

ABO-BTI: AN OPEN-SOURCE ABO BLOOD TYPING IMAGE DATASET FOR MEDICAL AI APPLICATIONS

Daas SARA¹ , Zehir HATEM^{1,*} , Chebli ASMA^{2,*} , Hafs TOUFIK^{1,*} ,
Hadeef CHAIMA^{3,*}

¹Laboratory of Study and Research in Instrumentation and Communication of Annaba,
Faculty of Technology, Badji Mokhtar-Annaba University, P.O. BOX 12, 23000, Annaba, Algeria

²L.R.I Laboratory, Department of Computer Science, Faculty of Technology,
Badji Mokhtar-Annaba University, P.O. BOX 12, 23000, Annaba, Algeria

³Department of Biology, Faculty of Science, Badji Mokhtar-Annaba University,
P.O. BOX 12, 23000, Annaba, Algeria

sara.daas@univ-annaba.dz, hatem.zehir@univ-annaba.dz, asma.chebli@univ-annaba.dz,
toufik.hafs@univ-annaba.dz, chaimahadeef3@gmail.com

*Corresponding author: Daas Sara; sara.daas@univ-annaba.dz

DOI: 10.15598/aece.v24ix.250705

Article history: Received Jul 07, 2025; Revised Aug 22, 2025; Accepted Sep 16, 2025; Published xx xx, 202x.
This is an open access article under the BY-CC license.

Abstract. Accurate blood type classification is crucial for safe transfusions and clinical decision-making, yet existing research is limited by the lack of standardized, publicly available datasets for training and evaluating machine learning models. To address this gap, we introduce ABO-BTI (ABO Blood Typing Image), the first open-source dataset dedicated to blood type classification using high-resolution agglutination images. The dataset comprises 144 cases, with 432 images standardized to a resolution of 1280×590 pixels after processing. This study evaluates the effectiveness of deep learning for blood type identification using the ABO-BTI database. Three models, ResNet50, MobileNetV2, and a proposed deep learning architecture, were trained and tested on the dataset to assess its suitability for machine learning applications. The proposed model achieved an accuracy of 96.51%, significantly outperforming MobileNetV2 (12.64%) and ResNet50 (72.41%). Comparative analysis with traditional machine learning methods further demonstrated that deep learning provides competitive performance while reducing reliance on handcrafted feature extraction. These results highlight ABO-BTI as a valuable benchmark for advancing AI-driven blood type classification. The findings also suggest the potential integration of deep learning-based classification into embedded systems for real-time blood typing in point of care and emergency settings. By providing a standardized

dataset and demonstrating the viability of deep learning models, this study lays the foundation for future research in automated blood classification, with implications for both clinical applications and AI-driven medical diagnostics.

Keywords

Blood type classification, ABO-BTI dataset, deep learning, machine learning, agglutination images, biomedical image analysis, ResNet50, MobileNetV2, automated blood typing, transfusion medicine, AI in healthcare, convolutional neural networks, medical imaging.

1. Introduction

Accurate blood typing is a critical component of transfusion medicine and organ transplantation, ensuring compatibility between donors and recipients [1,2]. The identification of blood groups, primarily the ABO and Rh systems, is essential for preventing hemolytic reactions, which can result in severe complications, including hemolysis, renal failure, shock, and even death. Blood typing is also crucial in maternal-fetal medicine,

where Rh incompatibility can lead to hemolytic disease of the newborn (HDN) [3,4], necessitating timely intervention to prevent neonatal morbidity and mortality.

Mismatched transfusions due to blood typing errors pose significant clinical risks [5]. Incompatible transfusions trigger immune-mediated hemolytic reactions, in which the recipient's antibodies attack transfused red blood cells, causing their destruction. Acute hemolytic transfusion reactions (AHTRs) [6] are among the most serious complications, leading to disseminated intravascular coagulation (DIC), multiple organ failure, and death in severe cases [7,8]. Additionally, minor discrepancies in antigen typing can result in delayed hemolytic reactions, complicating patient recovery and increasing hospitalization duration [9].

Beyond transfusion medicine, accurate blood typing is vital for organ transplantation, where mismatches can lead to hyperacute rejection, reducing graft survival rates [10,11]. Furthermore, in forensic science and genetic studies, blood group determination plays a role in paternity testing and population genetics research [12,13].

In recent years, artificial intelligence (AI) has shown great promise in advancing medical diagnostics, including blood type classification [14]. However, a significant challenge in this field is the lack of standardized, open-source datasets. Reliable and diverse datasets are important for training and validating AI models to ensure their robustness and generalizability across different populations.

Unlike other areas of medical imaging and diagnostics, where large public datasets exist [15,16], blood typing remains underrepresented in open-source medical AI repositories. This lack of accessible data hinders the development of AI-driven solutions and slows progress in automating blood typing processes. Additionally, the limited availability of standardized datasets raises concerns about potential biases in AI models, which may lead to misclassification risks in real-world clinical applications.

The absence of widely accepted datasets also complicates benchmarking efforts, making it difficult to compare the performance of different AI models objectively. This challenge underscores the need for collaborative initiatives between medical institutions, research organizations, and policymakers to create and share high-quality, anonymized datasets. Addressing this gap will be essential for advancing AI-based blood typing solutions, improving diagnostic accuracy, and enhancing patient safety in transfusion medicine and transplantation.

Several studies have explored the application of AI in blood analysis, particularly in tasks such as blood cell

classification, anemia detection, and disease diagnosis [14,17].

In [17], the authors proposed a rehearsal-based continual learning approach aimed at improving the classification of white blood cells in diverse clinical settings. The study utilizes three distinct datasets: Matek-19 [18], INT-20 and Acevedo-20 [19], comprising 13 and 10 classes, respectively. The study [20] by R. Nithya and K. Nirmala uses images collected from Apollo Hospital with the consent of participants and guided by a histopathologist. The data set comprises a total of 435 images from 29 slides with 15 field views, including 180 normal images and 255 anaemic images. Another study [21] involves a dataset developed in collaboration with Shaukat Khanum Hospital and Research Center in Pakistan. The dataset consists of blood smear images collected from 50 patients, with 25 non-anemic and 25 anemic subjects. A total of 500 images were captured using an Olympus Dp27 8.9-megapixel CMOS sensor at a resolution of 1920x1080 pixels.

The authors of [22] presents a deep learning algorithm to diagnose leukemia by analyzing microscopic images of blood samples. The datasets used for this study are ALL_IDB1 and ALL_IDB2, which are publicly available and consist of blood samples from leukemia patients. ALL_IDB1 includes 108 images with 39000 blood components, while ALL_IDB2 contains cropped regions of cells.

In this study, we introduce ABO-BTI, the first open-source dataset comprising high-resolution ABO blood typing images with standardized annotations. Unlike prior studies, which primarily focus on broader hematological analysis such as blood cell classification and disease detection, ABO-BTI is specifically designed to facilitate the development, benchmarking, and clinical validation of AI models for ABO blood group classification. By providing real-world agglutination images with verified annotations, this dataset establishes a foundation for reproducible research and cross-institutional collaboration in AI-driven blood typing. The ABO-BTI dataset offers a diverse collection of images capturing various blood types and agglutination patterns under standardized conditions, ensuring robustness and generalizability in model training. Its availability addresses the lack of publicly accessible datasets necessary for algorithm evaluation and comparative benchmarking. Furthermore, the inclusion of high-quality annotations supports the development of AI models capable of automating blood typing with high accuracy, thereby enhancing transfusion safety and mitigating the risks associated with human error. By fostering innovation in AI-assisted blood classification, ABO-BTI represents a significant step toward integrating machine learning into transfusion medicine.

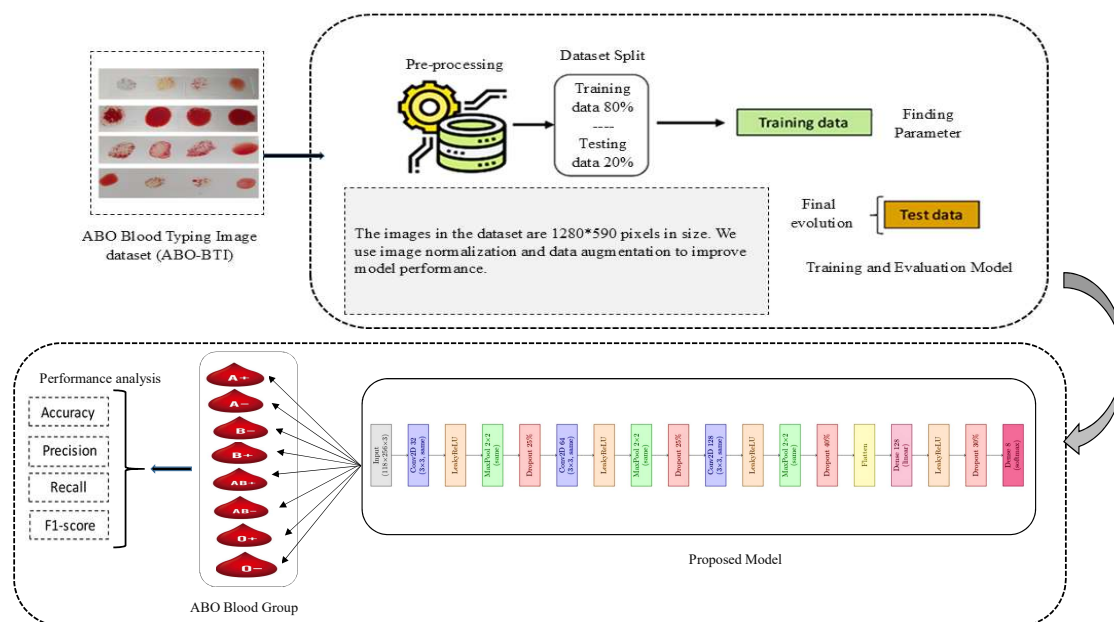


Fig. 1: Workflow of the proposed deep learning-based ABO blood typing classification model, including dataset preprocessing, training and evaluation, and performance analysis.

Beyond its primary application in AI-driven blood typing, ABO-BTI has the potential to support a wide range of additional use cases. In medical education and training, the dataset can serve as a valuable resource for teaching students and laboratory professionals how to interpret agglutination patterns accurately. By providing a standardized set of annotated images, ABO-BTI can aid in developing interactive learning tools and simulation-based training programs, improving diagnostic proficiency and reducing human error in manual blood typing procedures. Furthermore, ABO-BTI can contribute to quality control and standardization efforts in clinical laboratories. Automated systems trained on this dataset can assist laboratory technicians in verifying blood typing results, flagging ambiguous or potentially erroneous classifications, and ensuring consistency in testing procedures. This application is particularly relevant in high-throughput transfusion centers, where maintaining accuracy across large volumes of blood samples is critical for patient safety. The rest of this paper is organised as follows. In Section 2, materials and methods are described. Section 3 reports the experimental results with discussions. Finally, the last section 4 presents the conclusion and future works.

2. Materials and Methods

Figure. 1 illustrates the workflow of an ABO blood typing classification model based on deep learning techniques. The process begins with the ABO Blood Typing Image dataset (ABO-BTI), which comprises images of blood agglutination reactions. These images undergo pre-processing steps, including normalization and data augmentation, to enhance model performance. The dataset is then split into training (80%) and testing (20%) subsets. The training data is utilized to optimize model parameters, while the test data is used for final evaluation. The proposed model, based on a CNN architecture, is trained to classify blood samples into eight ABO blood groups (A+, A-, B+, B-, AB+, AB-, O+, and O-). Performance analysis is conducted using standard classification metrics, including accuracy, precision, recall, and F1-score, to assess the effectiveness of the model in blood group identification.

2.1. Data Collection

The data analyzed in this study were collected from two healthcare institutions in Annaba, Algeria: center Hospitalo-Universitaire Dr Dorban and Abdallah Nouaouria El Bouni Specialized Hospital. These facilities serve as regional referral centers, providing a

representative sample of demographic and clinical diversity within the province. Data extraction focused on the distribution of blood types in sex and age groups according to the ABO Blood Group System [23], with anonymized patient records spanning multiple years. Blood agglutination [24] assays were conducted by trained personnel following standardized laboratory protocols to ensure reproducibility and minimize variability. Images of agglutination patterns were captured using consistent imaging settings.

Collection of the ABO-BTI dataset was done using the World Health Organization (WHO) standards on ABO and RhD typing of blood [25] for clinical relevance and worldwide comparability. Every image file holds four sub-images arranged in 2x2 quadrants. These correspond to the four test tubes used in the routine ABO/D blood typing:

- Left—Anti-A reaction
- Middle right—Anti-B reaction
- Middle left—Positive control (Anti-AB)
- Right—Anti-D (Rh) reaction

The four tubes were: Prepared and drawn in equal aliquots from a single specimen of blood. Were incubated in the same conditions. Were captured in a fixed position using a digital camera and controlled lighting during a single imaging session to yield the same scale, orientation, and illumination. The inclusion of a positive control (Anti-AB) aids in confirming the presence of agglutination of red blood cells with antibodies, thereby validating sample reactivity. This also aids in distinguishing true negatives, like the absence of reactions because of a luxury of A/B antigens from non-reactive technical failures of cells, reagents, and surfaces.

Fig. 2 presents a representative image from the ABO-BTI dataset, demonstrating the characteristic morphological features utilized for subsequent deep learning-based classification.

1) Data Collection Procedures

The Source Data was meticulously curated to ensure a robust and representative data set for classification of blood types. The inclusion criteria required the availability of a documented blood type, limited to the eight primary categories. A+, A-, AB+, AB-, B+, B-, O+, and O-, alongside essential demographic metadata, including sex and age. To facilitate a comprehensive age-based analysis, participants were systematically stratified into seven distinct age groups: newborns (0-28 days), infants (1-12 months), children (1-12 years), adolescents (13-18 years), young adults (19-35 years), adults (36-64 years), and seniors/elderly (65+ years).



Fig. 2: Sample agglutination pattern captured during ABO-BTI dataset collection by a research team member.

adults (36-64 years) and seniors (65 + years). This stratification, described in Table 1, enables a structured evaluation of the distribution of the type of blood in different stages of development.

The composition of the initial data set comprised 164 cases, with a gender distribution of 66% female (108 cases) and 34% male (56 cases). The structured organization of age groups and the balanced inclusion of different blood types contribute to the suitability of the data set for machine learning applications in transfusion medicine, biomedical imaging, and AI-driven diagnostics.

Tab. 1: Age group distribution by gender. This table presents the demographic breakdown of participants in the ABO-BTI dataset, categorizing individuals into seven distinct age groups.

Age Group	Women	Men	Total
Newborns (0-28 days)	11	6	17
Babies (1-12 months)	11	6	17
Children (1-12 years)	22	11	33
Adolescents (13-18 years)	16	8	24
Young Adults (19-35 years)	27	14	41
Adults (36-64 years)	16	8	24
Seniors/Elderly (65+ years)	5	3	8
Total	108	56	164

To facilitate systematic and standardized data collection, a structured Google Form was designed and implemented to gather essential demographic and blood type information from participants. This approach ensured efficient data entry, real-time validation, and centralized management, enabling streamlined compilation and subsequent analysis. By utilizing an online form, data collection was conducted in a consistent and organized manner, minimizing errors and ensuring the reliability of recorded information.

The form was structured to capture critical variables necessary for constructing the dataset. Each participant was given an anonymized unique identification number, ensuring data traceability while maintaining confidentiality. Age categorization was an integral component, with participants stratified into seven predefined groups: newborns (0–28 days), infants (1–12 months), children (1–12 years), adolescents (13–18 years), young adults (19–35 years), adults (36–64 years), and seniors (65+ years). Gender identification was recorded as either male or female, contributing to demographic diversity within the dataset. Additionally, participants were required to report their blood type, selecting from the eight primary ABO and Rh blood group categories.

To ensure data integrity, the form was configured to enforce mandatory responses for all fields, preventing incomplete submissions and enhancing dataset completeness. Internal validation mechanisms minimized inconsistencies, while duplicate entries were identified and filtered during the preprocessing stage. No personally identifiable information beyond the anonymized unique identifier was collected, ensuring compliance with ethical research standards and protecting participant confidentiality.

2) Data Description

The ABO-BTI dataset consists of 164 systematically collected blood typing samples from two major hospitals in Annaba, Algeria: Centre Hospitalo-Universitaire Dr Dorban and E.H.S Abdallah Nouaouria El Bouni. This dataset has been meticulously curated to support advancements in medical imaging, AI-driven diagnostics, and automated blood type classification. By ensuring a balanced representation of blood types and demographic diversity, it facilitates the development of generalizable and clinically relevant machine learning models for transfusion medicine and computational diagnostics.

To mitigate potential biases in AI model training, gender representation within the dataset was carefully structured, with 66% of cases from female patients (108 samples) and 34% from male patients (56 samples). This proportion is essential for ensuring that trained

models maintain fairness, reliability, and unbiased predictive accuracy across different demographic groups.

Furthermore, the dataset maintains an equitable distribution of blood types, ensuring that each major category within the ABO and Rh systems is adequately represented. This structured approach is paramount for training robust classification models, capable of accurately predicting blood types across diverse populations. The breakdown of blood type distribution by gender is detailed in Table 2, reinforcing the dataset's suitability for AI-driven applications in transfusion medicine and clinical decision support systems.

Tab. 2: Blood Type Distribution by Gender. This table provides a detailed breakdown of blood type distribution within the ABO-BTI dataset, categorizing cases by gender.

Blood Type	Women (66%)	Men (34%)	Total Cases
A+	15	8	23
A-	13	7	20
AB-	12	6	18
AB+	13	6	19
B-	12	6	18
B+	14	7	21
O-	13	7	20
O+	17	8	25
Total	108	56	164

A summary of overall gender distribution is presented in Table 3, highlighting the dataset's demographic diversity. The inclusion of both sexes in well-defined proportions ensures equitable model performance, reducing gender-based discrepancies in classification accuracy while enhancing the dataset's reliability for machine learning applications in biomedical research.

Tab. 3: Overall Sex Distribution. This table presents the distribution of participants in the ABO-BTI dataset by gender.

Sex	Number of Cases	Percentage (%)
Women	108	66%
Men	56	34%
Total	164	100%

This structured dataset not only enhances the robustness of AI-driven classification models but also ensures equitable performance across patient demographics. By addressing potential sources of bias, this dataset serves as a valuable resource for developing machine learning applications in transfusion medicine, computational pathology, and biomedical diagnostics.

Each blood sample in the ABO-BTI dataset is accompanied by high-resolution agglutination images, which are critical for training machine learning mod-

els in automated blood type classification. These images capture a diverse range of agglutination patterns, enabling AI models to learn fine-grained distinctions between different blood types. The dataset contains a total of 492 images, with the distribution per blood type detailed in Table 4.

Tab. 4: Total Image Numbers per Blood Type. This table presents the distribution of images in the ABO-BTI dataset across the eight primary blood types.

Blood Type	Total Cases	Total Images
A+	23	69
A-	20	60
AB-	18	54
AB+	19	57
B-	18	54
B+	21	63
O-	20	60
O+	25	75
Total	164	492

The inclusion of multiple images per sample significantly enhances the dataset's robustness and applicability. This diverse collection of images not only supports classification tasks but also facilitates anomaly detection, quality control in blood testing laboratories, and educational initiatives in medical and biological sciences. By providing high-resolution images with standardized annotations, the ABO-BTI dataset establishes a valuable benchmark for AI-driven research in transfusion medicine.

This dataset holds substantial potential for advancing AI applications in clinical practice. By providing a well-balanced collection of high-resolution images, the ABO-BTI dataset enables researchers to develop and benchmark AI models for blood type classification, promote collaborative advancements in AI-driven diagnostics, and support clinical validation efforts. Beyond classification, the dataset is highly applicable to automated blood typing, laboratory quality assurance, and medical education, reinforcing its role as a foundational resource for AI-driven innovations in transfusion medicine.

The ABO-BTI dataset provides a comprehensive representation of blood type distribution across different age groups and genders. This section details the demographic breakdown of blood types in both male and female populations, which is essential for ensuring robust AI model training and clinical research. The dataset consists of 164 total cases, with 108 female cases (66%) and 56 male cases (34%). Understanding the variation of blood types across different age groups enhances the applicability of the dataset for research in transfusion medicine, immunohematology, and AI-driven diagnostics.

The male cohort consists of 56 cases, distributed across various age groups. Table 5 presents the detailed blood type distribution among men by age category.

The male population shows an even distribution across different blood types, with O+ and A+ being the most frequent. The highest concentration of cases is among young adults (19-35 years), indicating the age group most represented in the dataset.

The female cohort consists of 108 cases, covering a wider range of age groups. Table 6 provides a detailed overview of the blood type distribution among women by age category.

The female population shows a more varied age distribution, with a notable representation in the young adult (19-35 years) and child (1-12 years) categories. The most common blood type among women is O+, followed by A+.

The distribution of blood types across genders and age groups in this dataset ensures diversity, making it an invaluable resource for research in transfusion medicine and AI-driven diagnostics. The relatively high proportion of young adults in both genders suggests that the dataset is well-suited for training models aimed at clinical applications involving active donors and patients undergoing transfusions. Additionally, the inclusion of newborns and children provides opportunities for research in pediatric transfusion strategies and congenital blood disorders.

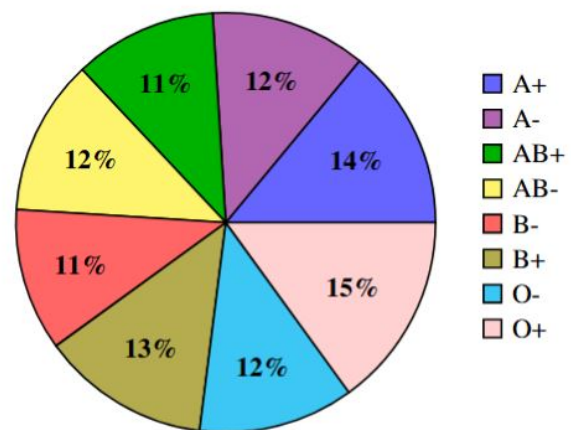


Fig. 3: Blood Type Distribution in the ABO-BTI Dataset. The figure presents the proportional representation of different blood types within the dataset, covering the eight primary ABO and Rh factor groups.

Figure 3 illustrates the distribution of blood types within the ABO-BTI dataset, highlighting the relative prevalence of each blood group in the studied population. As depicted in the pie chart, the most common blood type is O+, accounting for 15% of the total cases, followed by A+ at 14%. Other major blood groups include B+ (13%) and A- (12%). Meanwhile, AB-

Tab. 5: Blood Type Distribution Among Men by Age Group. This table presents the distribution of blood types within the male cohort of the ABO-BTI dataset, categorized across seven distinct age groups.

Blood Type	Newborns (0-28 days)	Babies (1-12 months)	Children (1-12 years)	Adolescents (13-18 years)	Young Adults (19-35 years)	Adults (36-64 years)	Seniors (65+ years)	Total
A+	1	1	2	1	2	1	0	8
A-	1	1	2	1	2	1	0	7
AB-	1	1	2	1	1	0	0	6
AB+	1	1	2	1	1	0	0	6
B-	1	1	2	1	1	0	0	6
B+	1	1	2	1	2	1	0	8
O-	1	1	2	1	1	1	0	7
O+	1	1	2	1	2	1	0	8
Total	8	8	15	8	12	5	0	56

Tab. 6: Blood Type Distribution Among Women by Age Group. This table presents the distribution of blood types within the female cohort of the ABO-BTI dataset, categorized across seven age groups.

Blood Type	Newborns (0-28 days)	Babies (1-12 months)	Children (1-12 years)	Adolescents (13-18 years)	Young Adults (19-35 years)	Adults (36-64 years)	Seniors (65+ years)	Total
A+	2	2	3	2	1	2	1	13
A-	1	1	2	2	4	2	1	12
AB-	1	1	3	2	3	2	0	12
AB+	1	1	3	2	3	2	1	13
B-	1	1	3	2	3	2	0	12
B+	2	2	3	2	4	2	1	16
O-	1	1	3	2	3	2	0	12
O+	2	2	3	3	4	3	1	18
Total	11	11	23	17	25	16	5	108

(11%), B- (11%), and O- (12%) demonstrate moderate prevalence, with AB- and B- being the least frequent at 11%. This distribution aligns with general population trends while emphasizing slight variations that may be attributed to regional genetic factors or sample demographics. Understanding such distributions is crucial for blood transfusion compatibility studies, clinical research, and the development of AI-based diagnostic models.

3) Cleaned Data

The preprocessing stage is a critical step in ensuring the accuracy, consistency, and reliability of the dataset. This process involved systematically addressing duplicate entries, removing poor-quality images, such as blurry ones, and resolving inconsistencies, such as misclassified blood types. These measures were essential to maintaining data integrity and ensuring that the dataset adhered to standardized demographic and clinical classifications.

Following the data cleaning process, the final dataset comprised 144 cases, consisting of 95 female and 49 male participants. To ensure uniformity in image representation, all collected images were resized to a fixed resolution of 1280×590 pixels. Standardizing image dimensions was necessary to maintain consistency across the dataset and facilitate efficient processing by deep learning models [26] [27].

In the data cleaning phase, we addressed the class imbalance in the ABO blood group image dataset to

ensure unbiased model performance for blood type classification. The initial dataset exhibited significant disparities across the A+, A-, AB+, AB-, B+, B-, O+, and O- classes, with sample sizes ranging from 18 cases and 54 images (AB-, B-) to 25 cases and 75 images (O+), totaling 164 cases and 492 images (Table 4). To resolve this imbalance, we implemented a downsampling strategy by aligning all classes with the smallest class size, which consisted of 18 cases and 54 images. During this process, we prioritized the removal of low-resolution images to maintain high-quality inputs for the model. This resulted in a balanced dataset with 18 cases and 54 images per class, totaling 144 cases and 432 images (Table 7).

Tab. 7: Blood Type Distribution by Sex

Blood Type	Women (66%)	Men (34%)	Total Cases
A+	12	6	18
A-	12	6	18
AB+	11	7	18
AB	12	6	18
B+	12	6	18
B	12	6	18
O	12	6	18
O+	12	6	18
Total	95	49	144

The pre- and post-balancing details are summarized in Table 4 and Table 7. This preprocessing step mitigated bias toward overrepresented classes, ensured consistent image quality, and enhanced the fairness and

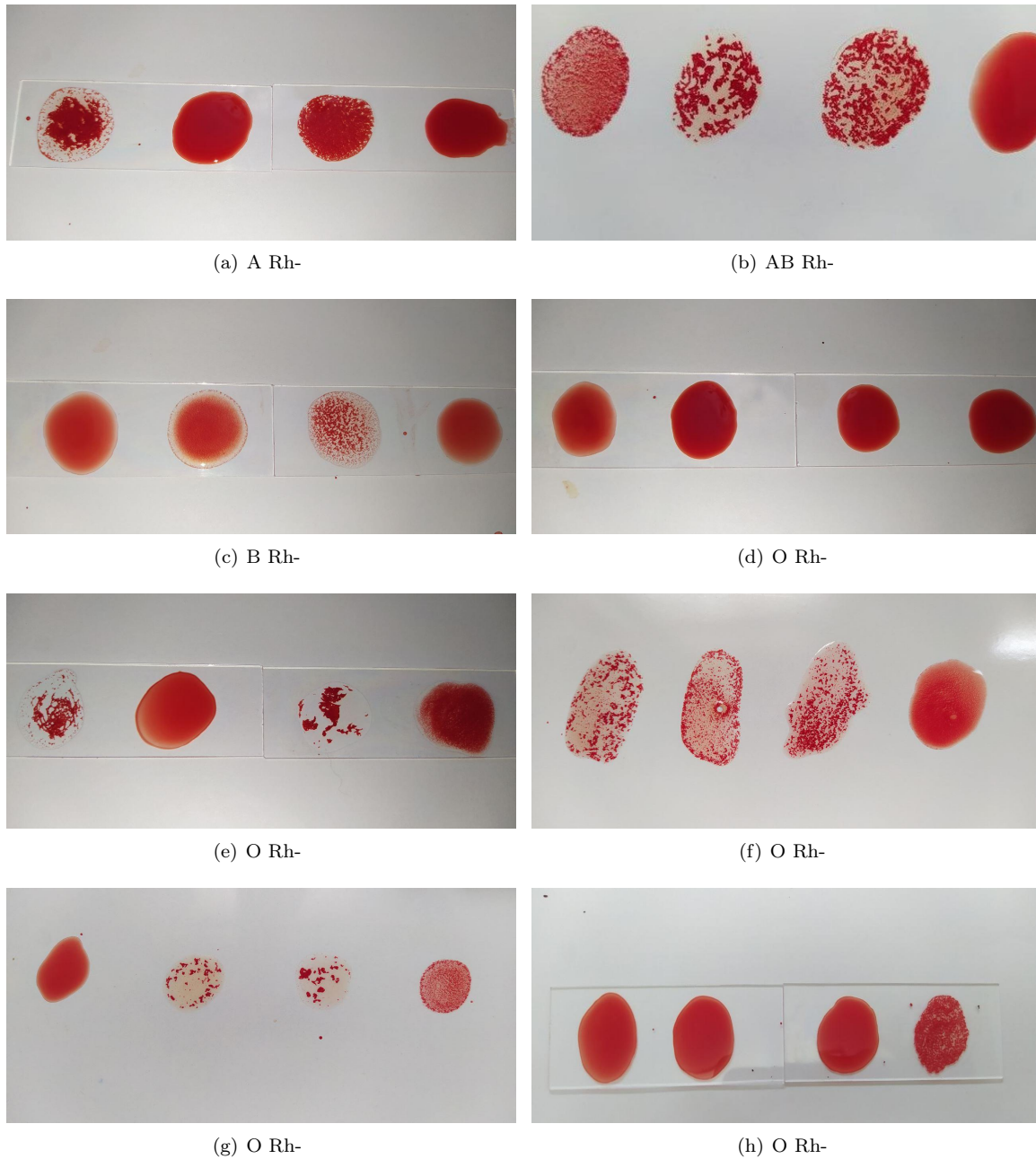


Fig. 4: Representative images from the dataset, illustrating samples from each blood type class after preprocessing. The figure showcases agglutination patterns corresponding to the eight primary blood type categories.

reliability of subsequent machine learning model training.

To provide a visual representation of the dataset composition, Figure 4 illustrates sample images from each blood type class. This figure serves as an important reference, highlighting the diversity and uniformity of the dataset across all categories.

By implementing these preprocessing steps, the dataset was refined to support reliable and reproducible experiments, enabling a rigorous assessment of its suitability for automated blood type classification.

The ABO-BTI dataset, including all high-resolution agglutination images and associated metadata, is publicly available for research purposes. The dataset can be accessed at: <https://doi.org/10.5281/zenodo.16971046>.

In order to protect donor privacy and comply with ethical protocols, individual-level metadata (e.g., exact age, gender, hospital ID) is not included in the public release. However, the aggregate demographic data displayed in Tables 5 and 6 were obtained from summary

Tab. 8: Age group distribution by gender. This table presents the demographic breakdown of participants in the ABO-BTI dataset, categorizing individuals into seven distinct age groups.

Age Group	Women	Men	Total
Newborns (0-28 days)	11	6	17
Babies (1-12 months)	11	6	17
Children (1-12 years)	22	11	33
Adolescents (13-18 years)	16	8	24
Young Adults (19-35 years)	27	14	41
Adults (36-64 years)	16	8	24
Seniors/Elderly (65+ years)	5	3	8
Total	108	56	164

records that are anonymized and were acquired with ethical approval and informed consent.

2.2. Preprocessing

Effective data preprocessing is crucial for ensuring the robustness and generalizability of machine learning models [28], particularly in medical imaging tasks such as automated blood type classification. The ABO-BTI dataset undergoes a series of preprocessing steps designed to normalize, augment, and enhance the training data while preserving critical diagnostic information. These preprocessing techniques not only improve model generalization by reducing overfitting but also ensure that the trained model can effectively handle real-world variations in blood agglutination images.

1) Normalization and Rescaling

To standardize pixel intensity values across all images, min-max normalization [29] applied by rescaling pixel values to the $[0,1]$ range. This is achieved by dividing each pixel intensity by 255, ensuring that all input values remain within a uniform scale. This normalization step is critical in deep learning models as it prevents numerical instability, accelerates convergence during training, and ensures that different images are processed on a consistent scale.

Mathematically, the rescaling operation is defined as follows:

$$I' = \frac{I}{255} \quad (1)$$

Where I represents the original pixel intensity (ranging from 0 to 255), and I' is the rescaled intensity in the range $[0,1]$.

2) Data Augmentation

To enhance the dataset's diversity and improve model generalization, extensive data augmentation is applied to the training images. Augmentation artificially expands the dataset by applying a series of random transformations, which simulate real-world variations in imaging conditions [32]. This process significantly reduces overfitting, enabling the model to learn robust and invariant features from blood agglutination patterns. The following augmentation techniques are incorporated into the preprocessing pipeline:

- **Random Rotation ($\pm 40^\circ$):** Introduces variability in the orientation of blood smear images, simulating natural variations in image capture.
- **Width and Height Shifts ($\pm 20\%$):** Offsets the images horizontally and vertically by up to 20% of their original dimensions to account for slight misalignments in blood sample imaging.
- **Shear Transformation (0.2 Factor):** Applies affine transformations to modify the geometric perspective of the images, reinforcing robustness against viewpoint variations.
- **Zooming ($\pm 20\%$):** Simulates differences in magnification by randomly zooming in or out, ensuring that the model remains scale-invariant.
- **Horizontal Flipping:** Introduces left-right symmetry by randomly flipping images, aiding in the detection of agglutination patterns regardless of their orientation.

These augmentation techniques are applied exclusively to the training dataset to enhance variability, while the test dataset remains unaltered to ensure that performance evaluation is conducted on unaltered, real-world images.

2.3. Model Development

To assess the suitability of the proposed ABO-BTI dataset for blood type classification, three deep learning models were implemented and evaluated: a custom CNN, ResNet50, and MobileNetV2. These models were selected to encompass a range of architectural complexities, from a relatively shallow CNN to deeper, more computationally intensive networks. The objective of this evaluation was to determine whether the dataset provides adequate feature representation to facilitate accurate blood type classification.

1) Proposed CNN Model

The proposed model is a custom CNN architecture, as shown in Table 9, designed to serve as a baseline for classification performance on the ABO-BTI dataset. It comprises three convolutional layers [33], each followed by LeakyReLU activation [34], max pooling [35], and dropout regularization to mitigate overfitting. The feature maps are subsequently flattened and passed through a fully connected layer, followed by a softmax activation function [36] to produce the final classification output.

This model was designed to test whether a relatively lightweight architecture is sufficient for distinguishing blood types based on the given dataset. The inclusion of dropout layers enhances generalizability by reducing reliance on specific neurons, thereby improving robustness against overfitting.

2) ResNet50

In this study, ResNet50 [30] was initialized with ImageNet-pretrained weights to benefit from prior feature representations learned from large-scale image datasets. The architecture was subsequently fine-tuned by modifying the final classification layers. The original fully connected layers were replaced with a Global Average Pooling (GAP) layer, followed by a fully connected layer with 1,024 neurons and ReLU activation. The final classification was performed using a softmax output layer with 8 neurons, corresponding to the number of blood type categories, as outlined in Table 10.

The inclusion of ResNet50 in this study allows for an assessment of whether a deep, highly expressive model provides significant performance improvements over shallower architectures when trained on the ABO-BTI dataset.

3) MobileNetV2

Similar to ResNet50, MobileNetV2 [31] was initialized with ImageNet-pretrained weights to leverage transferable feature representations. To adapt the model for the ABO-BTI dataset, the original classification head was replaced with a GAP layer, followed by a fully connected layer with 1,024 neurons and ReLU activation. The final classification was performed using a softmax output layer with 8 neurons, ensuring compatibility with the dataset's categorical labels, as detailed in Table 11.

The implementation of MobileNetV2 in this study facilitates an evaluation of lightweight neural networks in medical AI applications. By assessing its performance in comparison to deeper architectures such

as ResNet50, this analysis provides valuable insights into the feasibility of deploying efficient and memory-conscious blood type classification models in real-world, low-power computing environments.

3. Results and Discussion

This section presents the experimental results of the proposed deep learning model and the transfer learning architectures evaluated for ABO-BTI blood typing classification. The evaluation includes the experimental setup, training configuration, performance metrics, and confusion matrix analysis.

3.1. Experimental Setup

The computational environment was designed to leverage both CPU and GPU resources for optimized performance. The hardware setup included an Intel Core i3-10100F processor with four cores and eight threads, along with an NVIDIA GTX 1650 graphics card featuring 4GB of GDDR6 VRAM. The system was equipped with 8GB of DDR4 RAM and operated on Linux Mint 22.1 (64-bit). The GTX 1650, while categorized as an entry-level GPU, provided CUDA-enabled acceleration for deep learning computations. Given the constraints of 4GB VRAM, careful optimization was required, including batch size adjustments and memory-efficient training strategies to prevent resource exhaustion which will be discussed later on this section.

The software environment was designed to ensure compatibility with deep learning frameworks and maximize computational efficiency. Python 3.12.3 was used as the primary programming language. The deep learning models were implemented using TensorFlow 2.18.0, which was optimized for GPU acceleration through CUDA 12.4. The integration of TensorFlow with CUDA allowed for significant reductions in training time compared to CPU-based execution.

3.2. Training Configuration

The deep learning models were trained using an 80/20 data split, where 80% of the dataset was allocated for training and 20% for validation. This partitioning ensures that the models are exposed to a sufficiently diverse set of samples during training while preserving an independent subset for evaluating generalization performance. Given the hardware constraints, particularly the GTX 1650 GPU with 4GB GDDR6 VRAM, specific optimizations were implemented to efficiently manage memory usage while maintaining high model performance.

Tab. 9: Proposed CNN Architecture for Blood Type Classification.

Layer	Type	Output Shape	Parameters
1	Conv2D (32, 3×3) + LeakyReLU (0.1)	(118, 256, 32)	896
2	MaxPooling2D (2×2)	(59, 128, 32)	0
3	Dropout (0.25)	(59, 128, 32)	0
4	Conv2D (64, 3×3) + LeakyReLU (0.1)	(59, 128, 64)	18,496
5	MaxPooling2D (2×2)	(30, 64, 64)	0
6	Dropout (0.25)	(30, 64, 64)	0
7	Conv2D (128, 3×3) + LeakyReLU (0.1)	(30, 64, 128)	73,856
8	MaxPooling2D (2×2)	(15, 32, 128)	0
9	Dropout (0.4)	(15, 32, 128)	0
10	Flatten	(61440)	0
11	Dense (128) + LeakyReLU (0.1)	(128)	7,864,448
12	Dropout (0.3)	(128)	0
13	Dense (8, softmax)	(8)	1032

Tab. 10: Modified Classification Layers in ResNet50.

Layer	Description
Initial Weights	ImageNet Pretrained Weights
Global Average Pooling (GAP)	Reduces feature maps to a single value per channel
Fully Connected Layer	1,024 neurons, ReLU activation
Softmax Output Layer	8 neurons (corresponding to blood type classes)

Tab. 11: Modified Classification Layers in MobileNetV2.

Layer	Description
Initial Weights	ImageNet Pretrained Weights
Global Average Pooling (GAP)	Reduces feature maps to a single value per channel
Fully Connected Layer	1,024 neurons, ReLU activation
Softmax Output Layer	8 neurons (corresponding to blood type classes)

To standardize input dimensions, ResNet50 and MobileNetV2 were trained using input images resized to 224×224 pixels, whereas the proposed model was trained with images resized to 118×256 pixels. These resizing strategies were chosen to optimize computational efficiency while ensuring compatibility with the architectures' input layer requirements.

To enhance model robustness and prevent overfitting, data augmentation techniques were applied to the training set using the ImageDataGenerator class. The augmentation pipeline included random rotations ($\pm 40^\circ$), width and height shifts ($\pm 20\%$ of image dimensions), shear transformations ($\pm 20\%$), zoom variations ($\pm 20\%$), and horizontal flipping. Additionally, all input images were normalized by rescaling pixel values to the range [0,1]. These transformations increased dataset variability, allowing the models to learn invariant features and improve generalization.

The training batch size was set to 8, primarily due to the limited VRAM capacity of the GTX 1650 GPU. This batch size was chosen to prevent memory overflow while ensuring stable model training. The categorical cross-entropy loss function was employed for multi-

class classification, with the models producing probability distributions across blood type categories. To address class imbalances, class weights were computed dynamically using the balanced class weight strategy, ensuring that underrepresented blood types had a proportional influence on the optimization process.

The models were compiled using the Adam optimizer with a learning rate of 1×10^{-4} . The choice of Adam as the optimization algorithm ensures adaptive learning rates for efficient convergence, while the selected learning rate balances fast convergence with stable gradient updates. Training was conducted for 20 epochs, with the number of steps per epoch dynamically determined based on dataset size.

The validation set underwent the same preprocessing pipeline as the training data but without augmentation, ensuring an unbiased evaluation of model performance. Moreover, shuffling was disabled for the validation generator to maintain label consistency during performance assessment.

Tab. 12: Performance comparison of ResNet50, MobileNetV2, and the proposed model in terms of accuracy, precision, recall, and F1-score.

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Proposed CNN	96.90 $\pm$ 0.67	97.24 $\pm$ 0.55	96.90 $\pm$ 0.67	96.87 $\pm$ 0.67
ResNet50	70.50 $\pm$ 12.18	71.06 $\pm$ 15.75	70.50 $\pm$ 12.18	65.58 $\pm$ 14.60
MobileNetV2	10.34 $\pm$ 3.04	1.39 $\pm$ 0.18	10.34 $\pm$ 3.04	2.55 $\pm$ 0.37

Note: Values represent mean $\pm$ standard deviation from 3 independent runs.

3.3. Deep Learning Performance

The results presented in Table 12 demonstrate a substantial variation in classification performance among the three models, ResNet50, MobileNetV2, and the proposed model, when evaluated on the ABO-BTI dataset.

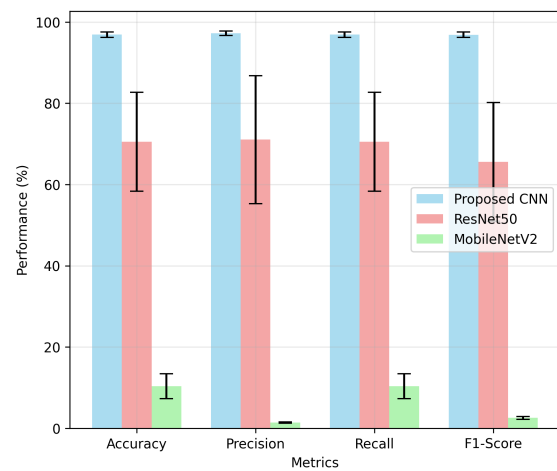
ResNet50 achieved an accuracy of 70.50 ± 12.18 , with a precision of 71.06 ± 15.75 and an F1-score of 65.58 ± 14.60 , indicating a moderate ability to differentiate between blood types. The recall value of 70.50 ± 12.18 suggests that while the model correctly identified a reasonable proportion of positive instances, it still encountered difficulties in distinguishing certain blood types. This performance aligns with expectations, as ResNet50 is a deep convolutional model with robust feature extraction capabilities. However, its computational requirements and architectural complexity may have limited its effectiveness in this specific classification task.

In contrast, MobileNetV2 exhibited significantly lower performance across all metrics, attaining an accuracy of 10.34 ± 3.04 , a precision of 1.39 ± 0.18 , a recall of 10.34 ± 3.04 , and an F1-score of 2.55 ± 0.37 . These results suggest that MobileNetV2 struggled to learn meaningful representations from the dataset, likely due to its lightweight design, which prioritizes computational efficiency over expressive feature extraction. The particularly low precision indicates a high frequency of misclassifications, while the poor recall suggests a general failure in correctly identifying blood type categories. The overall weak performance of MobileNetV2 underscores its limitations when applied to agglutination images that require more intricate feature extraction mechanisms.

The proposed model significantly outperformed both baseline architectures, achieving an accuracy of 96.90 ± 0.67 , with a precision of 97.24 ± 0.55 , a recall of 96.90 ± 0.67 , and an F1-score of 96.87 ± 0.67 . These results demonstrate the model's superior ability to capture and distinguish relevant features from the agglutination patterns, leading to a substantial improvement in classification accuracy. The high precision and recall indicate that the model is capable of correctly classifying blood types with minimal errors while maintaining a balanced distribution of correctly identified instances across different categories. The F1-score fur-

ther supports the model's reliability, confirming its robustness in handling the classification task with high consistency.

The comparative results highlight the importance of model selection in the context of blood type classification from agglutination images. While deeper architectures such as ResNet50 offer reasonable performance, their reliance on extensive computational resources may limit their practical applicability. On the other hand, lightweight models such as MobileNetV2, while efficient, may lack the representational capacity required for accurate classification in this domain. The proposed model demonstrates that a well-structured architecture, specifically optimized for this task, can achieve superior performance, making it a more suitable candidate for automated blood typing applications in clinical and laboratory settings.

**Fig. 5:** Performance comparison of Proposed CNN, ResNet50, and MobileNetV2 models on blood type classification using mean $\pm$ standard deviation.

The bar plot presented in Figure 5 illustrates the mean performance ($\pm$ standard deviation) of three models, proposed model, ResNet50, and MobileNetV2, on blood type classification across four key evaluation metrics across three runs. The Proposed model consistently outperforms the other models, achieving high scores (above 95%) with minimal variability across all metrics, indicating high reliability and effectiveness. In comparison, ResNet50 demonstrates moderate performance (approximately 70%) but with large error bars, reflecting greater inconsistency. MobileNetV2

performs significantly worse, with scores below 15% on all metrics and relatively high variance, suggesting limited suitability for this task.

3.4. Confusion Matrix Analysis

The confusion matrix, presented in Figure 6, illustrates the classification performance of MobileNetV2 across multiple classes. It is evident that the model consistently predicts a single class for all test samples, as indicated by the high concentration of values in one column. This behavior suggests a severe classification failure, likely due to suboptimal feature extraction or convergence issues during training. The absence of true positive classifications across multiple classes indicates that the model struggles to generalize effectively, potentially as a result of poor weight updates or insufficient feature separability in the dataset.

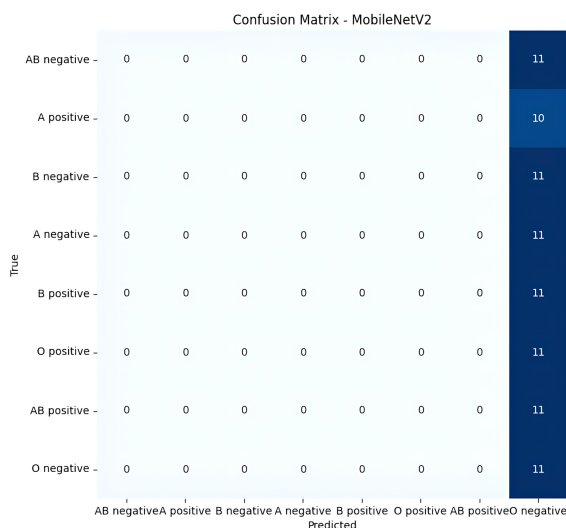


Fig. 6: Confusion matrix for MobileNetV2 classification results. The model exhibits a strong bias toward a single class, indicating poor generalization across multiple categories. This suggests potential class imbalance, inadequate feature representation, or convergence issues during training.

The training history, depicted in Figure 7, shows the variation in accuracy and loss over training epochs. The rapid decline in loss during the initial epochs, accompanied by a sharp increase in accuracy, suggests that the model quickly adapts to the training data. However, the high accuracy observed in training does not correspond to meaningful class separability, as evident in Figure 6. This discrepancy implies that the model may have overfitted to dominant patterns in the dataset while failing to learn robust class distinctions. Furthermore, the relatively stable accuracy curve in the later epochs suggests that the model has converged, but on a suboptimal solution that does not generalize well to unseen samples.

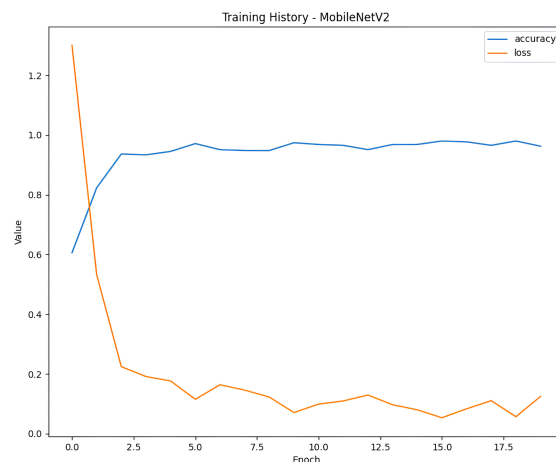


Fig. 7: Training history of MobileNetV2, showing accuracy and loss over epochs. The sharp decline in loss during the early epochs indicates rapid adaptation to training data. However, the high accuracy does not correspond to meaningful class separability, suggesting potential overfitting or inadequate learning of class-specific features.

MobileNetV2 is designed for efficiency in low-resource environments, prioritizing computational speed over deep feature extraction. While this makes it suitable for embedded systems such as the Raspberry Pi, it also limits its capacity to capture intricate agglutination patterns.

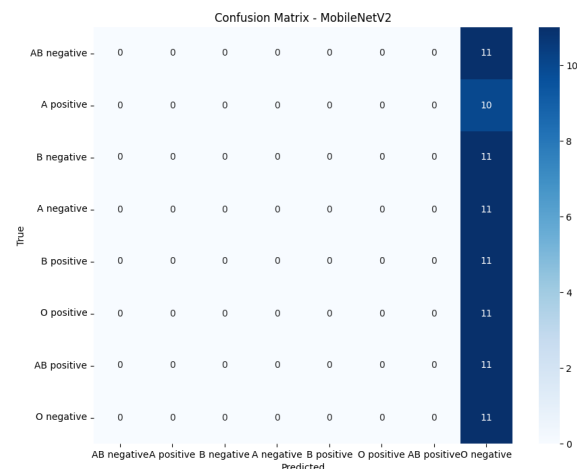


Fig. 8: Confusion matrix corresponding to the classification performance of the ResNet50 model. The matrix provides insights into the model's ability to correctly classify blood types, with the color intensity representing the frequency of each classification.

The performance of the ResNet50 model in the classification task can be assessed through the confusion matrix and training history. Figure 8 presents the confusion matrix, which provides insights into the model's ability to correctly classify different blood types. The diagonal elements of the matrix represent correct classifications, while off-diagonal elements indicate misclassifications. The results suggest

that the model achieves high accuracy for most blood types, particularly for B-negative, A-negative, and AB-positive, which exhibit strong diagonal values. However, some misclassifications are observed, notably for AB-negative, where a proportion of instances are classified as O-negative or AB-positive. This suggests potential challenges in distinguishing between these specific blood types, likely due to feature similarities in the learned representations.

The training dynamics of the ResNet50 model are illustrated in Figure 9, which depicts the evolution of accuracy and loss over training epochs. The accuracy curve shows a rapid increase in early epochs, stabilizing close to 1.0, indicating effective learning and convergence. Simultaneously, the loss function demonstrates a steep decline, which further corroborates the stability of the training process.

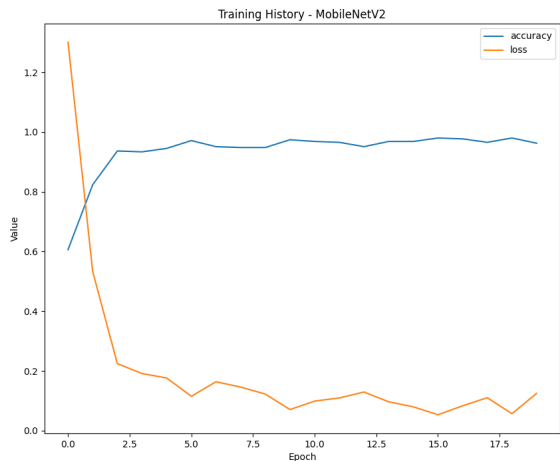


Fig. 9: Training performance of the ResNet50 model, showing the evolution of accuracy and loss over training epochs. The steady increase in accuracy and reduction in loss indicate the model’s convergence during training.

ResNet50, despite being a deeper and more expressive model, achieved only 72.41% accuracy, indicating difficulties in fully adapting to the ABO-BTI dataset.

The real effectiveness of the ABO-BTI database is assessed through the performance of the proposed model, as demonstrated by its confusion matrix (Figure 10) and training history (Figure 11). These results provide crucial insights into how well the dataset supports the development of a robust classification model for blood type identification. The following discussion elaborates on the implications of the confusion matrix and training history in evaluating the dataset’s suitability for deep learning-based classification.

The confusion matrix in Figure 10 illustrates the classification performance of the proposed model, offering valuable information regarding the dataset’s quality, class separability, and potential biases. The results indicate that the model exhibits a high degree of classification accuracy, with the majority of predictions

aligning correctly with the ground truth labels. The strong diagonal pattern observed in the matrix suggests that the ABO-BTI dataset provides sufficiently distinct and well-defined features for each blood type category, enabling the model to achieve accurate predictions.

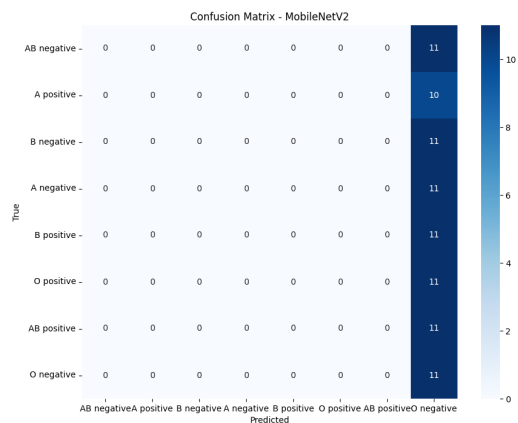


Fig. 10: Confusion matrix of the proposed model on the ABO-BTI dataset, illustrating the classification performance across different blood types. The high diagonal values indicate strong classification accuracy, with minimal misclassifications, demonstrating the dataset’s effectiveness in enabling clear class separability.

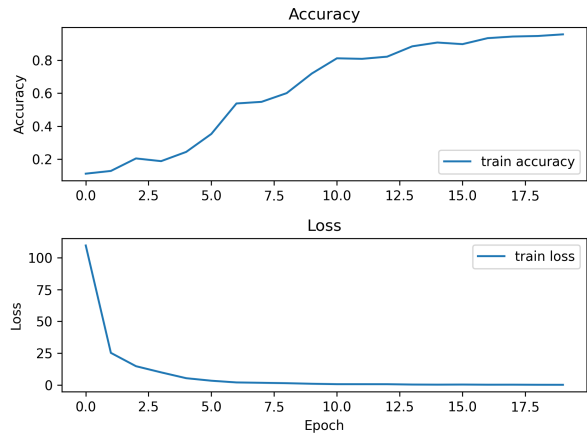


Fig. 11: Training history of the proposed model, showing accuracy (top) and loss (bottom) over multiple epochs. The steady increase in accuracy and smooth decrease in loss indicate efficient learning, suggesting that the ABO-BTI dataset provides high-quality and well-structured data for robust model training.

The high classification accuracy observed in the matrix indicates that the ABO-BTI dataset is structured in a manner that facilitates effective learning. Certain blood types, such as “A negative” and “AB positive”, demonstrate particularly strong classification performance, as evidenced by their high correct prediction counts. The minimal number of off-diagonal misclassifications further suggests that the dataset allows for

a clear differentiation between the blood type classes. For instance, “A negative” is correctly classified in all cases, while “AB positive” achieves a high recognition rate with only one misclassification. These findings highlight the effectiveness of the dataset in providing meaningful and discriminative features necessary for robust classification.

Despite the overall strong performance, a few instances of misclassification suggest the presence of subtle overlaps in feature representation between certain blood types. The confusion matrix reveals that “A positive” is misclassified as “A negative” in two cases, while “AB negative” is predicted as “A positive” twice. These misclassifications indicate that some blood type categories may share intrinsic similarities in the dataset, potentially due to overlapping physiological or morphological characteristics. While such minor errors are expected in biomedical classification tasks, their low frequency in this case suggests that the ABO-BTI dataset successfully minimizes ambiguities in feature space. The limited number of incorrect classifications further reinforces the notion that the dataset effectively supports accurate blood type identification.

The training history, as depicted in Figure 11, provides additional insights into the dataset’s effectiveness in facilitating deep learning-based blood type classification. The accuracy curve, shown in the upper panel, demonstrates a steady improvement in classification performance as training progresses. This trend indicates that the model is able to learn meaningful patterns from the ABO-BTI dataset without encountering significant challenges related to data inconsistencies or poor feature representation. The smooth and monotonic increase in accuracy suggests that the dataset is well-structured, allowing the model to gradually refine its predictive capabilities over successive epochs.

The training history further suggests that the dataset exhibits a balanced distribution across blood type classes, as evidenced by the stable and progressive learning curve. In scenarios where class imbalances are present, model accuracy often exhibits fluctuations or prolonged plateaus due to difficulty in learning underrepresented classes. However, the observed accuracy curve suggests that the ABO-BTI dataset does not suffer from severe class imbalance issues, enabling the model to generalize effectively across all blood type categories.

The loss curve, presented in the lower panel of Figure 11, corroborates these findings by demonstrating a consistent decline in loss values over the course of training. The sharp initial drop in loss values suggests that the ABO-BTI dataset provides clear and distinctive features, allowing the model to rapidly learn discriminative patterns in the early training stages. The subsequent gradual stabilization of loss values indicates

that the model converges effectively without encountering excessive noise or ambiguities in the training data. A well-behaved loss curve such as this is a strong indication that the dataset is well-labeled and contains high-quality feature representations.

A key factor contributing to the smooth training dynamics observed in the model is the high-quality annotation and feature representation in the ABO-BTI dataset. The absence of significant fluctuations in the loss curve suggests that the dataset contains minimal label noise, ensuring that the model learns from accurately labeled examples. In datasets with noisy or ambiguous labels, loss values often exhibit irregular fluctuations, leading to unstable training behavior. However, the observed stability in both accuracy and loss metrics confirms that the ABO-BTI dataset is effectively curated to support deep learning-based classification tasks.

The analysis of the confusion matrix and training history confirms that the ABO-BTI database is a highly effective resource for training deep learning models in blood type classification. The dataset’s high-quality feature representation, minimal noise, and balanced class distribution allow for robust and accurate model training. The minimal misclassification rates and stable learning curves further reinforce the notion that the dataset is well-suited for biomedical classification tasks. With a final accuracy of 96.51%.

The ABO-BTI database, despite its significance for blood type classification, presents several limitations that contribute to the poor performance of deep learning models such as ResNet50 and MobileNetV2. These architectures, which have demonstrated high performance on large-scale image classification tasks, struggle to achieve satisfactory results when applied to the ABO-BTI dataset. This discrepancy arises due to fundamental differences between the nature of the ABO-BTI data and the feature extraction mechanisms these models employ.

ResNet50 and MobileNetV2 are CNNs originally designed for processing natural images, where spatial patterns, edges, and textures provide critical information for classification. ResNet50, with its deep residual learning framework, is particularly adept at extracting hierarchical features, while MobileNetV2, optimized for efficiency, utilizes depthwise separable convolutions to reduce computational overhead. However, the nature of blood type data differs significantly from that of traditional image datasets. Blood type classification does not rely on complex spatial patterns, but rather on subtle biochemical and morphological features that may not be optimally captured by CNN-based architectures.

A key factor contributing to the poor performance of these models is the limited dataset size. Deep learn-

Tab. 13: Performance Comparison of Blood Type Classification Methods.

Method	Model	Feature Technique	ROI Shape	F1-Score (%)
Ferraz et al. [14]	SVM	HoG	Square	52.63
		HoG	Round	79.17
		FFT	Round	87.18
Proposed Method	Proposed Model	Deep Features	–	96.87 $\pm$ 0.67

Note: HoG = Histogram of Oriented Gradients, FFT = Fast Fourier Transform. ROI = Region of Interest. Dashes (–) indicate not applicable.

ing models, particularly those as complex as ResNet50, require vast amounts of training data to generalize effectively. The ABO-BTI database does not provide a sufficiently large dataset for these architectures to learn discriminative patterns without overfitting. Without adequate training samples, these models fail to establish robust decision boundaries, leading to suboptimal accuracy, precision, recall, and F1-score. In contrast, the proposed model, which is specifically designed for the nature of blood type classification, exhibits significantly improved performance, as it is better aligned with the characteristics of the ABO-BTI dataset.

Beyond the challenges posed to deep learning models, the ABO-BTI database itself presents inherent limitations that impact classification performance. One of the most critical issues is the high inter-class similarity between blood types. Unlike natural image classification tasks, where classes often have distinct visual features, the differences between blood types are more subtle, making it challenging for models to distinguish between them with high confidence. This can result in increased misclassification rates, particularly for architectures that rely on spatial differentiation.

Another limitation of the ABO-BTI database is the potential lack of representational diversity. If the dataset consists of a relatively homogeneous set of samples, models trained on it may fail to generalize to new, unseen examples. Effective blood type classification requires training data that encompasses a wide range of variations, including different sample conditions and potential measurement artifacts. A lack of diversity within the ABO-BTI dataset may lead to biased model predictions and reduced robustness in real-world applications.

Table 13 presents a performance comparison between the proposed blood type classification method and prior approaches by Ferraz et al. [14], highlighting differences in model architectures, feature extraction techniques, and ROI design. Ferraz et al.'s method, based on SVM classifiers and handcrafted features (HoG and FFT), achieved variable results depending on the ROI shape, with F1-Scores ranging from 52.63% (HoG with square ROI) to 87.18% (FFT with round ROI). In contrast, the proposed method, which integrates a custom deep learning model and deep fea-

ture representations, significantly outperforms earlier methods with an F1-score of 96.87 $\pm$ 0.67, while eliminating the need for explicit ROI design. This demonstrates the superiority and robustness of the proposed approach in accurately classifying blood types.

4. Conclusion

This study introduced the ABO-BTI database as a benchmark dataset for blood type classification and evaluated its effectiveness using deep learning models. The experimental results demonstrated that the proposed deep learning-based classification model achieved an accuracy of 96.51%, highlighting its potential for automated blood type identification. Compared to traditional machine learning approaches that rely on handcrafted feature extraction, the deep learning model successfully learned meaningful feature representations directly from the data, reducing the need for extensive preprocessing and domain-specific feature engineering. However, the comparative analysis with state-of-the-art methods revealed that handcrafted approaches using SVM with certain feature extraction techniques, such as histogram and FFT, still achieve higher classification accuracy in some cases.

The primary research question addressed in this study was whether deep learning models, when trained on the ABO-BTI database, could serve as reliable and effective tools for blood type classification. The findings confirm that while deep learning models can achieve high classification performance, they may still face challenges in outperforming optimized traditional methods under certain conditions. Nevertheless, the results support the viability of deep learning for this task and underscore the importance of dataset quality and representation in achieving optimal performance. By providing a standardized and publicly available dataset, this work lays the foundation for further advancements in automated blood type classification.

A key contribution of this research is the introduction of the ABO-BTI database, which fills a critical gap in the field by providing a dedicated dataset for blood type classification. Unlike previous studies that relied on smaller, private datasets, ABO-BTI offers a

structured and diverse dataset that can facilitate the development and benchmarking of new classification models. Additionally, the study presents an extensive evaluation of deep learning performance on this task, shedding light on both its strengths and limitations in comparison to conventional machine learning techniques. This work not only advances the field of automated blood classification but also contributes to the broader area of medical image analysis by demonstrating the potential of deep learning in biomedical classification tasks.

The findings of this study have several implications for both theoretical and practical applications. Theoretically, they contribute to the ongoing discourse on the effectiveness of deep learning in classification tasks where data variability and feature representation are critical. The results suggest that deep learning can generalize well to blood type classification, although it still benefits from careful dataset construction and pre-processing. From a practical perspective, the ability to automate blood type classification has significant applications in healthcare, particularly in emergency medical scenarios where rapid and accurate blood type identification is essential. Furthermore, as machine learning continues to integrate into medical diagnostics, the ABO-BTI database can serve as a valuable resource for researchers and practitioners seeking to refine classification techniques and enhance the robustness of predictive models.

Despite the promising results, certain limitations should be acknowledged. The study primarily evaluated deep learning models on the ABO-BTI dataset, which, while comprehensive, may still require further expansion to ensure greater generalization across diverse imaging conditions and population groups. Additionally, while deep learning models offer automation advantages, their dependency on large amounts of labeled training data remains a challenge, and their susceptibility to misclassification in certain cases warrants further investigation.

Future research should explore strategies to further enhance deep learning performance in blood type classification. Potential directions include incorporating hybrid models that combine deep learning with handcrafted feature extraction techniques, investigating self-supervised learning approaches to reduce dependency on labeled data, and leveraging explainability methods to interpret model decisions more effectively. Moreover, expanding the ABO-BTI dataset with additional samples and variations could further improve the robustness of classification models, making them more suitable for real-world deployment.

Beyond traditional machine learning applications, future work should focus on the deployment of blood type classification models in embedded systems. The

integration of deep learning-based blood type classification into low-power embedded platforms could enable real-time and on-site blood type identification, which would be particularly beneficial in resource-constrained medical environments, ambulatory emergency units, and military field hospitals. Edge AI devices, such as those using TinyML frameworks, could facilitate fast and efficient blood classification directly from mobile diagnostic tools without requiring cloud-based processing. Furthermore, wearable health monitoring systems and point-of-care testing devices could incorporate automated blood typing to streamline transfusion compatibility assessments in decentralized healthcare settings. These advancements would significantly enhance the accessibility and responsiveness of blood type identification, ultimately improving patient outcomes in critical situations.

Acknowledgment

We would like to thank the healthcare institutions in Annaba, Algeria: Centre Hospitalo-Universitaire Dr Dorban and Abdallah Nouaouria El Bouni Specialized Hospital, for their valuable assistance in the collection of the database.

Author Contributions

S. D. developed the theoretical methodology, carried out data analysis, supervised the research, contributed to the conception and design of the study, and finalized the editing of the manuscript. H. Z. implemented the computational models, performed the experiments, and contributed to the drafting of the manuscript. A. C. participated in the data pre-processing and conducted the literature review. T. H. contributed to manuscript revision, validation of results, and editing of the manuscript. C. H. contributed to the collection of data.

References

- [1] WESTHOFF, C. M., A. FLOCH. Blood group genotype matching for transfusion. *British Journal of Haematology*. 2024, vol. 206, iss. 1, pp. 18-32. DOI: 10.1111/bjh.19664.
- [2] WEI, H., et al. CRISPR/Cas13a-based single-nucleotide polymorphism detection for reliable determination of ABO blood group genotypes. *The Analyst*. 2024, iss. 7. DOI: 10.1039/d3an02248j.

- [3] LI, H.-Y. and K. GUO. Blood Group Testing. *Frontiers in Medicine*. 2022, vol. 9. DOI: 10.3389/fmed.2022.827619.
- [4] MATZHOLD, E., et al. Characterization of Novel RHD Allele Variants and Their Implications for Routine Blood Group Diagnostics. *Biomedicines*. 2024, vol. 12, no. 2. DOI: 10.3390/biomedicines12020456.
- [5] BROWN, M., et al. Incorrect blood typing and mis-transfusion due to low-titer group O whole blood resuscitation. *Transfusion*. 2025. DOI: 10.1111/trf.17840.
- [6] PANCH, S. R., C. MONTEMAYOR. Hemolytic transfusion reactions. In: *Rossi's Principles of Transfusion Medicine*. 5th ed. Wiley-Blackwell, 2022, pp. 543–552.
- [7] DEVECI, B., et al. Severe Acute Hemolytic Transfusion Reaction Treated with Ruxolitinib and Plasma Exchange. *Transfusion Medicine and Hemotherapy*. 2021, vol. 48, no. 4, pp. 250–253. DOI: 10.1159/000513056.
- [8] GEHRIE, E., B. SAVANI, G. S. BOOTH. Risk factors for hemolytic transfusion reactions resulting from ABO and minor red cell antigen incompatibility: From mislabeled samples to stem cell transplant and sickle cell disease. *Blood Reviews*. 2021, vol. 45. DOI: 10.1016/j.blre.2020.100719.
- [9] HASAN, R., et al. Rates of delayed hemolytic transfusion reactions observed in a trauma center. *Transfusion Medicine*. 2021, vol. 61, pp. 2035–2040. DOI: 10.1111/trf.16433.
- [10] ZHOU, Y., et al. The Potential Significance of ABO Genotyping for Donor Selection in Kidney Transplantation. *Frontiers in Immunology*. 2020, vol. 11. DOI: 10.3389/fimmu.2020.608716.
- [11] MENECHINI, M., et al. On the clinical relevance of using complete high-resolution HLA typing for an accurate interpretation of posttransplant immune-mediated graft outcomes. *Frontiers in Immunology*. 2022, vol. 13. DOI: 10.3389/fimmu.2022.924825.
- [12] SKLAVOUNOS, A. A., et al. Digital Microfluidic Hemagglutination Assays for Blood Typing, Donor Compatibility Testing, and Hematocrit Analysis. *Clinical Chemistry*. 2021, vol. 67, iss. 12, pp. 1699–1708. DOI: 10.1093/clinchem/hvab180.
- [13] WU, P. C., et al. Epigenetic dissection of human blood group genes reveals regulatory elements and detailed characteristics of KEL and four other loci. *Transfusion*. 2024, vol. 64, pp. 1083–1096. DOI: 10.1111/trf.17840.
- [14] FERRAZ, A., J. H. BRITO, V. CARVALHO, J. MACHADO. Blood type classification using computer vision and machine learning. *Neural Computing and Applications*. 2017, vol. 28, pp. 2029–2040. DOI: 10.1007/s00521-015-2151-1.
- [15] GOLDBERGER, A. L., et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation*. 2000, vol. 101, no. 23, pp. e215–e220. DOI: 10.1161/01.cir.101.23.e215.
- [16] CLARK, K., et al. The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. *Journal of Digital Imaging*. 2013, vol. 26, pp. 1045–1057. DOI: 10.1007/s10278-013-9622-7.
- [17] SADAFI, A., et al. A continual learning approach for cross-domain white blood cell classification. In: *MICCAI Workshop on Domain Adaptation and Representation Transfer*. 2023, pp. 136–146. DOI: 10.1007/978-3-031-45857-6_14.
- [18] MATEK, C., S. SCHWARZ, K. SPIEKERMANN, C. MARR. Human-level recognition of blast cells in acute myeloid leukaemia with convolutional neural networks. *Nature Machine Intelligence*. 2019, vol. 1, no. 11, pp. 538–544. DOI: 10.1038/s42256-019-0101-9.
- [19] ACEVEDO, A., et al. A dataset of microscopic peripheral blood cell images for development of automatic recognition systems. *Data in Brief*. 2020, vol. 30. DOI: 10.1016/j.dib.2020.105474.
- [20] NITHYA, R., K. NIRMALA. Detection of Anaemia using Image Processing Techniques from microscopy blood smear images. In: *Journal of Physics: Conference Series*. 2022, vol. 2318, no. 1. DOI: 10.1088/1742-6596/2318/1/012043.
- [21] SHAHZAD, M., et al. Identification of anemia and its severity level in a peripheral blood smear using 3-tier deep neural network. *Applied Sciences*. 2022, vol. 12, no. 10. DOI: 10.3390/app12105030.
- [22] BUKHARI, M., S. YASMIN, S. SAMMAD, A. A. ABD EL-LATIF. A deep learning framework for leukemia cancer detection in microscopic blood samples using squeeze and excitation learning. *Mathematical Problems in Engineering*. 2022, vol. 2022, no. 1. DOI: 10.1155/2022/2801227.
- [23] ROMANOS-SIRAKIS, E. C., D. DESAI. ABO Blood Group System. [Updated 2025 Apr 26]. In: <https://www.ncbi.nlm.nih.gov/books/NBK580518/>.
- [24] FALCON, R. M. G., F. M. M. CLIMACOSA, S. E. C. CAOILI. Agglutination of yeast-binding

- antibodies from human blood plasma products. *BMC Microbiology*. 2025, vol. 25, no. 1. DOI: 10.1186/s12866-025-03457-2.
- [25] WORLD HEALTH ORGANIZATION. *Blood transfusion safety: Guidelines on assuring the quality of blood grouping, compatibility testing and blood component preparation*. Geneva: World Health Organization; 2021. Available from: <https://www.who.int/publications/i/item/9789240032408>.
- [26] VO, V.-A., V.-D. PHAN, V.-M. BUI, T.-N. DO. Design of Deep Learning Model Applied For Smart Parking System. *Advances in Electrical and Electronic Engineering*. 2023, vol. 21, no. 4. DOI: 10.15598/aeec.v21i4.5366.
- [27] PRASAD, A. G., et al. A Hybrid Predictive Architecture Formulation Using Deep Learning and Histogram of Gradients for Compound Emotion Recognition. *Advances in Electrical and Electronic Engineering*. 2024, vol. 22, no. 1. DOI: 10.15598/aeec.v22i1.5467.
- [28] CABELLO-SOLORZANO, K., et al. The impact of data normalization on the accuracy of machine learning algorithms: A comparative analysis. In: *International Conference on Soft Computing Models in Industrial and Environmental Applications*. 2023, pp. 344–353. DOI: 10.1007/978-3-031-42536-3_33.
- [29] ALI, P. J. M. Investigating the Impact of Min-Max Data Normalization on the Regression Performance of K-Nearest Neighbor with Different Similarity Measurements. *ARO-The Scientific Journal of Koya University*. 2022, vol. 10, no. 1, pp. 85–91. DOI: 10.14500/aro.11027.
- [30] HE, K., X. ZHANG, S. REN, J. SUN. Deep Residual Learning for Image Recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 770–778. DOI: 10.1109/CVPR.2016.90
- [31] SANDLER, M., et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018, pp. 4510–4520. DOI: 10.1109/CVPR.2018.00474.
- [32] NGUYEN, T. T., T.-H. NGUYEN, B.-V. NGO, T.-N. NGUYEN. Largest ROI Segmentation for Breast Cancer Classification Using a VGG16 Deep Learning Network. *Advances in Electrical and Electronic Engineering*. 2024, vol. 22, no. 4. DOI: 10.15598/aeec.v22i4.240303.
- [33] O'SHEA, K. and R. NASH. An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*. 2015.
- [34] DUBEY, A. K. and V. JAIN. Comparative study of convolution neural network's relu and leaky-relu activation functions. In: *Applications of Computing, Automation and Wireless Systems in Electrical Engineering: Proceedings of MARC 2018*. 2019, pp. 873–880. DOI: 10.1007/978-981-13-6772-4_76.
- [35] GHOLAMALINEZHAD, H., H. KHOSRAVI. Pooling methods in deep neural networks, a review. *arXiv preprint arXiv:2009.07485*. 2020.
- [36] BRIDLE, J. S. Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition. In: *textitNeurocomputing: Algorithms, Architectures and Applications**. Springer. 1990, pp. 227–236. DOI: 10.1007/978-3-642-76153-9_28.