Advances in Science and Technology Research Journal

Advances in Science and Technology Research Journal, 2026, 20(1), 510–524
https://doi.org/10.12913/22998624/210475
ISSN 2299-8624, License CC-BY 4.0

Received: 2025.09.01
Accepted: 2025.10.08
Published: 2025.11.21

Artificial intelligence: Assisted fundus image analysis for medical diagnostics in conflict zones

Magdalena Matysiak[1], Arkadiusz Podkowiński[2], Tomasz Chorągiewicz[3], Robert Karpiński[4*], Michał Dolecki[5], Rafał Stęgierski[5], Andriy Zimenkovsky[6], Volodymyr Shybinskyi[6], Katarzyna E. Jonak[7], Robert Rejdak[3]

[1] Department of Technical Computer Science, Lublin University of Technology, Nadbystrzycka 38, 20-618 Lublin, Poland
[2] Da Vinci NeuroClinic, Tomasza Zana 11a, 20-601 Lublin, Poland
[3] Chair and Department of General and Pediatric Ophthalmology, Medical University of Lublin, Aleje Archaicize 1, 20-059 Lublin, Poland
[4] Department of Machine Design and Mechatronics, Lublin University of Technology, Nadbystrzycka 38, 20-618 Lublin, Poland
[5] Lublin University of Technology, Department of Computational Intelligence, ul. Nadbystrzycka 38, 20-618 Lublin, Poland
[6] Department of Health Care Management, Pharmacotherapy & Clinical Pharmacy, Danylo Halytsky Lviv National Medical University, Ukraine
[7] Doctoral School, The John Paul II Catholic University of Lublin, Al. Racławickie 14, 20-950 Lublin, Poland
* Corresponding author's e-amil: r.karpinski@pollub.pl

ABSTRACT

Artificial intelligence (AI) has become an important tool for recognizing changes in the ocular fundus, but most existing studies are conducted in peacetime clinical environments with advanced diagnostic equipment and stable infrastructure. In contrast, wartime conditions impose severe constraints, including limited access to sophisticated imaging devices, reduced medical resources, and the urgent need for rapid decision-making. This article addresses this research gap by examining AI-assisted classification of retinal fundus images collected under conflict conditions in Ukraine. Three approaches were employed: feature extraction combined with deep neural networks, convolutional neural network (CNN)-based models, and Microsoft's Custom Vision platform. The dataset consisted of 448 retinal images divided into five groups: normal findings, trauma-related injuries, optic nerve disc changes, vascular lesions, and macular degeneration. Despite the small and imbalanced dataset, and the challenging acquisition environment, each pre-processing method achieved at least 80% classification accuracy, with the CLAHE method yielding the best results. This study demonstrates, for the first time, that AI can provide reliable ophthalmic diagnostics in extreme and resource-limited wartime settings, bridging the gap between peacetime and conflict healthcare.

Keywords: retinal fundus imaging, wartime healthcare, automated diagnosis, ophthalmology, artificial intelligence.

INTRODUCTION

The ongoing armed conflict has caused a dramatic increase in traumatic injuries, creating unprecedented challenges for healthcare systems. Eye injuries are particularly concerning, as they often lead to irreversible vision loss if not diagnosed and treated quickly [1]. While artificial intelligence (AI), including machine learning (ML) and deep learning (DL), has shown great promise in ophthalmology, particularly in analyzing fundus and optical coherence tomography (OCT) images, existing studies have largely been conducted under controlled laboratory or hospital conditions. This leaves a significant research gap: little is known about the feasibility of

AI-based diagnostic tools in resource-limited and wartime settings, where conventional equipment is unavailable and medical staff face urgent time constraints. Our study addresses this gap by evaluating AI-assisted fundus image analysis in real-world wartime conditions in Ukraine, highlighting both the challenges and the potential clinical value of such tools in emergency and field medicine.

Military operations put a strain on the health care system. New military technologies contribute to permanent damage to health. The health care system becomes inefficient and helping the injured is difficult. Eyes are exposed to a wide range of damage from impact, heat, chemistry, radiation, or laser. The rise of new military technologies is accelerating the increase in eye injuries. The most common are penetrating and perforating injuries, often accompanied by intraocular diseases that create specific diagnostic and treatment challenges. Modern weapons are great risk for ophthalmic injuries. Soldiers and civilians are at risk of injury from ammunition fragments, explosions and chemical burns during armed conflicts [2]. Technology and the growing advancement of weapons increases combat effectiveness, but also causes more damage, including to the eye. In the research, we want to develop faster methods of diagnosing eye diseases that will save the eyesight of the injured even in difficult conditions. By applying artificial neural network methods, diagnostics can be performed more rapidly and with greater accuracy, enabling physicians to provide timely and effective treatment for patients with injuries [3, 4].

In contrast to peacetime ophthalmic diagnostics, which typically benefit from stable infrastructure, advanced imaging technologies such as OCT, and access to specialized medical teams, wartime conditions impose severe limitations. Medical professionals must often rely on portable, less sophisticated equipment, operate under time pressure, and deliver care in environments where resources and safety are compromised. These constraints create a distinct research gap, as most prior AI-based ophthalmology studies have been conducted in controlled, resource-rich settings. Our study addresses this gap by evaluating the feasibility and effectiveness of AI-assisted fundus diagnostics under wartime conditions, providing insights into how AI can support medical decision-making in extreme and resource-limited environments.

Artificial Intelligence is a multidimensional technology that encompasses various components, including advanced algorithms, ML, and DL. DL employs representation-learning techniques with multiple levels of abstraction, enabling the processing of input data without manual feature engineering, for example in tasks such as recognizing complex structures [5, 6]. Compared to conventional techniques, DL provides significantly higher accuracy which can be used for medical imaging analysis with highly effective results in detection of various diseases. In ophthalmology, DL is most commonly used for fundus analysis and optical coherence tomography (OCT) [7]. In fact, with the advent of the Internet, ML has become an important part of the information revolution. AI, ML, DL is expected to provide ophthalmologists with an automated device for early diagnosis and timely treatment [8].

The study highlights that hybrid machine learning models can significantly accelerate the diagnosis of eye diseases using OCT images. However, a major challenge remains the limited availability of costly ophthalmic equipment in conflict zones. The lack of appropriate equipment makes diagnosis difficult. Correct diagnosis is key to introducing appropriate treatment. It is not just the price that matters. The size of the medical device is also important. Equipment must be lightweight and mobile, must be ready to be transported at any time due to different situations on the front line [9]. Previous studies [10, 11] have shown that neural networks can substantially enhance medical diagnostics. These studies focus on the accurate analysis of OCT images. Although OCT imaging offers greater diagnostic accuracy than fundus photography, its high cost and limited availability significantly constrain its use in resource-limited settings. Studies with fundus cameras, despite their lower resolution and accuracy are generally available. The availability and ease of taking images is significant during difficult conditions [12].

Research shows that the combination of artificial intelligence and retinal images can improve diagnostics. Easier diagnosis can make it significantly easier for doctors to diagnose a disease. The main goal of this study is to develop diagnostic tools that are reliable and easy to use in places with limited hospital access [13]. In the recent years, the combination of neural networks (CNN) and vision transformers (ViT) has become increasingly important medical imagine analysis. Combining these two methods has helped improve the precision of diagnosing ophthalmologic diseases. The article concentrated on triple stream features from OCT images which had advantages over single stream models. The triple stream had very high test accuracy [14].

The fast development of AI technology enables quicker and simpler diagnostics. Most of the research is based on laboratory conditions. This research is innovative because it is conducted in war conditions. The new image analysis techniques introduced in this article help doctors make rapid diagnoses. Combining camera fundus with AI can solve many problems. In recent years, advances in artificial intelligence (AI) have significantly accelerated the development of image analysis, thereby facilitating more efficient and accurate medical image interpretation.

The main aim of the study is to develop a new method for diagnosing different eye injuries using convolution neural networks (CNN). The use of advanced image analysis techniques that allow consideration of texture and intensity based on images obtained from the fundus camera allows significant improvements in the accuracy of injury classification. Modern image processing methods, based on ML and deep learning algorithms, allow more precise extraction of features characteristic of different types of injuries. Texture analysis makes it possible to identify even very subtle changes in tissue structure, which can indicate injuries even in the form of microdamage, as well as degenerative changes or pathological processes. These tools are designed to operate quickly and effectively, assisting physicians in diagnosing groups of eye injuries, including normal findings, trauma-related injuries, optic nerve disc changes, vascular lesions, and macular degeneration. This means that many injured people are not given the chance for a quick diagnosis, which can lead to permanent vision loss. We want to create a simpler and more easily accessible tool that can detect eye damage, even in the difficult conditions of war. This technology enables accurate verification of specific eye injuries. Doctors will be able to implement better diagnostics and faster treatment.

The remainder of this article is organized as follows. Next section presents the materials and methods, including details on patient data, imaging devices, and machine learning approaches. Then we report the experimental results obtained from the three AI-based classification methods. In next section we discuss the findings in the context of existing literature, highlights limitations such as dataset size and imbalance, and outlines future research directions. Final section concludes the study by emphasizing the unique contribution of AI-assisted fundus diagnostics in wartime conditions compared to conventional peacetime healthcare.

MATERIALS AND METHODS

The war in Ukraine has exposed a new threat. The hybrid war on Ukraine carries with it the direct threat of using unconventional weapons. Modern military technologies have an impact on public health. The authors of the article [15] discuss the immediate as well as the long-term consequences. The majority of patients at risk were men between 25 and 63 years of age. The most common injuries were to the stomach, face and limbs. Limited transportation conditions made it difficult to help the injured. The ongoing conflict poses new challenges for health care workers, who must be prepared to deal with the complexities of dealing with victims [16].

Information about patients and location

The research was conducted during the war in Ukraine. The analyses included 570 patients who presented with eye injuries. The majority of eye injuries resulted from bombings, shrapnel, and battlefield clashes. Participants were recruited at the Lviv Regional Clinical Hospital. The patients then had specialized eye examinations. Patients whose health status allowed for full medical record were examined first. Each patient signed an informed consent for the eye examination. Ethical approval was granted by the Bioethics Committee of the Medical University of Lublin. The criteria for including patients in the study is that participants had to have experienced an eye injury directly to military operations and be in stable health during the study. Patients were excluded from the study if their condition was unstable or if they did not provide informed consent. Reported cases included eye injuries, orbital trauma, and post-traumatic ocular pain. Some patients presented with only superficial injuries, with no detectable changes in the ocular fundus.

Data acquisition

The fundus images were primarily collected using Optomed Aurora portable cameras, which are independent, lightweight devices capable of producing diagnostic-quality images (50-degree field of view, 5-megapixel resolution) without requiring a computer. These cameras were used both in hospital settings and in more challenging wartime environments. Although the equipment was consistent, there were introduced effects knows from the wartime conditions, such as variable lighting, operator experience, and urgency of examinations, introduced heterogeneity in image quality, resolution, and color distribution. This explains why, despite standardized hardware, the dataset displays significant variability. Images were saved in JPEG format to ensure data integrity and facilitate straightforward analysis [17].

Patients were dropped with 1% tropicamide solution. When the eyes were dropped, the pupil dilated so that better photos of the fundus could be taken. Patients were analyzed on a scale of 0 (no rating) – 2 (excellent). Images that do not meet standards for clarity, centering or visibility of key structures are excluded. This rigorous approach ensures good data quality. The data has been anonymized in accordance with Polish and Ukrainian data protection regulations (RODO). Visual acuity is measured on a decimal scale and compared with the Eye Trauma Score. Patient demographics such as age, gender, profession and social status were taken.

Machine learning

Warfare comes with significant challenges, such as limited equipment and incomplete data sets. Adaptation of neural networks in terms of standardizing image analysis and adapting to different input data reduces these difficulties. This adaptation is particularly important for analyzing results under difficult conditions.

The research used machine learning. CNN were mainly used to analyze images of the eye. CNN automate the image analysis process and classify fundus injuries. Using these advanced algorithms, it was possible to reduce diagnostic time and increase accuracy in condition that require quick decision making [18, 19].

This study highlights their potential for developing automated diagnostic tools that support physicians in making faster and more accurate diagnoses. Studies have been conducted to enable early detection and timely treatment using deep learning algorithms for fundus images. Quick diagnosis and treatment planning can be made easier with the Deep learning techniques have revolutionized medical imaging, markedly enhancing the accuracy of eye disease detection and classification. ability of deep learning models to process images quickly and deliver results immediately. Our research aims to provide a non-invasive method for early detection and rapid treatment of eye diseases using a CNN. This article presents practical applications of artificial intelligence for the analysis of various retinal injuries under challenging conditions. Furthermore, it discusses specific applications of AI in the classification of retinal diseases [20–22].

RESULTS

The analyzed dataset contained 448 retinal images categorized into five groups: normal images, trauma-related injuries, optic disc changes, vascular changes, and macular degeneration. The images were collected under extremely challenging conditions during wartime using various devices, resulting in images with varying resolutions and inconsistent color distributions. The initial resolution of the images ranged from 436 × 333 pixels to 2368 × 1776 pixels. The color histograms for six randomly selected images from class normal only are presented in Figure 1.

Additionally, the individual image classes contain different numbers of images, with the normal class having 25 images, trauma-related injuries 24, optic disc changes 192, vascular changes 197, and macular degeneration 295 [20].

The retinal fundus images are being analyzed using binary classification models in three different approaches: based on feature extraction and deep neural networks, convolutional neural networks, and utilizing Microsoft Azure's Custom Vision tools for image classification. The results from these methods are being consolidated for comparison and analysis providing a foundation for potential future aggregation using the approach based on Choquet integral extensions, which is being developed and was published by our team [22].

Feature based classification

The fundamental task at the beginning of the analysis is to standardize the input data and to ensure that the data set is split in a way that preserves the distribution of the predicted feature. Because the acquired images varied substantially in resolution, visual quality, and illumination, they were

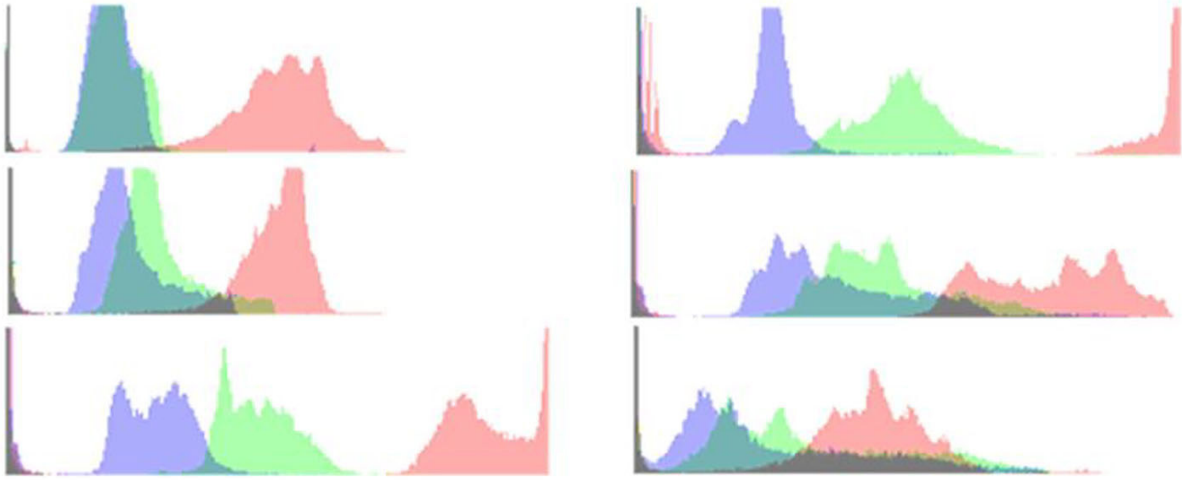

Figure 1. Color histograms for six chosen images from class normal

subjected to a series of preprocessing steps, detailed below. This process included isolating the retinal area from each image, improving contrast by separating and adjusting the LAB color channels by applying CLAHE method to the L (lightness) channel, and subsequently cropping the images. The preprocessing approach similar to proposed in the work [23] is used. It involves four key steps: (1) removing irrelevant margins, such as black borders, by identifying the boundary between these and the retinal region; (2) detecting the circular retina or, if unsuccessful, estimating the circle based on the image center and pixel distribution; (3) cropping the image according to the detected retina circle, and (4) adding black borders during training to prevent the accidental removal of important areas after augmentation, such as random rotations and cropping. Pre-processing is performed on images at their maximum available resolution to ensure that the process is as efficient as possible. Subsequently, all images were resized to a uniform size of 300 × 300 pixels to create a consistent dataset for training machine learning models. It should be noted that due to the poor quality of the input images, selecting the circle containing the retinal image results in black borders of varying thickness across different images, which affects the feature extraction process.

The images were pre-processed using the standard Contrast Limited Adaptive Histogram Equalization (CLAHE) approach to enhance local contrast and improve the visibility of fine details. In this approach, the image is divided into small blocks, referred to as tiles, and histogram equalization is applied separately to each region. This ensures that contrast enhancement is localized to specific areas, rather than being applied uniformly across the entire image. However, if noise is present, it can be amplified during this process. To mitigate this issue, contrast limiting is introduced, where excessively high histogram bins are clipped and their values are redistributed before equalization. Finally, to smooth transitions between adjacent tiles and reduce artifacts at their borders, interpolation is applied. Following this, three linear combinations of the original image and its Gaussian blurred version were applied to emphasize relevant features and suppress noise. The Gaussian blur was used to create smoothed versions of the image, which were then combined in varying proportions with the original to highlight different structural elements. After pre-processing, the circular border of each image, often introduced during image acquisition, was removed to ensure uniformity and to focus solely on the region of interest. Figure 2 presents four example images of the fundus from subset normal. In subsequent rows, the original images are displayed, followed by those processed using the CLAHE method and three mentioned linear combinations.

In the first experiment, an attempt is made to extract features from the obtained images and apply this data to build models for image classification. To enhance the diversity of the dataset and improve model robustness, data augmentation techniques were employed. A horizontal flip was applied randomly to a subset of the images, effectively doubling the available perspectives. Additionally, each image was subjected to a random rotation by an angle within the range of -20 to 20

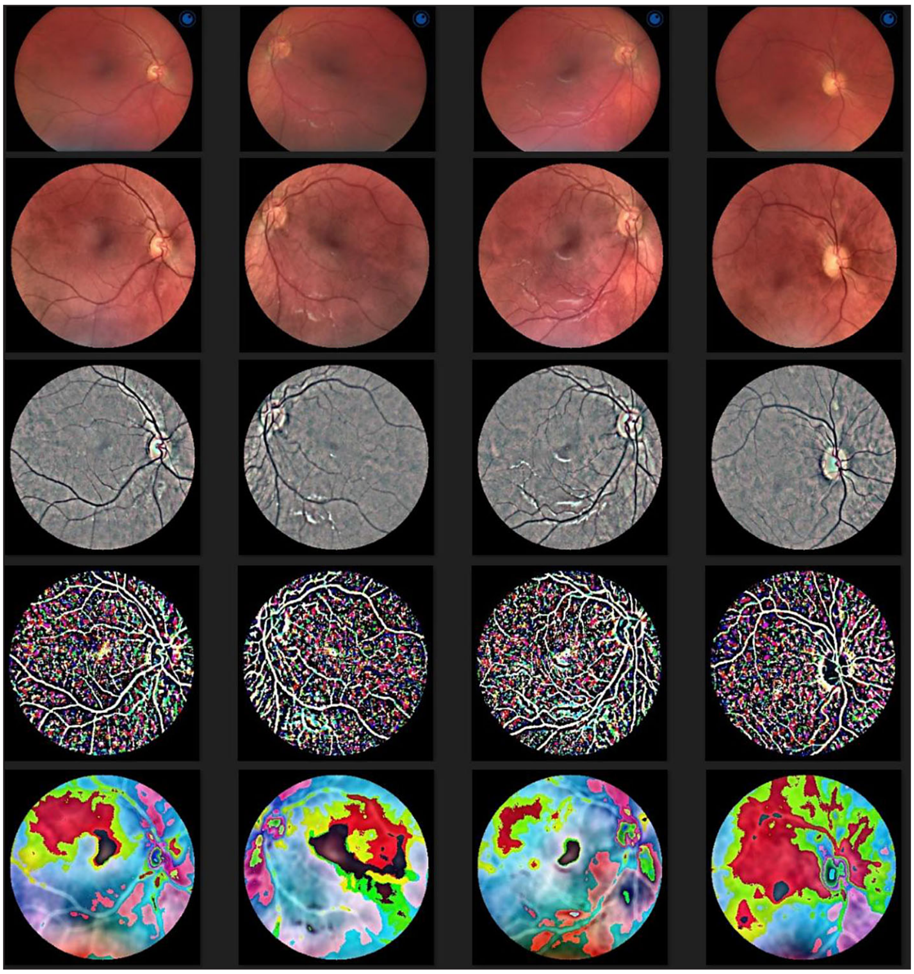

Figure 2. Preprocessed images

degrees, simulating various orientations encountered in real-world scenarios. These augmentations helped mitigate overfitting and ensured better generalization during model training. The qMaZda tool is used for feature extraction [24, 25]. This program is designed for digital image analysis and recognition. It calculates features related to color, texture, and shape in user-defined regions of interest. Color analysis includes models such as RGB, CMY, HSY, YIQ, YUV, CIELab, and CIEXYZ. Texture feature extraction is conducted using methods like co-occurrence matrices, run-length matrices, autoregression models, brightness distribution statistics, local binary patterns, histograms of oriented gradients, Haar, and Gabor transforms. In addition, morphological features such as inertia moments, Feret diameters, and various shape descriptors are computed. A script has been prepared to automate these operations, leveraging the tool's advantage of ease in conducting analyses through script execution. A feature dataset consisting of 567 columns is extracted. The dataset is refined by removing columns containing only zeros, columns with constant values, and columns exhibiting low data diversity, as measured by their standard deviation. After these operations, the dataset contains 420 columns. In the next step, standard scaling was performed using the Z-score method to normalize the features values. This approach ensured that the data had a mean of zero and a standard deviation of one, promoting consistency across the dataset and improving the convergence of the model during training.

Artificial neural networks were employed for binary classification for each class pair in the dataset. For five classes, this resulted in ten pairs for which networks were built and trained. The AutoKeras library was utilized for model creation and training, automating the selection of network architecture. Training was conducted for up to 2500 epochs, utilizing early stopping with patience 100 and restoring the best model weights. The final classification results of this experiment

Table 1. Classification results using deep neural networks

Parameter	Normal	Trauma-related	Vascular	Macular	Optic disc
Normal	X	0.556	0.956	0.953	0.909
Trauma-related	0.556	X	0.841	0.906	0.907
Vascular	0.956	0.841	X	0.646	0.641
Macular	0.953	0.906	0.646	X	0.612
Optic disc	0.909	0.907	0.641	0.612	X

are presented in Table 1, that contains accuracy calculated for testing set.

Artificial neural networks are used for classification for each subset. Multiclass classification was conducted for each preprocessing algorithm to evaluate their effectiveness in distinguishing between different classes. The study initially explored larger neural network architectures, starting with a configuration of hidden layers with the following number of neurons: (500, 250, 250, 100, 50, 5), comprising over 430,000 trainable parameters. Despite the application of DropOut layers, this model exhibited noticeable overfitting. Subsequently, the number of neurons in the hidden layers was gradually reduced, including a configuration of (500, 250, 200, 50, 5) with over 390,000 parameters, which also failed to generalize well. Ultimately, a significantly smaller architecture of (300, 50, 5) neurons and additional DropOut layers, consisting of approximately 130,000 trainable parameters, was found to achieve satisfactory accuracy without signs of overfitting. By applying the preprocessing techniques individually and assessing their impact on classification performance, insights were gained into how each method contributed to feature extraction and model accuracy. The fitting process is conducted for up to 2000 epochs, utilizing early stopping with patience set to 20 and restoring the best model weights.

The final classification results of this experiment are presented in Table 2, which presents the validation accuracy and loss obtained on datasets with different preprocessing methods for multiclass classification. Each analyzed preprocessing method achieved at least 80% accuracy. Notably, the CLAHE method yielded the highest accuracy, while the third linear combination approach resulted in the minimum loss function value.

Table 2. Classification results using deep neural networks

Preprocessing	Accuracy	Loss
CLAHE	0.833	0.227
Combination 1	0.800	0.251
Combination 2	0.810	0.264
Combination 3	0.816	0.224

Image-based classification

In the second model, images were likewise used as the input data. This time, however, we use the visual information contained in them directly and after multistep preprocessing. Nowadays, in most cases, when building classifiers for an image, deep networks based on convolution are used. Their architectures can be different, which largely depend on the characteristics of images, their diversity, and expected effectiveness. In our case, imaged data are transferred both at the training stage and for later inference to the multilayer convolutional network. This approach is relatively common in the case of retinal change classification, and the proposed processing pipelines differ in way how input images are treated, selected model architecture, and used hyperparameters values [21, 23]. The images are processed based on a modified method proposed by B. Graham for the Diabetic Retinopathy Detection competition [26, 27]. After finding the region representing the captured retinal image, scaling and cropping is applied and Gaussian blurred version of image is created (Figure 3).

The image was then normalized to the maximum range of pixel values using the levels function. The two variants are then added together. Finally processed 224 × 224 pixel color images are used as input for the model (Figure 4). The model selected for this task was Imagenet [28]. SSD-based object detection model trained on Open Images V4 with ImageNet pre-trained MobileNet V2 as image feature extractor, which is lightweight convolutional neural network providing high quality features for downstream classifications [29]. This version of the model improves upon the original MobileNet by introducing inverted residual blocks and a linear bottleneck, which reduce computational complexity while maintaining accuracy.

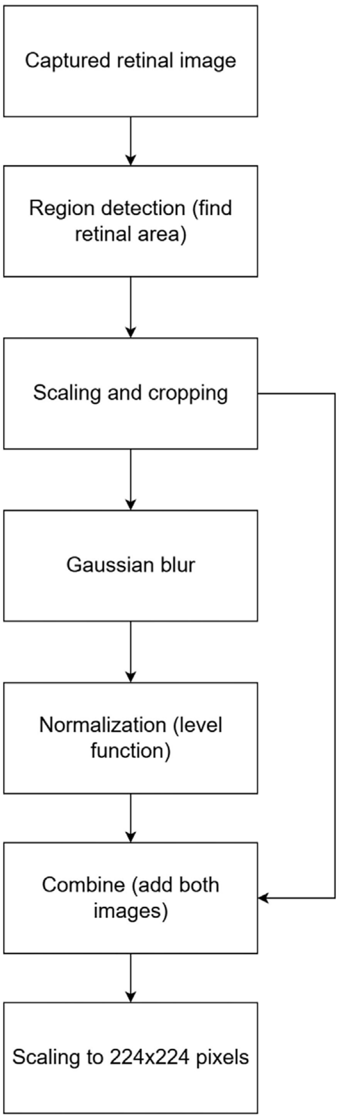

Figure 3. Preprocessing flow

Total over 1.3 M, from total number of 3.6 M, of parameters was used during training.

ImageNet pre-trained MobileNet V2 use input as 224 × 224 × 3 RGB images. Initial layers are convolutive 3 × 3 with striding 2 and 32 filters, then there are inverted residual blocks series of blocks with varying expansion factors *t*, output channels *c*, number of repetitions *n,* and stride *s* (Table 3). Final convolution layer is 1 × 1 and global average pooling layer which reduces feature map to 1280 channels vector. Fully connected dense layers with 1000 units with softmax functions forms outputs. During training and validation, the data were split into sets at a ratio of 1:5. Pretraining on a large, diverse dataset gives the model a rich set of general features. When fine-tuning on smaller datasets, this acts as a strong regularizer, reducing overfitting. Learning process was split in two main steps. The first model was trained using the Soft F1-score loss function, which is differentiable and derived from the F1-score, taking into account both precision and recall:

Table 3. Inverted residual block configuration

t (expansion)	c (output channels)	n (repeats)	s (stride)
1	16	1	1
6	24	2	2
6	32	3	2
6	64	4	2
6	96	3	1
6	160	3	2
6	320	1	1

$$Soft\ F1\ Loss = 1 - 2 \cdot TP / (2 \cdot TP + FP + FN) \quad (1)$$

where: *TP* is true positive value, *FP* is false positive and *FN* is false negative.

In case of multilabel classification selection of this loss function is quite typical for step which is focused on preparation of set of features.

In the second step, pairs of data from different labels were used to train the same model with the binary cross-entropy loss function. This resulted in a multilabel classifier. The outcomes of the processed images are presented in Figure 5. Training was conducted in two rounds, each consisting of 200 epochs with a batch size of 256.

After first, feature extraction, round we got finally F1 score for training set equal to 0.69 and 0.53 for validation set. In case of second, binary classification round, F1 score for training set was equal to 0.78 and 0.53 for validation set. Results, are quite average, but probably there is still space for improvement which could be made by model tuning, modification or extension to data set to get more balanced input. Figure 6 shows training and validation for a round of binary loss functions.

Custom Vision classification

We also tested as third possible solution Microsoft Azure AI Custom Vision service. We employed a model and workflow designed for multi-tag classification. This Microsoft service employs deep learning-based architectures, primarily CNNs. The specific model architectures may vary depending on the chosen configuration, but Microsoft does not disclose the exact models used. From the available information, Custom

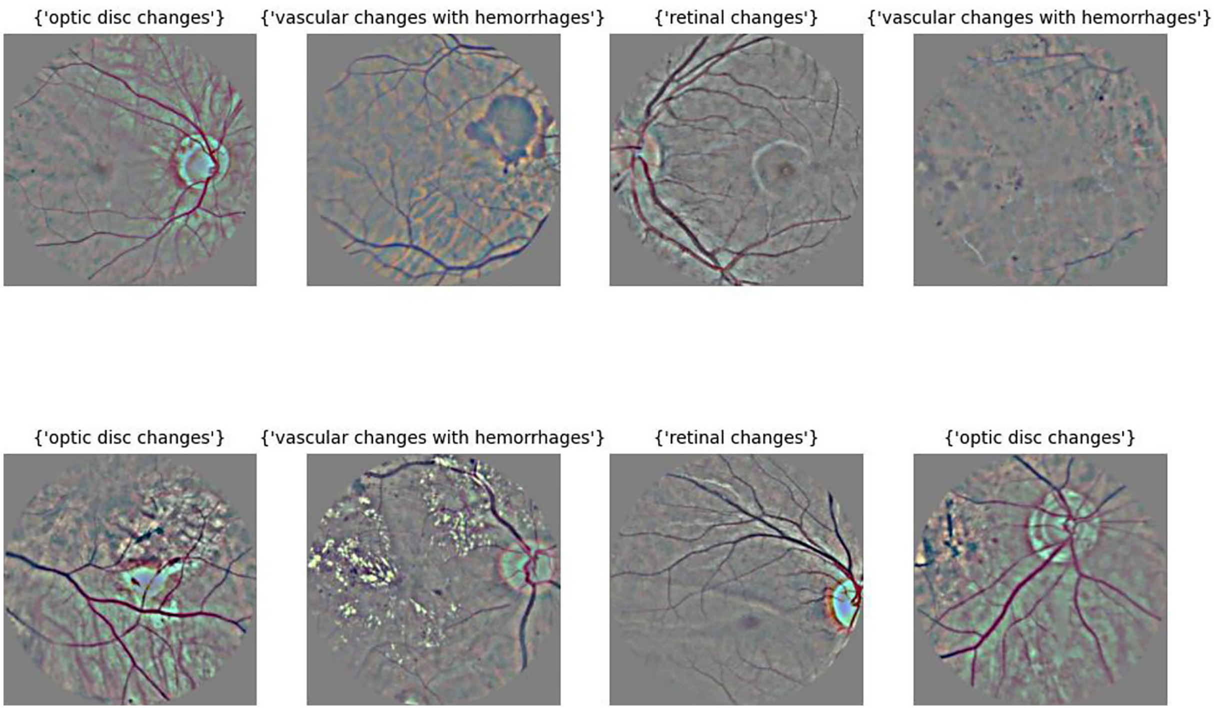

Figure 4. Eight processed images from training set with assigned labels

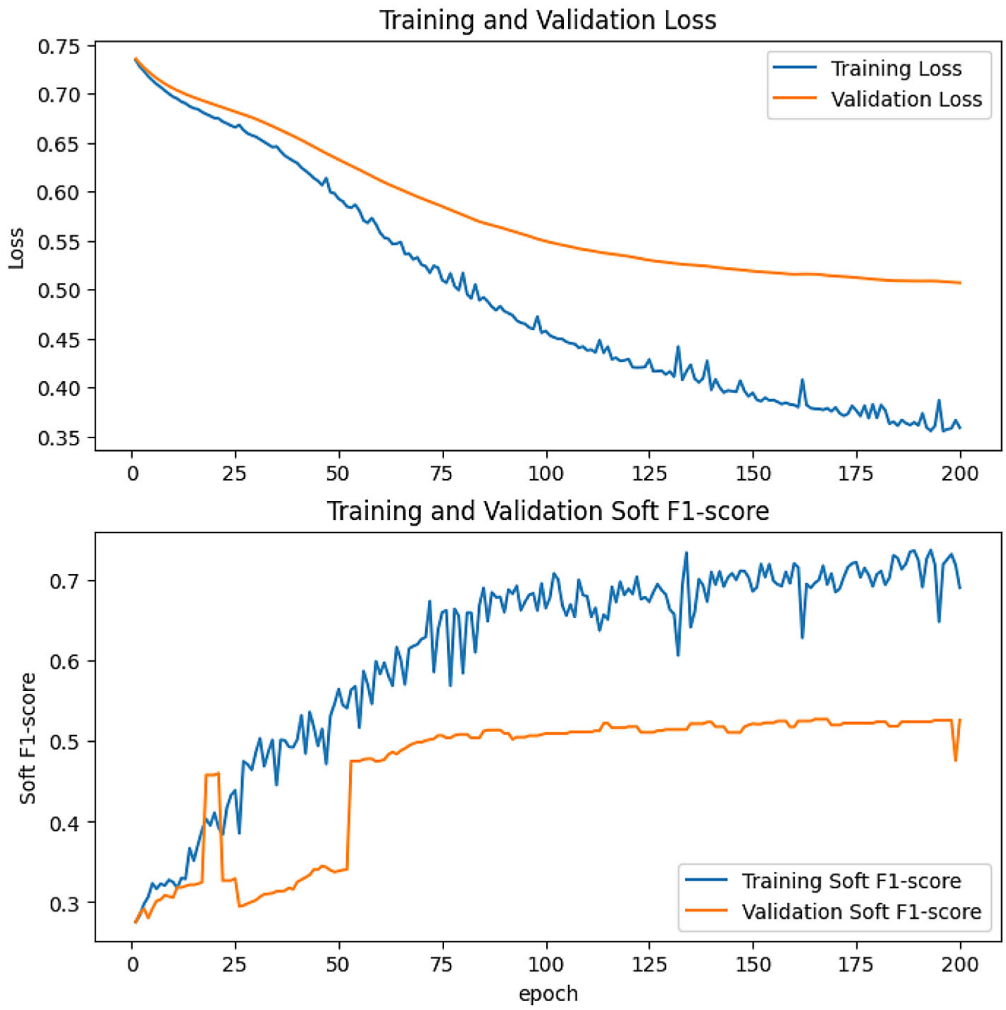

Figure 5. Training and validation loss and F1-score for first training round

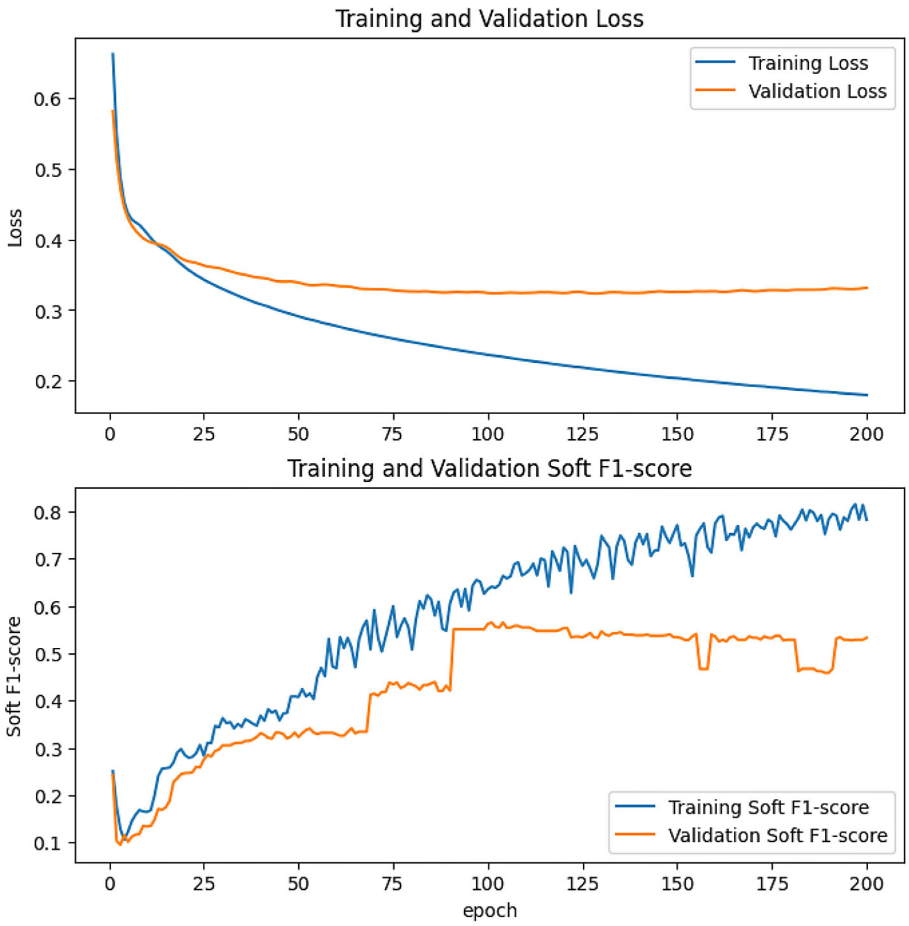

Figure 6. Training and validation for binary loss function round

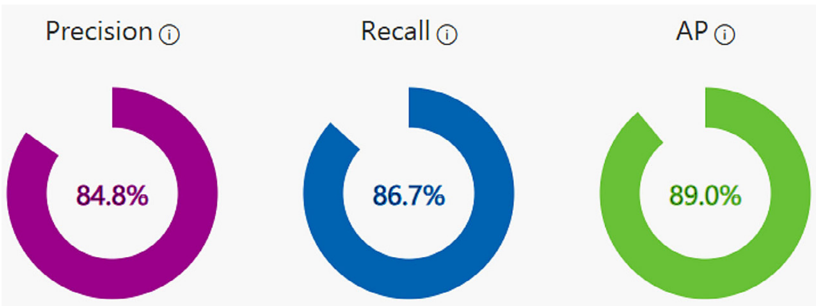

Figure 7. Graphical representation of results generated bu Custom Vision

Vision utilizes optimized CNN models similar to ResNet, MobileNet, and EfficientNet and allows training models on custom datasets by employing transfer learning, enabling the adaptation of pre-trained networks to new applications. Custom Vision offers different levels of accuracy and performance (fast vs. more precise models) corelated with payment for training process. It is possible to export the trained model into various formats, such as ONNX, TensorFlow, or Core ML. In our case we started with images preprocessed the same way as in case of image based classification. This method could be treated as a kind of ad hoc solution because of limited configuration possibilities. Custom Vision reports classifier performance in terms of average precision; therefore, accuracy was calculated independently based on the content of the test subsets.

The results from the Custom Vision model indicate above-average performance, with metrics such as recall and accuracy. The results shown in Figure 7 represent of the Microsoft Azure AI

Custom Vision service test – average precision 0.89, recall 0.87 and accuracy 0.85.We could compare this results with F1 score because the later is a balance between precision and recall. Of course for imbalanced datasets or cases where both precision and recall are critical, the F1 score is a better metric than accuracy alone. It means that despite the fact that Custom Vision is a closed, proprietary solution provided by Microsoft for building and deploying image classification and object detection models it could be treated as a practical starting point for many applications or basic performance level. The results in Table 4 show scores for Custom Vision multilabel classification.

As you can see on Figure 8 the red outline points out misclassifications in validation set. Image should have been classified as a pigment cluster class with probability equal to 98.8% and as a vascular changes with probability 40.4%. The second class was correct; however, in this case, the model also recalled pigment cluster images, which is not unexpected.

Table 4. Scores for Custom Vision multilabel classification

Tag	Presion	Recall	Average precision
Pigment clusters	1.00	0.67	0.68
Normal	1.00	1.00	1.00
Vascular changes with hemorrhages	0.91	0.86	0.98
Retinal changes	0.82	0.92	0.88
Optic disc changes	0.80	0.76	0.84

Overall, the feature-based approach combined with CLAHE preprocessing achieved the highest multiclass accuracy (0.833), indicating strong performance despite the limited and imbalanced dataset. The image-based CNN approach (MobileNet V2) achieved lower validation F1-scores (0.53), though it demonstrated the advantage of an end-to-end pipeline without manual feature engineering. The Custom Vision service performed comparably (average precision 0.89, recall 0.87, accuracy 0.85), confirming the viability of transfer learning with pre-trained CNN architectures. It should be noted that these results are not strictly equivalent, since each method used a slightly different setup, handcrafted features with ANN classification, MobileNet-based deep learning, and Microsoft's proprietary transfer learning framework. Therefore, the outcomes should be interpreted as complementary evidence rather than directly comparable benchmarks.

DISCUSSION

The aim of this study was to investigate the potential clinical applications of AI-based algorithms for the detection of retinal injuries during armed conflict. We found that CNNs can accurately identify different types of retinal injuries. Importantly, the results of our analysis are comparable to those obtained using more advanced technologies, such as optical coherence tomography (OCT). Although we used less sophisticated

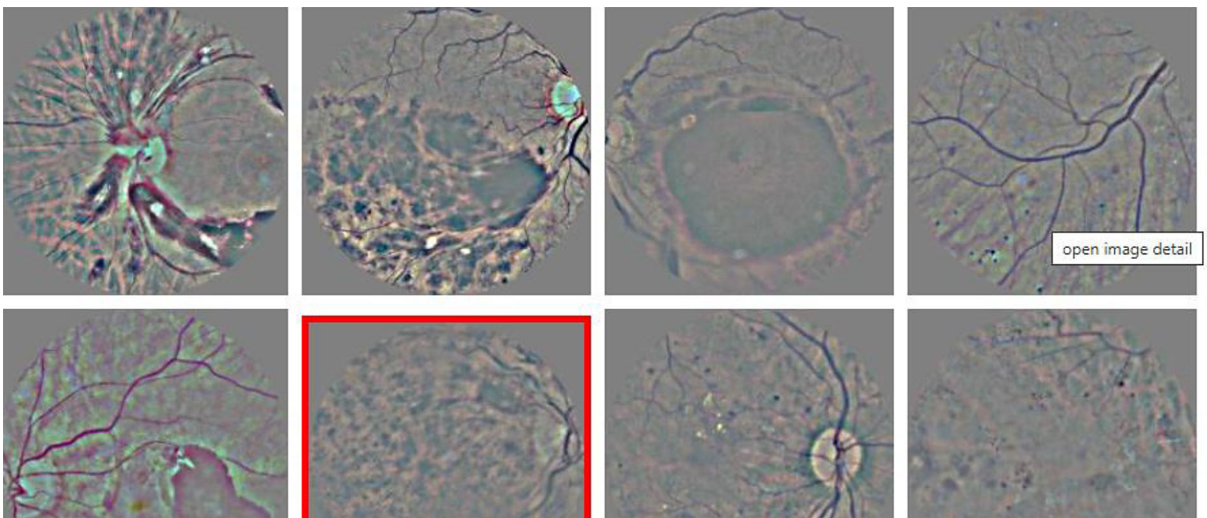

Figure 8. Part of performance test for model trained in Custom Vision

equipment, AI has shown that it can provide very precise diagnoses. This information is particularly valuable in situations where access to modern diagnostic tools is limited. Previous works discusses current applications of artificial intelligence in ophthalmic diagnostics. Some articles like [30], highlights how quickly AI can analyze images of the eye. The article also provides a broader picture of deep learning, in the context of eye diseases as well as highlights challenges in ethics and integration with traditional treatments. Additionally, previous articles [30–32] mainly focus on the use of deep learning and AI in the diagnosis of eye diseases. They point to the use of neural networks in analyzing fundus images to accurately detect pathological changes. These studies underscore the effectiveness of artificial intelligence in speeding up the diagnostic process and improving the precision of diagnoses.

Research [33–35] focuses on artificial intelligence in the diagnosis of retinal diseases. They mainly study the macula and on using neural networks to analyze fundus images. This research highlights the challenges of accessing technology in a constrained environment. Comparing our research, this too is conducted under limited conditions. These solutions aim to create diagnostic tools that are effective in difficult field conditions.

Our results confirm previous findings that AI has great potential in medical image analysis. For example, the Hemalakshmi et al. study [36] showed how advanced techniques such as sparse vision transformer (SViT) can be used in medical diagnosis. Although we focused on fundus camera images, our results clearly show that CNNs speed up diagnosis and improve accuracy. This is crucial when making quick decisions matters. Similar observations appeared in a study [10] that dealt with the classification of OCT images in diabetic retinopathy. The results from this article, which dealt with the classification of retinal trauma caused by war trauma, overlap with the article [10]. This shows that AI-based techniques can be used not only for the diagnosis of classic retinal diseases, but also for the analysis of injuries in conflict zones. Our observations are also in line with studies that combine CNN techniques with feature extraction methods, significantly improving the precision of diagnosis [11]. This is especially important in situations where access to advanced diagnostic tools such as OCT is severely limited. Thus, we see that CNNs can provide valuable support in the analysis of eye injuries resulting from warfare, offering an effective diagnostic solution even with limited resources.

In addition to these observations, several practical and ethical considerations must be emphasized. One is data variability. Retinal fundus images collected in conflict zones often differ in resolution, illumination, and overall quality due to diverse acquisition devices and unstable environmental conditions. Moreover, demographic variability across patient populations (e.g., age, ethnicity, comorbidities) can affect retinal presentation and may influence the robustness of AI models. While our current dataset provides valuable proof-of-concept evidence, future studies should incorporate larger and more diverse datasets from multiple centers to enhance generalizability. Another important issue is patient safety. AI-assisted diagnostics in ophthalmology must be viewed strictly as decision-support tools rather than autonomous diagnostic systems. Clinicians should remain responsible for final medical decisions, using AI outputs as supportive evidence. Safeguards such as clear uncertainty quantification, fail-safe mechanisms when input quality is inadequate, and clinician training on model interpretability are essential to mitigate risks of misdiagnosis. Finally, limitations in generalization should be acknowledged. Models trained on data from specific populations or conflict-affected regions may not fully capture variations present in other geographic or socioeconomic contexts. Ensuring equitable performance across different patient groups requires further validation, ideally through multicenter collaborations and prospective clinical trials. Addressing these aspects will be crucial for safe and reliable integration of AI-assisted pipelines into real-world healthcare workflows, particularly in fragile healthcare systems.

In the context of medical image segmentation, UTNet's hybrid transform architecture, discussed in [12], shows promise in separating areas of interest in images, even under difficult conditions. Although we did not address segmentation in our study, we believe that in the future it could further improve the precision of diagnoses by allowing more accurate assessment of damaged areas of the retina. Similarly, a study [13] suggesting combining CNNs with transformers for medical image segmentation highlights that hybrid approaches can improve diagnostic performance. Although our work has focused primarily on classification, we see potential in extending the study to include segmentation to obtain even more detailed information about retinal injuries from warfare.

Regarding the application of AI in the diagnosis of eye injuries in conflict zones, our study clearly shows that CNNs can effectively identify and classify retinal damage, even under extreme conditions. Moreover, the results suggest that AI's potential may extend beyond the retina itself. In the future, we could extend diagnostics to other structures of the eye, such as the cornea. This expansion of the diagnostic system could significantly improve the efficiency of emergency care.

The main limitation of our study was the quality of the fundus camera images, which was sometimes compromised by unfavorable conditions and the limitations of the equipment itself. Still, the use of CNNs for automated image analysis made it possible to standardize the diagnostic process and reduce the impact of variability in image quality. Although some images were not suitable for analysis, the vast majority of data were good enough to confirm the validity of our approach.

This study attempts to create a diagnostic model to facilitate the work of doctors. Data for the model was taken under wartime conditions. It is worth noting that similar research under wartime conditions has not been conducted before, which makes our project unique and adapted to the realities of battlefield operations. It should be noted that earlier works did not address the realities of war, which makes our approach particularly valuable [10–14]. We would like to emphasize that the presented research represents the first stage of clinical work to evaluate the usefulness of the diagnostic system in a field setting. In addition, the expanded analysis of error rates provides a better understanding of the potential limitations of the model in terms of practical application.

To ensure the highest ethical standards, all participants in the study were fully informed about the objectives and methodology, and the research produres were conducted in accordance with applicable ethical standards. Each patient gave informed consent for the tests conducted.

Limitations

A limitation of this study is the relatively small dataset of 448 images, which may restrict the generalizability of the findings. Larger datasets are typically required to fully capture the variability of retinal injuries and to train deep learning models with stronger robustness across diverse patient populations. Furthermore, the dataset is imbalanced, with some diagnostic categories being underrepresented (e.g., only 24 trauma-related cases compared to 295 macular degeneration cases). This imbalance may have biased the classification outcomes toward the majority classes and reduced the reliability of results for rarer but clinically critical categories such as trauma-related injuries. Nevertheless, the dataset presented here is unique, as it was collected under actual wartime conditions, where access to medical facilities and diagnostic tools was severely constrained. This makes it not only a valuable proof-of-concept resource but also one of the first such datasets available from conflict zones. While the limited size and imbalance may have introduced biases and constrained the statistical power of the models, the consistently high accuracy across different approaches suggests that AI-based fundus analysis can still provide meaningful diagnostic support. Future research should expand the dataset, incorporate multicenter collaborations, and apply advanced techniques such as transfer learning, data augmentation, and class rebalancing strategies to improve the generalizability and clinical applicability of the models. The quality of images was sometimes compromised due to unfavorable acquisition conditions and the limitations of portable fundus cameras used in the field. Finally, the wartime setting constrained the range of available diagnostic data and limited opportunities for repeated examinations.

CONCLUSIONS

Our study confirms that the use of fundus cameras combined with Artificial Intelligence, particularly artificial neural networks, serves as an effective clinical diagnostic alternative in wartime conditions where access to advanced equipment, such as OCT, is limited. Convolutional and Deep Networks prove efficient in classifying retinal injuries, such as vessel damage, detachments, and foreign bodies, significantly speeding up diagnoses. Despite variability in image quality under challenging conditions, AI-based automation enables the standardization of diagnostic procedures. The study demonstrates that even less advanced technologies yield satisfactory results. Image segmentation and hybrid AI architectures further enhance precision. These findings highlight the potential of AI applications in conflict and crisis situations with limited resources, guiding future research in both military and civilian contexts.

It is important to stress the differences between diagnostics in wartime and peacetime contexts. In peacetime, ophthalmologists usually have access to advanced technologies, comprehensive patient records, and adequate time for thorough examinations. In contrast, wartime diagnostics are constrained by scarce resources, reliance on portable devices, and the urgent need for rapid decision-making, often under stressful and hazardous conditions. Our study demonstrates that, despite these challenges, AI-assisted fundus analysis can provide meaningful diagnostic support and help bridge the gap created by the absence of advanced equipment. These findings highlight the potential of AI not only as a complementary tool in conventional healthcare systems but also as a critical resource in crisis and conflict settings where conventional diagnostic pathways are disrupted. The future clinical work directions will, among other, the application of multiple classifiers in the process of aggregation of their results to improve the accuracy measure. Moreover, we are going to investigate other algorithms dedicated to medical image classification. Finally, a larger collection of eye injuries coming from the war conflicts areas would be helpful to improve the quality of methods.

Acknowledgements

Project financed by the Polish National Agency for Academic Exchange under the Urgency Grant program, entitled "Cross-border development telemedicine system on the example of ophthalmological examinations in conditions of armed conflict – analyzes and recommendations", project number: BPN_GIN_2022_1_00120.

REFERENCES

1. Jonak, K. et al. War-related eye trauma: A study of civilian and military cases from Ukraine's Ongoing conflict. Front. Public Health 2025; 13.
2. Wong, T. Y., Seet, M. B., Ang, C.-L. Eye injuries in twentieth century warfare: A historical perspective. Survey of Ophthalmology 1997; 41, 433–459.
3. Akhlaghi, F., Aframian-Farnad, F. Management of maxillofacial injuries in the Iran-Iraq war. Journal of Oral and Maxillofacial Surgery 1997; 55, 927–930.
4. Biehl, J. W., Valdez, J., Hemady, R. K., Steidl, S. M., Bourke, D. L. Penetrating eye injury in war. Mil Med 1999; 164, 780–784.
5. Li, F. et al. Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs. Eye 2022; 36, 1433–1441.
6. Akter, N. et al. Glaucoma diagnosis using multi-feature analysis and a deep learning technique. Sci Rep 2022; 12, 8064.
7. Ting, D. S. W. et al. Artificial intelligence and deep learning in ophthalmology. Br J Ophthalmol 2019; 103, 167–175.
8. Promising artificial intelligence-machine learning-deep learning algorithms in ophthalmology. Asia Pac J Ophthalmol (Phila) 2019. https://doi.org/10.22608/APO.2018479
9. Hemalakshmi, G. R., Murugappan, M., Sikkandar, M. Y., Begum, S. S., Prakash, N. B. Automated retinal disease classification using hybrid transformer model (SViT) using optical coherence tomography images. Neural Comput & Applic 2024; 36, 9171–9188.
10. Chavan, R., Pete, D. Automatic multi-disease classification on retinal images using multilevel glowworm swarm convolutional neural network. J. Eng. Appl. Sci. 2024; 71, 26.
11. Islam, M. M., Yang, H.-C., Poly, T. N., Jian, W.-S., (Jack) Li, Y.-C. Deep learning algorithms for detection of diabetic retinopathy in retinal fundus photographs: A systematic review and meta-analysis. Computer Methods and Programs in Biomedicine 2020; 191, 105320.
12. Gao, Y., Zhou, M., Metaxas, D. N. UTNet: A Hybrid Transformer Architecture for Medical Image Segmentation. in Medical Image Computing and Computer Assisted Intervention – MICCAI 2021 (eds. De Bruijne, M. et al.) 12903, 61–71. (Springer International Publishing, Cham, 2021).
13. Kuwayama, S. et al. Automated detection of macular diseases by optical coherence tomography and artificial intelligence machine learning of optical coherence tomography images. Journal of Ophthalmology 2019; 1–7.
14. Dutta, P., Sathi, K. A., Hossain, Md. A., Dewan, M. A. A. Conv-ViT: A convolution and vision transformer-based hybrid feature extraction method for retinal disease detection. J. Imaging 2023; 9, 140.
15. Jarrassier, A. et al. Lessons learned from the war in Ukraine for the anesthesiologist and intensivist: A scoping review. Anaesthesia Critical Care & Pain Medicine 2024; 43, 101409.
16. Kricorian, K., Khoshnood, K., Chekijian, S. Hybrid warfare and public health: Conflicts in Ukraine and Nagorno-Karabakh raise the alarm. Public Health in Practice 2022; 4, 100342.
17. Optomed Aurora IQ. Optomed https://www.optomed.com/optomed-aurora-iq/

18. Aurangzeb, K., Alharthi, R. S., Haider, S. I., Alhussein, M. An efficient and light weight deep learning model for accurate retinal vessels segmentation. IEEE Access 2023; 11, 23107–23118.
19. Ashok, N., Gangadhara Rao, K. Detection of Retinal Diseases Using Advanced Deep Learning Algorithms. In: 2023 International Conference on Sustainable Communication Networks and Application (ICSCNA) 1503–1509 (IEEE, Theni, India, 2023). https://doi.org/10.1109/ICSCNA58489.2023.10370475
20. Verma, G. Efficient Retinal Disease Detection with ResNet-18: A Deep Learning Approach. In: 2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI) 812–816 (IEEE, Coimbatore, India, 2024). https://doi.org/10.1109/ICoICI62503.2024.10696615
21. Ejaz, S. et al. A deep learning framework for the early detection of multi-retinal diseases. PLoS ONE 2024; 19, e0307317.
22. Parmar, U. P. S. et al. Artificial intelligence (AI) for early diagnosis of retinal diseases. Medicina 2024; 60, 527.
23. Cen, L.-P. et al. Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. Nat Commun 2021; 12, 4828.
24. Szczypiński, P. M., Strzelecki, M., Materka, A., Klepaczko, A. MaZda – A software package for image texture analysis. Computer Methods and Programs in Biomedicine 2009; 94, 66–76.
25. Szczypinski, P. M., Klepaczko, A., Kociolek, M. QMaZda – Software tools for image analysis and pattern recognition. in 2017 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA) 217–221 (IEEE, Poznan, 2017). https://doi.org/10.23919/SPA.2017.8166867
26. Diabetic Retinopathy Detection. https://kaggle.com/diabetic-retinopathy-detection
27. Navaneethan, R., Devarajan, H. Enhancing diabetic retinopathy detection through preprocessing and feature extraction with MGA-CSG algorithm. Expert Systems with Applications 2024; 249, 123418.
28. Deng, J. et al. ImageNet: A large-scale hierarchical image database. in 2009 IEEE Conference on Computer Vision and Pattern Recognition 248–255 (IEEE, Miami, FL, 2009). https://doi.org/10.1109/CVPR.2009.5206848
29. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. 2018. https://doi.org/10.48550/ARXIV.1801.04381
30. Lim, J. I., Rachitskaya, A. V., Hallak, J. A., Gholami, S., Alam, M. N. Artificial intelligence for retinal diseases. Asia-Pacific Journal of Ophthalmology 2024; 13, 100096.
31. Bali, A., Mansotra, V. Analysis of deep learning techniques for prediction of eye diseases: A systematic review. Arch Computat Methods Eng 2024; 31, 487–520.
32. Senapati, A., Tripathy, H. K., Sharma, V., Gandomi, A. H. Artificial intelligence for diabetic retinopathy detection: A systematic review. Informatics in Medicine Unlocked 2024; 45, 101445.
33. Anton, N. et al. Comprehensive review on the use of artificial intelligence in ophthalmology and future research directions. Diagnostics 2022; 13, 100.
34. Ting, D. S. W. et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA 2017; 318, 2211.
35. Myślicka, M., Włodarczyk, K., Kawala-Sterniuk, A., Mikołajewski, D., Mikołajewska, E. The role of artificial intelligence in ophthalmology – brief review. Ophthalmol J 2024; 9, 106–113.
36. Hemalakshmi, G. R., Murugappan, M., Sikkandar, M. Y., Begum, S. S., Prakash, N. B. Automated retinal disease classification using hybrid transformer model (SViT) using optical coherence tomography images. Neural Comput & Applic 2024; 36, 9171–9188.