracy was achieved. Similar approaches have been applied in other studies [16-21]. With the exception of one study, no information regarding the testing processes was shared in the other studies [16-20]. According to the findings of the study conducted by Odeh et al., the proposed method demonstrated a 99.6% accuracy rate [21]. In the

context of machine learning research, classification studies have been conducted employing various classifiers, including Support Vector Machine (SVM), Decision Tree (DT), Linear Discriminant Analysis (LDA), Logistic Regression (LR), and K-Nearest Neighbor (KNN), Artificial Neural Network (ANN) in conjunction with features such as entropy, mean, standard deviation, contrast, energy, correlation, and homogeneity, subsequent to segmentation [22-25]. In the study by Rosales et al., features were generated with the gray-level co-occurrence matrix (GLCM), which is utilized in texture analysis and training processes were carried out [22]. Variants of SVM, DT, KNN LDA and LR machine learning methods were used as classifiers. A classification approach is proposed based on the presence of sedimentation in segmented image patches with binary classification. It has been reported that the presence of sedimentation was detected with 97% accuracy using Coarse DT. In the study by Dannana et al., training was performed with an SVM classifier using mean, standard deviation, entropy, variance and kurtosis features after segmentation [23]. The study reported an accuracy value of 92.37% for the 8-class classification. In the study conducted by Mahmood, it was reported that 99.7% accuracy rate was achieved with the ANN classifier [24]. In the study of Ferraz et al., the presence of sedimentation was classified using SVM from individual features obtained for different types of ROI image patches [25]. It was reported in the study that F1-Score values ranging from 52.63% to 100% were obtained. Several studies have been conducted employing the deep learning approach, which represents one of the most advanced methodologies in blood group classification. In the study by Balaji et al., the classification performance of the convolutional neural network (CNN) structure proposed by the study group was compared with that of the AlexNet and LeNet networks [26]. In the study, it was reported that 46.07% training accuracy was achieved with AlexNet, 97.79% with LeNet, and 96.69% with the proposed model. In the study conducted by Titus et al., the classification process was performed by detecting sedimentation and non-sedimentation cases with the CNN-based segmentation network suggested by the study group [27]. In the study, it was reported that the average sensitivity value was 86.29% and the average specificity value was 83.98%. In a further study undertaken by Aboubaker, the classification performances of the CNN structure proposed by the researcher were compared with those of ResNet, AlexNet and VGG-16 networks to investigate the potential correlation between fingerprints and blood groups [28]. In a separate study, Shen et al. examined the classification performance of gel test method images using the improved AlexNet network [29]. In the study, it was reported that the test accuracy value of the gel test method images was obtained as 96.9%.

In recent years, one of the most current approaches that can compete with CNN-based methods is that of vision transformers (ViT). ViTs are one of the most frequently preferred models in the field of natural language processing used in processing texts, and ViTs emerged by combining the self-attention mechanism with patch-based architecture and applying it to images [30,31]. A gap in literature was identified by the absence of studies examining the performance of vision transformers in blood group classification of slide images. Therefore, in the present study, the classification performance of vision transformers was investigated using a dataset of slide images.

2. Material and Method

In the material and method section, a concise overview of the methods employed, a comprehensive description of the analyzed data, and the precise specifications of the experimental parameters of the analysis are provided under the respective headings.

2.1. Vision transformers

ViT represents a deep learning model that assigns distinct weights to each component of the image input data, incorporating an attention mechanism [32]. The ViT model, as depicted in Figure 1, involves the transformation of the input image ($x; x \in R^{H\times W\times C}$) into a series of smaller image pieces, each with dimensions smaller than the original image. Each reconstructed image part (p) is projected onto the z_0 vector using a learned embedding matrix (E) with the equation (1).

$$z_0 = \left[x_{class}; x_p^1 E; x_p^2 E; \ldots; x_p^N E\right] + E_{pos} \qquad (1)$$

Here E is the element of $R^{(P^2\times C)\times D}$, E_{pos} is the element of $R^{(N+1)\times D}$, and x_p is the element of $R^{N\times(P^2\times C)}$. N represents the number of extracted image parts and C represents the number of channels. x_{class} represents the mark added to the image parts and used for classification. Transformer encoder consists of L identical layers. Each layer consists of multi-head self-attention (MSA), multi-layer perceptron (MLP), layer normalization (LN) and residual connection. When the input sequence z_0 is

given as input to the encoder layer, the entire process applied to obtain the output map z_l with the same length was summarized by equations (2) and (3). l represents the relevant transformer encoder layer number.

$$\acute{z_l} = MSA\big(LN(z_{l-1})\big) + z_{l-1} \quad (2)$$

$$z_l = MLP\big(LN(\acute{z_l})\big) + \acute{z_l} \quad (3)$$

Here it is expressed as $l = 1 \ldots L$. The output of each transformer encoder layer is applied as input to the subsequent encoder layer. The final layer of the MLP placed at the output of the L-identical encoder layer is employed in conjunction with the SoftMax classifier to facilitate the classification of the learned image representations.

Given a sequence $X \in R^{n \times d}$ as input, the self-attention mechanism predicts the relationship of one element to another. X represents the input vector. In basic terms, a self-attention layer updates each component in the array by combining the whole information from all input elements. The interaction obtained from the inputs is realized by defining three learnable weight matrices to transform the matrices called query (W^Q), key (W^K), and value (W^V) obtained from the input vectors. The input array X is initially traced to these weight matrices through the following equations, $Q = XW^Q$, K= XW^K, V= XW^V. The output of the self-attention layer is calculated by equation (4).

$$\underset{Attention}{Self}(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (4)$$

In the present study, pre-trained (ImageNet 2012) a vision transformer model (ViT-Base16) was utilized, which was fed with an input image of 384×384×3 dimensions for the input layer, consisted of 11 transformer encoder layers, and was trained with the fine-tuning method. The base-sized model has 86.8 million trainable parameters with a patch size of 16.

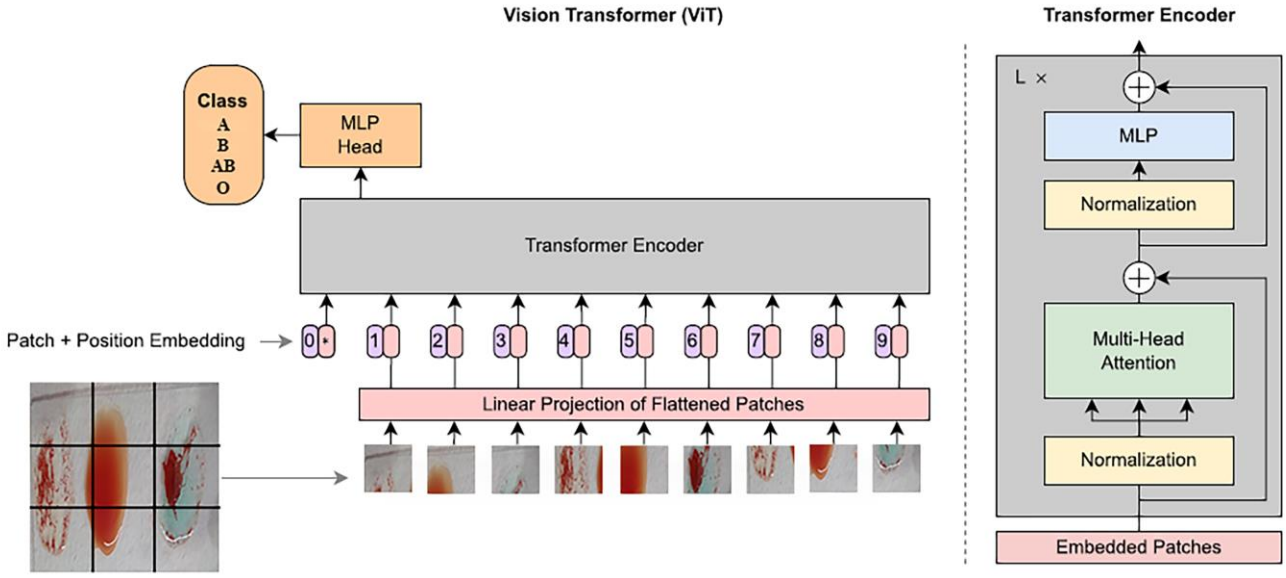

Figure 1. Structural architecture of classifying an image using vectoral input and transformer encoders with attention mechanisms

2.2. Data preparation and data augmentation

The present study utilized an open-source dataset [33]. The dataset contains slide images where the sedimentation status can be examined as a result of the application of anti-A, anti-B, and anti-D serum to the blood samples taken. The image numbers of blood groups A, B, AB, and O in the dataset were summarized in Table 1.

Table 1. Image distribution of dataset used in blood group classification

Blood Group	Rh (+)	Rh (-)	TOTAL
A	264	504	**768**
B	445	373	**818**
AB	366	392	**758**
O	382	391	**773**
TOTAL	**1457**	**1660**	**3117**

The data augmentation method was implemented to enhance the utilization of the data contained within the dataset and to mitigate the model's tendency to memorize (overfitting) during the training process. An

overview of the transform features employed for data augmentation, along with the parameter values associated with these features, was provided in Table 2. While the translation and shearing parameters were chosen at the value that produced the least noise in the image (by calculating the PSNR), the rotation parameter was selected from values frequently applied and recommended in applications [34].

Table 2. Transform parameters of data augmentation method

Transform	Parameter Range
Rotation	(-45) ° - (45) °
Translation of X Axis	(-3) pixel – (3) pixel
Translation of Y Axis	(-3) pixel – (3) pixel
Shearing of X Axis	(-3) pixel – (3) pixel
Shearing of Y Axis	(-3) pixel – (3) pixel

2.3. Proposed method and training parameters

The training process of the detailed vision transformer model was carried out in two stages. In the initial phase of the training, a 4-class (A, B, AB, O) training was conducted solely for the purpose of distinguishing the sedimentation status of A and B antigens, without taking into account the sedimentation status of the Rh protein. In the subsequent phase of the training, images of blood groups were utilized individually to classify the sedimentation status of Rh proteins. At this stage, four independent binary classification training procedures were executed for four distinct blood groups. The algorithm of the proposed method in the present study was summarized in Figure 2. The presented approach yielded models that can be used in a structure capable of performing gradual classification (Appendix A).

The fine-tuning method was implemented by training specifically the fully connected layer of the vision transformer model. To mitigate the risk of overfitting during the training process, a 10-fold cross-validation method was employed. Two distinct optimizers, Adaptive Moment Estimation (ADAM) and Stochastic Gradient Descent with Momentum (SGDM), were employed during the training process. The learning rate was set to 0.0001, and the mini-batch size was chosen as 8. Cross-entropy was used for calculating loss function. The training process was executed for a total of 10 epochs, and the weights from the epoch that exhibited the best validation loss value were selected as the final weights.

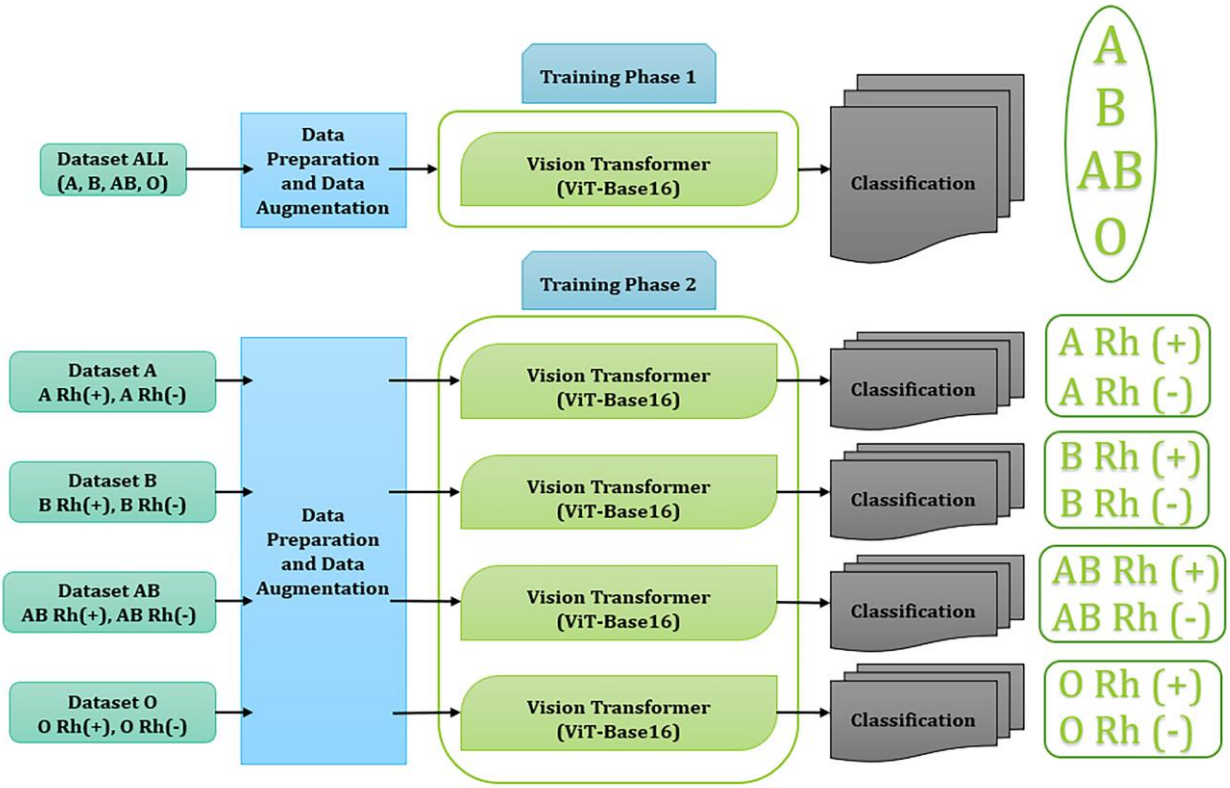

Figure 2. The algorithm of the proposed method in the present study

2.4. Performance metrics

The assessment of the classification outcomes was performed using various metrics. Metrics including

accuracy, recall, precision, and F1-score computed for each class and in total. The classification performance of the model was expressed by the mean and standard deviation ($\overline{X} \pm SD$) of the test data obtained as a result of the cross-validation method used for training. The mathematical formulations of these performance measures are provided below:

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$Recall=\frac{TP}{TP+FN} \quad (6)$$

$$Precision=\frac{TP}{TP+FP} \quad (7)$$

$$F1\text{-}score=2\ x\ \frac{Precision\ x\ Recall}{Precision\ +\ Recall} \quad (8)$$

True positive (TP) items are those for which the model correctly identifies a positive classification, and which are, in fact, positive. False positive (FP) items are those for which the model incorrectly identifies a positive classification, and which are, in fact, negative. True negative (TN) items are those for which the model correctly identifies a negative classification, and which are, in fact, negative. False negative (FN) items are those for which the model incorrectly identifies a negative classification, and which are, in fact, positive.

3. Results

As detailed in the materials and methods section, the four-class classification results obtained using ADAM and SGDM optimizers in the first stage training process were summarized in Table 3. Upon examination of the results, it was determined that both optimizers were capable of attaining elevated discrimination outcomes. It is evident that both optimizers demonstrate an average accuracy that exceeds 98%, with a negligible disparity of less than 1% between their respective means. Upon examination of the results according to class, it becomes evident that the average accuracy values exceed 97%. However, since there was no prioritization of classes, an examination of the F1-score values of the classes reveals that the ADAM optimizer was the most successful. While the F1-Score values obtained with the ADAM optimizer were above 96% when standard deviations are considered, these values were determined as 92% for SGDM. For this reason, ADAM optimizer comes to the fore in blood group classification based on antigen sedimentation.

In the subsequent phase of the study, ViT models were trained to classify the sedimentation status of Rh proteins for images of classified blood groups. Subsequently, the capacity of these models to detect the Rh protein in each blood group was examined. The Rh protein classification success of the models in the second stage of the study was summarized in Table 4. In this step, while the ADAM optimizer produced the most successful performance results, the performance results of the SGDM optimizer did not reach the level of the results in the previous step. An examination of the average accuracy values obtained from binary classifications reveals that the ADAM optimizer achieved an accuracy range of 95% to 99%, while the SGDM optimizer attained an accuracy range of 80% to 95%. Upon examination of the F1-Score values in the context of the imbalanced distributions across certain classes (group of A and B) within the data set, the outcomes align with those of the accuracy values. Upon examination of both the mean accuracies and the mean F1-score values, it is evident that the discrimination success for the two optimizers was AB, O, B, and A, starting from the most successful, respectively. In addition, when the averages of the F1-Score results of binary classifications were calculated, an average value of 90.04% was obtained for the SGDM optimizer, while an average value of 97.10% was obtained for the ADAM optimizer. Upon consideration of the results obtained, it was determined that the vision transformer models which were trained using the ADAM optimizer were the most successful model.

Table 3. Performance results of four class classification

Optimizer	Blood Group	Accuracy	Recall	Precision	F1-Score
ADAM	A	0,9868 ± 0,0074	0,9791 ± 0,0153	0,9683 ± 0,0246	0,9735 ± 0,0149
	AB	0,9887 ± 0,0085	0,9683 ± 0,0226	0,9854 ± 0,0209	0,9766 ± 0,0176
	B	0,9852 ± 0,0075	0,9815 ± 0,0226	0,9637 ± 0,0272	0,9721 ± 0,0141
	O	0,9891 ± 0,0069	0,9703 ± 0,0277	0,9859 ± 0,0160	0,9777 ± 0,0145
	Overall	0,9874 ± 0,0054	0,9748 ± 0,0109	0,9758 ± 0,0102	0,9753 ± 0,0105
SGDM	A	0,9817 ± 0,0161	0,9648 ± 0,0287	0,9623 ± 0,0425	0,9632 ± 0,0315
	AB	0,9871 ± 0,0106	0,9472 ± 0,0438	1 ± 0	0,9724 ± 0,0242
	B	0,9765 ± 0,0167	0,9668 ± 0,0564	0,9453 ± 0,0240	0,9551 ± 0,0340
	O	0,9781 ± 0,0166	0,9676 ± 0,0392	0,9486 ± 0,0509	0,9568 ± 0,0312
	Overall	0,9808 ± 0,0087	0,9616 ± 0,0172	0,9640 ± 0,0156	0,9628 ± 0,0164

Table 4. Performance results of Rh protein classification

Optimizer	Performance Metric	A Rh	AB Rh	B Rh	O Rh
ADAM	Accuracy	0,9623 ± 0,0530	0,9907 ± 0,0164	0,9584 ± 0,0321	0,9753 ± 0,0248
	Recall	0,9542 ± 0,0659	0,9906 ± 0,0165	0,9600 ± 0,0302	0,9755 ± 0,0246
	Precision	0,9617 ± 0,0543	0,9909 ± 0,0164	0,9595 ± 0,0306	0,9763 ± 0,0229
	F1-Score	0,9579 ± 0,0599	0,9907 ± 0,0164	0,9597 ± 0,0304	0,9759 ± 0,0238
SGDM	Accuracy	0,8006 ± 0,1375	0,9591 ± 0,0373	0,8973 ± 0,0565	0,9366 ± 0,0487
	Recall	0,7960 ± 0,1164	0,9597 ± 0,0367	0,8980 ± 0,0535	0,9366 ± 0,0489
	Precision	0,8104 ± 0,1124	0,9607 ± 0,0350	0,9020 ± 0,0496	0,9411 ± 0,0418
	F1-Score	0,8027 ± 0,1129	0,9602 ± 0,0358	0,9000 ± 0,0514	0,9388 ± 0,0452

4. Discussion and Conclusion

The study's main objective was to ascertain blood groups with a high level of accuracy, contingent on their sedimentation status after implementing anti-A, anti-B, and anti-D serum on test slides. To achieve this determination, the vision transformer approach was selected as the preferred method. ViTs are among the state-of-the-art applications in literature. The classification success of the ViTs in the relevant images was evaluated. Unlike other studies in the literature, this evaluation was performed with an open-source data set. Most studies indicate that the datasets in question were collected with the specific intention of being utilized in the aforementioned studies. Conversely, a study indicated that the dataset was uploaded to the data repository (Kaggle), yet it remained inaccessible [26]. Therefore, a direct comparison of the performance outcomes of the methodology employed in this study with those of other studies was not feasible. Conversely, due to the incomplete characterization of the dataset employed in certain studies, performance outcomes were compared based on overall values. Studies using image processing techniques not directly related to the method used in this study, reported test accuracies between 84.61% and 99.6%. Although promising results have been presented, it is noteworthy that the number and diversity of images reported to be used for testing are low in these studies [14-20]. It has been reported that machine learning studies have achieved a high rate of success in classification [22-25]. Nevertheless, the employment of image processing techniques, a prevalent component of these studies, may result in scenarios necessitating human involvement. Consequently, deep learning methods, which are automated approaches to producing features while minimizing human involvement, have become more prominent [26-29]. Very successful results have also been obtained in studies carried out with the deep learning approach. Despite the absence of data regarding the number of classes that were evaluated in the study conducted by Balaji et al., the classification performance of the proposed model was reported to be 96.69% [26]. In Amballa's study, an overall classification success rate of 98.50% was reported, achieved through a method that combined machine learning and deep learning [35]. Due to the absence of specified details regarding the dataset and the number of classes studied, direct comparisons between classes were not possible [26,35]. However, if the evaluation was based on overall values, it can be concluded that the ViT model trained for four classes using the ADAM optimizer was one of the most successful classifiers among these studies.

A thorough examination of the study's findings reveals that the ADAM optimizer emerges as the most effective optimizer among the two. While there was no substantial superiority in terms of optimizers in the results obtained for the four-class classification, it can be posited that the ADAM optimizer was more adept at determining the Rh protein of blood groups. It has been documented that the generalization capability of

the ADAM optimizer, which is less influenced by the initial conditions and requires less fine-tuning, is surpassed by that of SGDM [36]. However, the findings of the present study demonstrated an opposing conclusion. One of the reasons for this situation is that it has been reported that the ADAM optimizer may be a better option if hyperparameter optimization was performed [37]. Therefore, it is thought that the hyperparameters selected in the study cause the ADAM optimizer to produce better results. On the other hand, another reason can be associated with the number and distribution of data in the dataset. While the data numbers employed for the four-class classification demonstrate a balanced distribution across classes, the underperformance of the Rh classification in A and B blood groups for SGDM can be attributed to the imbalanced data numbers. On the other hand, the results indicate that the ADAM optimizer was less affected by this situation, optimizes faster, and performs better in tests. Another finding of the obtained results is that the efficacy of the training process is directly proportional to the quantity of data utilized for training in ViT models. This observation aligns with the findings reported in other studies in the literature [32,38]. A detailed examination of recall metrics revealed a direct correlation between the true positive rate and the number of data points classified within specific classes.

In the presented study, the success of vision transformers in blood group classification as well as the role of optimizers in this situation were investigated. In antigen detection, an average accuracy of 98.08% for SGDM optimizer and 98.74% for ADAM optimizer was obtained. In the detection of Rh protein, an average accuracy of 89.84% for SGDM optimizer and 97.16% for ADAM optimizer was obtained. The results obtained are quite promising. In the contemporary era, artificial intelligence approaches designed for machine vision, in conjunction with automated systems, have become a standard component of technological applications. A notable aspect of the presented work was the evaluation of the potential of vision transformers for decision-making in an automated system. In the future perspective, hyperparameter optimization with more data is suggested to improve the results of the presented study.

Declaration of Ethical Code

In this study, we undertake that all the rules required to be followed within the scope of the "Higher Education Institutions Scientific Research and Publication Ethics Directive" are complied with, and that none of the actions stated under the heading "Actions Against Scientific Research and Publication Ethics" are not carried out.

References

[1] Skull, A. 2022. Circulation - anatomy, physiology and pathophysiology. ss 1-86. McGloin, S., Skull, A. ed. 2022. Principles of Acute Care Nursing, SAGE Publications, UK, 280s.

[2] Roback, J. D., Grossman, B. J., Harris, T., Hillyer C. D. 2011. Technical Manual. 17th edition. American Association of Blood Banks (AABB). U.S.A., 1078s.

[3] Harmening, D. M. 2012. Modern Blood Banking & Transfusion Practices. 6th edition. F.A. Davis Company. U.S.A., 672s.

[4] Malomgre, W., Neumeister, B. 2009. Recent and future trends in blood group typing. Analytical and Bioanalytical Chemistry, 393, 1443-1451.

[5] Okroi, M., McCarthy, L. 2010. The original blood group pioneers: the Hirszfelds. Transfusion Medicine Reviews, 24(3), 244-246.

[6] Landsteiner, K., Wiener, A. S. 1940. An Agglutinable Factor in Human Blood Recognized by Immune Sera for Rhesus Blood. Experimental Biology and Medicine, 43(1), 223-223.

[7] Şentuna, C. 1982. Rh Gen Frekansları Yönünden Türkiye'nin Yeri. Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Dergisi, 30, 153-179.

[8] Ferraz, A., Carvalho, V., Machado, J. 2017. Determination of human blood type using image processing techniques. Measurement, 97, 165-173.

[9] Shaban, S. A., Elsheweikh, D. L. 2022. Blood Group Classification System Based on Image Processing Techniques. Intelligent Automation & Soft Computing, 31(2), 817-834.

[10] Mankar, J., Neve, S., Parkhe, M., Kumawat, P., Kale, N. R. 2021. Automated Blood Group Detection System Using Image Processing. International Journal of Advance Research and Innovative Ideas in Education, 7(3), 3237-3244.

[11] Malhotra, A., Chethana, N., Nagapranitha, D. R., Anand, S., Shruthi, B. 2021. Automated Blood Group Determination Using Image Processing Techniques with Integration of Raspberry Pi. Journal of Emerging Technologies and Innovative Research, 8(6), 954-965.

[12] Jayakumar, P., Padmanabhan, S., Suthendran, K., Kumar, Y. N., Sujith, M. 2020. Identification and

Analysis of Blood Group with Digital Microscope Using Image Processing. IOP Conference Series: Materials Science and Engineering, 923, 012013.

[13] Ferraz, A., Soares, F., Carvalho, V. 2013. A Prototype for Blood Typing Based on Image Processing. SENSORDEVICES 2013: The Fourth International Conference on Sensor Device Technologies and Applications, 25-31 August, Barcelona, 139-144.

[14] Ayan, E., Yıldırım, E. K. 2016. Real Time Blood Type Determination by Gel Test Method on an Embedded System. International Journal of Applied Mathematics, Electronics and Computers, 4, 412-415.

[15] Atıcı, H., Koçer, H. E., Kader, S. 2020. Dijital Görüntüler Kullanılarak Kan Grubunun Görüntü İşleme Tabanlı Tespiti. BŞEÜ Fen Bilimleri Dergisi, 7(2), 649-659.

[16] Dong, Y., Chen, N., Fu, W., Liu, M., Zhou, Z., Chen, S. 2017. ABO Blood Group Detection Based on Image Processing Technology. 2017 2nd International Conference on Image, Vision and Computing, 2-4 June, Chengdu, China, 655-659.

[17] Dhande, A., Bhoir, P., Gade, V. 2018. Identifying the blood group using Image Processing. International Research Journal of Engineering and Technology, 5(3), 2639-2642.

[18] Ravindran, G., Titus, T. J., Pravin, M., Pandiyan, P. 2017. Determination and Classification of Blood Types using Image Processing Techniques. International Journal of Computer Applications, 157(1), 12-16.

[19] Chandra, K. B. S., Madgula, V. B., Raghavaraju, D. 2023. Prototype for Blood Typing Based on Image Processing. Advanced Engineering Science, 55(4), 1-11.

[20] Rishitha, G. S., Jerusha, D. S., Shehanaz, S. 2022. Blood Detection Using Image Processing. International Journal of Creative Research Thoughts, 10(5), 365-369.

[21] Odeh, N., Toma, A., Mohammed, F., Dama, Y., Oshaibi, F., Shaar, M. 2021. An Efficient System for Automatic Blood Type Determination Based on Image Matching Techniques. Applied Sciences, 11, 5225-5252.

[22] Rosales, M. A., de Luna, R. G. 2022. Computer-Based Blood Type Identification Using Image Processing and Machine Learning Algorithm. Journal of Advanced Computational Intelligence and Intelligent Informatics, 26(5), 698-705.

[23] Dannana, S., Prasad, D. Y. V. 2022. Blood group detection using ML classifier. Journal of Health Sciences, 6(S1), 4395–4408.

[24] Mahmood, M. F. 2024. Recognition And Categorization of Blood Groups by Machine Learning and Image Processing Method. Innovative Biosystems and Bioengineering, 8(2), 53-68.

[25] Ferraz, A., Brito, J. H., Carvalho, V., Machado, J. 2017. Blood type classification using computer vision and machine learning. Neural Computing and Applications, 28, 2029-2040.

[26] Balaji, B., Jeyasakthi, R., Julius Fusic, S., Rishwana, M., Swathilakshmi, P. R. K., Reshma, K. K. 2021. A novel approach of classifying ABO blood group image dataset using deep learning algorithm. International Conference on Computational Performance Evaluation, 1-3 December, Meghalaya, India, 393-398.

[27] Titus, A., Devi, K. M., Divya, G., Nimmagadda, P., Shankar, S. V. 2023. A Systematic Procedure to Identify Human Blood Groups by using Image Processing Assisted Learning Principle. International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems, 14-15 December, Chennai, India, 1-6.

[28] Aboubaker, C. D. 2023. Blood Group Prediction Using Deep Learning. Kasdi Mebah University, Computer Science Department, Master Thesis, 70s, Algeria.

[29] Shen, R., Wen, J., Zhu, P. 2023. Blood Group Interpretation Algorithm Based on Improved AlexNet. Electronics, 12, 2608-2622.

[30] Devlin, J., Chang, M., Lee, K., Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. North American Chapter of the Association for Computational Linguistics, 1, 4171-4186.

[31] Vaswani, A., Ramachandran, P., Srinivas, A., Parmar, N., Hechtman, B., Shlens, J. 2021. Scaling local self-attention for parameter efficient visual backbones. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 20-25 June, Nashville, U.S.A., 12894-12904.

[32] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N. 2020. An image is worth 16x16 words: Transformers for image recognition at scale. International Conference on Learning Representations. International Conference on Learning Representations, 26 April – 1 May, Virtual, 1-22.

[33] Minor Project 5th Semester. 2024. Blood Group Detection Dataset. http://universe.roboflow.com/minor-project-5th-semester/blood-group-detection-4yvdx (Erişim Tarihi: 01.03.2025).

[34] Shorten, A., Khoshgoftaar T. M. 2019. A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(60), 1-48.

[35] Amballa, D. P. 2022. Automated Blood Group Identification using Machine Learning and Deep Learning: A Novel Approach for Laboratory Settings. International Journal of Science and Research, 11(10), 1390-1393.

[36] Wilson, A. C., Roelofs, R., Stern, M., Srebro, N., Recht, B. 2017. The marginal value of adaptive gradient methods in machine learning. 31st Conference on Neural Information Processing Systems, 4-9 December, U.S.A., 1-11.

[37] Choi, D., Shallue, C. J., Nado, Z., Lee, J., Maddison, C. J., Dahl, G. E. 2019. On empirical comparisons of optimizers for deep learning. arXiv, 1-27.

[38] Neyshabur, B. 2020. Towards Learning Convolutions from Scratch. 34th Conference on Neural Information Processing Systems, 6-12 December, 1-11.

Appendices

Appendix A. Gradual classification via trained models