BURNS 50 (2024) 966–979

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/burns

Segmentation and classification of skin burn images with artificial intelligence: Development of a mobile application

Metin Yıldız [a,*,1], Yakup Sarpdağı [b,2], Mehmet Okuyar [c,3], Mehmet Yildiz [d,4], Necmettin Çiftci [e,5], Ayşe Elkoca [f,6], Mehmet Salih Yildirim [g,7], Muhammet Ali Aydin [h,8], Mehmet Parlak [i], Bünyamin Bingöl [j,9]

[a] Department of Nursing, Sakarya University, Sakarya, Turkey
[b] Department of Nursing Van Yuzuncu Yil University, Turkey
[c] Sakarya University of Applied Sciences Biomedical Engineering, Sakarya, Turkey
[d] Sakarya University of Applied Sciences, Distance Education Research and Application Center, Sakarya, Turkey
[e] Muş Alparslan University, Faculty of Health Sciences, Department of Nursing, 49100 Muş, Turkey
[f] Gaziantep Islamic University of Science and Technology Faculty of Health Sciences, Midwifery, Turkey
[g] Vocational School of Health Services, Agri Ibrahim Cecen University School of Health, Agri, Turkey
[h] Department of Nursing, Ataturk University, Erzurum, Turkey
[i] Ataturk University, Department of Nursing, Erzurum, Turkey
[j] Sakarya University, Electrical and Electronics Engineering, Sakarya, Turkey

ARTICLE INFO

Article history:
Accepted 10 January 2024

Keywords:
Burn
Segmentation
Classification

ABSTRACT

Aim: This study was conducted to determine the segmentation, classification, object detection, and accuracy of skin burn images using artificial intelligence and a mobile application. With this study, individuals were able to determine the degree of burns and see how to intervene through the mobile application.

Methods: This research was conducted between 26.10.2021–01.09.2023. In this study, the dataset was handled in two stages. In the first stage, the open-access dataset was taken from https://universe.roboflow.com/, and the burn images dataset was created. In the

* Corresponding author.
E-mail address: metinyildiz@sakarya.edu.tr (M. Yıldız).
[1] ORCID ID: 0000-0003-0122-5677.
[2] ORCID ID: 0000-0002-1608-649X.
[3] ORCID: 0000-0002-3879-557X.
[4] ORCID: 0000-0002-9523-3805.
[5] ORCID: 0000-0002-4713-4212.
[6] ORCID ID: 0000-0002-3936-4122.
[7] ORCID ID: 0000-0003-2632-4583.
[8] ORCID ID: 0000-0002-2137-5114.
[9] Orcid ID: 0000-0002-9040-3107.

https://doi.org/10.1016/j.burns.2024.01.007
0305-4179/© 2024 Elsevier Ltd and ISBI. All rights reserved.

Mobile application
Object detection

second stage, in order to determine the accuracy of the developed system and artificial intelligence model, the patients admitted to the hospital were identified with our own design Burn Wound Detection Android application.

Results: In our study, YOLO V7 architecture was used for segmentation, classification, and object detection. There are 21018 data in this study, and 80% of them are used as training data, and 20% of them are used as test data. The YOLO V7 model achieved a success rate of 75.12% on the test data. The Burn Wound Detection Android mobile application that we developed in the study was used to accurately detect images of individuals.

Conclusion: In this study, skin burn images were segmented, classified, object detected, and a mobile application was developed using artificial intelligence. First aid is crucial in burn cases, and it is an important development for public health that people living in the periphery can quickly determine the degree of burn through the mobile application and provide first aid according to the instructions of the mobile application.

© 2024 Elsevier Ltd and ISBI. All rights reserved.

1. Introduction

Burn injuries can lead to high morbidity and mortality and are an important public health problem [1,2]. Globally, burns are the fourth most common type of injury after road traffic accidents, falls, and physical violence [3]. It is known that an estimated 11 million people worldwide are affected by burns each year, with 180,000,000 deaths due to burns, most of which occur in low- and middle-income countries [4]. Burns are preventable but potentially life-altering injuries that can have a significant impact on a person's health and quality of life and can lead to death [5,6].

Because first aid to injured patients plays a critical role in burn care, first aid can help reduce burn complications [7,8]. The classification of wounds is very important in terms of management, diagnosis, selection of the right treatment, the time required for wound healing, and prediction of risks and infections that may occur during the wound healing process [9,10]. A burn injury is usually classified by the healthcare professional into one of three categories: superficial (first-degree burns), superficial-partial or deep-partial (second-degree burns), and full-thickness (third-degree burns), with each category having different healing times and characteristics [11].

The most commonly used method to assess burn depth is clinical diagnosis by visual observation and physical examination [12]. Laser Doppler imaging [13], harmonic ultrasound imaging [14], optical coherence tomography [15], and high-resolution infrared thermography [16] have been developed and incorporated into the limited clinical diagnosis of burns to assess precise burn depth. The use of artificial intelligence as a diagnostic test has evolved [17].

The Deep Learning (DL) approach is a subcategory of Machine Learning (ML), where threshold logic was introduced in 1943 [18] to build a computer model very similar to the biological pathways of humans. This research area is still evolving; its evolution can be divided into two time periods, 1943–2006 and 2012 to date. In the first phase, several developments were observed, such as backpropagation [19,20], chain rule [21], handwritten text recognition (LeNET architecture) [22] and solving the training problem [23]. In the second phase, driverless cars [24,25], healthcare [26,27], text recognition [22], earthquake predictions [28], finance [29] and image recognition [30,31].

Classification of the medical image is the main task in deep learning to investigate clinically relevant issues for early treatment of the patient. Classification can be classical or multiple images as input with a single diagnostic variable (disease yes or no) as the outcome. In these cases, each diagnostic test is a model and the dataset sizes are characteristically small compared to the sizes in computer vision [32].

Advances in image analysis mean that AI algorithms can accurately segment images and identify areas of interest, such as burned tissue, to allow for estimates of burn size/segmentation or focus on these areas to estimate burn depth [33]. The basis of automatic burn diagnosis is the segmentation of the burn image area. Accurate segmentation of the burn area and estimation of the burn depth provide important aids for subsequent clinical treatment [34]. Artificial intelligence (AI) has great potential to advance burn care and can help improve patient experience, promote public health, reduce costs, and enhance provider experience [35]. A systematic review and meta-analysis by Taib et al. is promising for the potential application of AI in the management of burn patients, especially given its positive outcomes across a range of dimensions, including burn depth, extent, mortality, associated sepsis, and acute kidney injuries [17].

A simple way to obtain burn images is through common cameras and smartphones [34]. With the advancing technology, the world of mobile technology is also developing rapidly [36]. In the medical field, emerging programs and applications are gaining importance to aid medical care [36–39]. In burn care, automatic burn size calculations have been proposed through the use of various mobile applications [37,38]. Indeed, the visual nature of burn injuries makes them an excellent candidate for the use of such image-based technologies [40].

The aim of this study is to identify an autonomous classifier that can detect burn severity from real-time photographs of skin burns, segment the burn image area, and categorize them as first-, second-, or third-degree burns using an artificial intelligence-based mobile application.

2. Methods

This research was conducted between 26.10.2021–01.09.2023. In this study, the dataset was handled in two stages. In the first stage, the open-access dataset was taken from https://universe.roboflow.com/, and the skin burn images dataset was created. In the second stage, the accuracy of the developed system and artificial intelligence model was determined with our own design Burn Wound Detection Android application from patients who came to a Public Hospital with burn wounds.

2.1. *Inclusion criteria*

In the second stage of the study (mobile application part), all individuals who agreed to participate in the study, who were over 18 years of age, and who volunteered were included.

2.2. *Data evaluation*

The neural network selected in this study belongs to the supervised learning type. That is, the model for deep learning is subject to supervision. The result is based on a certain prediction. Python programming language was preferred while creating the model architecture. The functional and easy-to-understand nature of this language, as well as its ability to analyze data, are among the reasons for its choice. Python libraries such as Keras, NumPy, Pandas, and OpenCV were selected for the architecture. These libraries are frequently used in deep learning algorithms.

2.3. *Preparation of the dataset*

In creating the dataset, the dataset named skin-burn-degree-classification was used from the website https://universe.roboflow.com/, which shares open-access datasets [41]. The dataset consists of three classes. The degrees of burn are considered as 1,2,3. The dataset is divided into train and test. There are 7006 data in the original dataset. Data augmentation was performed to prevent overfitting (memorization) and to improve model performance by generating variations in the dataset created with image processing techniques in the original dataset. Data augmentation can be applied to any data format, and the number of images in the dataset can be increased by translating, cropping, scaling, rotating, flipping, and changing the brightness and contrast of the original image. In our study, a total of 21018 data were obtained after data augmentation, and 80% of them were used as training data and 20% as test data. In the data preprocessing part, cropping was applied to the image to prevent the model from focusing on irrelevant dark areas (Fig. 1).

Class-label distribution is given in Table 1.

2.4. *Algorithms and software architecture*

Deep learning algorithms can be analyzed under three headings. These can be considered as classification, object detection, and segmentation models. Object detection models were utilized throughout this study. The purpose of object detection is to predict the class information of the object to be detected and its location by drawing a frame around it when an image is presented to the model [42]. In this study, the YOLOv7 architecture was preferred. YOLOv7 is an effective option for object detection tasks due to its high popularity, reliability, and time management. This choice was made because it provides better results than other architectures through speed performance, especially in real-time applications [43].

The main reasons for choosing YOLOv7 are:

Speed Performance: YOLOv7 stands out for its ability to perform object detection tasks quickly. This is critical for real-time applications.

Reliability: YOLOv7 has the ability to successfully recognize various objects and classes, making it suitable for use in a variety of application scenarios.

Development community: YOLOv7 is supported by a large development community and is constantly receiving improvements and updates.

For these reasons, the study tackled object detection tasks using the YOLOv7 architecture and aimed to achieve successful results. This choice also reflects the goal of keeping up with the latest developments and performance in object detection (Fig. 2).

2.5. *Mobile application interface*

Flutter is a free and open source application interface toolkit developed by Google and released in 2017. Flutter consists of two important parts.

2.5.1. *Software development kit*
A collection of tools to help develop applications. Also known as SDK. The code written can be adapted for both IOS and Android.

2.5.2. *Framework*
Enables the user interface components needed in the software development phase to be run. Thanks to these components, mobile application projects can be converted into personal versions. Flutter can be used in different IDEs. One of them is Android Studio. After defining Android Studio to Flutter, the necessary interface was created with the DART programming language. On the opening screen, there is a short text explaining the purpose of the mobile application, the application logo, and the "START" button to run the application. After clicking this button, the user is redirected to the next page and is expected to upload the skin burn image to the application. The upload process is performed either by selecting a photo from the gallery or by taking a new photo with the phone camera. After the relevant photo is approved, it is sent to the server. At the last stage, the burn degree estimation result is shown to the user. According to the degree of the burn, the first aid to be provided is explained (Fig. 3).

2.6. *Ethical principles of the study*

Approval for the study was obtained from the Scientific Research and Publication Ethics Committee of a university (Date and Number: 26.10.2021–293). Written permission was

Fig. 1 – Sample images of the dataset.

Table 1 – Class - label distribution.

Classes	Label Count
1.degree	3869
2.degree	4507
3.degree	4408

acquired from the institution where the research would be conducted (16/12/2022-E-35465298–605.01). The individuals who would participate in the study were informed face-to-face by the researcher about the aim of the study, the method, the time they would spend on the study, that participation in the research would not cause any harm, that it was based on the principle of complete voluntariness, and their written consent was obtained. Because the rights of the

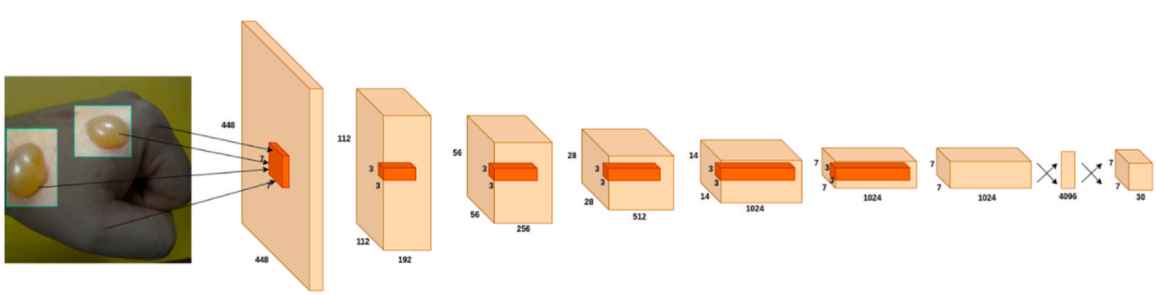

Fig. 2 – The network structure of the YOLO algorithm.

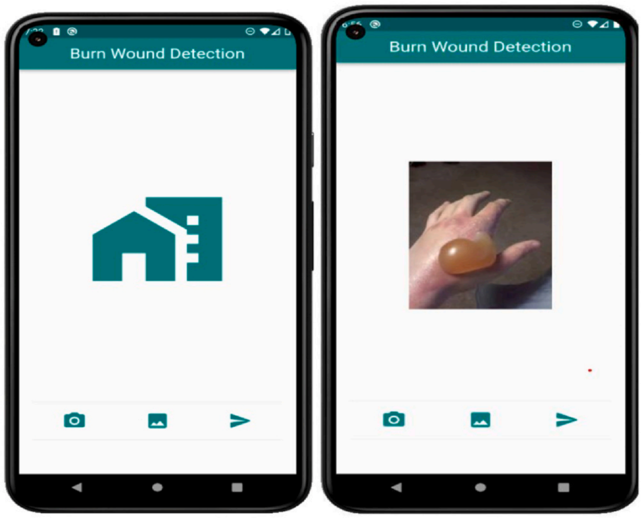

Fig. 3 – Mobile Application Login Section.

individuals should be protected in the research, the Helsinki Declaration of Human Rights was adhered to throughout the study.

3. Results

3.1. Performance metrics

In our study, confusion matrix, recall, precision and f-1 score were used as evaluation criteria.

The confusion matrix is used to evaluate the performance of a classification model after comparing it to the predictions of a classification model on a set of test data where the true values are known.

TP (True Positive): When the actual value is 1 and the predicted value is 1.

TN (True Negative): When the actual value is 0, but the predicted value is also 0.

FP (False Positive): When the actual value is 0, but the predicted value is 1.

FN (False Negative): When the actual value is 1, but the predicted value is 0.

Accuracy: In general, it is a measure of how often the classifier predicts correctly. Eq. 1 shows how accuracy is calculated.

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{1}$$

Recall is known as the hit rate and is expected to be high. It is a measure of how many TP or TN values the classifier predicts correctly. Eq. 2 shows how the recall is calculated.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

Precision is a measure of how much is correctly and incorrectly predicted from all classes. Like recall, precision is expected to be high. Eq. 3 shows how precision is calculated.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

The **f-1 score** is a measure of the classifier's performance. It is the harmonic mean of recall and precision. Eq. 4 shows how the f-1 score is calculated.

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{4}$$

For each bounding box, **IoU** is used to measure the overlap between the predicted bounding box and the actual bounding box. Fig. 4 shows the structure. In the output given as a result of YOLOv7 training, Map values are the average accuracy (Average Precision) values calculated for different IoU (Intersection over Union) values. IoU is the ratio of the intersection of a prediction box with the real object to the area of the real object [44].

The following hyperparameters were used in model training: 0.0001 learning rate (lr0), 0.937 momentum, 0.0005 wt decay, and 35 batch sizes. Using the Stochastic Gradient Descent (SGD) optimizer for 300 epochs, the best model weights were recorded with 100 patience. Our training results on the YOLOv7 model showed positive results in terms of overall mAP and individual class performance. The model achieved an overall mAP@0.5 of 0.7512 and mAP@0.5:0.95 of 0.3925 on the validation set. For mAP@0.5 shows that the model has the ability to accurately detect and classify burn wounds with a high level of confidence. The accuracy value we specified for mAP@0.5:0.95 (0.3925) was found. This value is calculated according to mAP@0.5:0.95 values and this shows that it is difficult to detect small or similar wounds with a high threshold. It shows that the model is not able to detect certain object detection tasks at the desired level for the specified IoU range (Table 2). In the output given as a result of YOLOv7 trainings, mAP@0.5 and mAP@0.5:0.95 results are the average precision values calculated for different IoU (Intersection over Union) values. IoU is the ratio of the intersection of a prediction box with the real object to the area of the real object. mAP@0.5: This term refers to the average accuracy calculated over a single threshold value, usually 50% Intersection over Union (IoU). That is, it is the average accuracy calculated for forecasts with IoU values greater than 0.5 [45].

This indicates that the prediction boxes intersect the real objects to a large extent. That is, if the mAP@0.5 value for a YOLOv7 model is high, the model is able to accurately detect a wide range of objects, including small objects [44].

The YOLOv7 training output also includes the value mAP@0.5:0.95. This value is the average accuracy calculated for all IoU values between 0.5 and 0.95. This value can be used

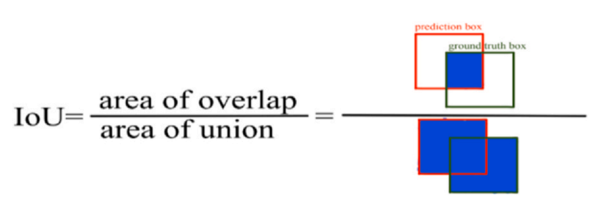

Fig. 4 – IoU (Overlap between predicted bounding box and actual bounding box).

Table 2 – Classification report of trained models.

Model	P	R	mAP@ 50	mAP@ 95
Yolov7	0.8081	0.6797	0.7512	0.3925

to evaluate the overall performance of a model for different IoU values [44]. mAP@0.5:0.95: This term refers to the average accuracy between different IoU threshold values. More specifically, mAP@0.5:0.95 measures the following:

The model's performance in detecting small objects: IoU values less than 0.5 make small objects more difficult to detect.

The model's performance in detecting medium-sized objects: IoU values between 0.5 and 0.75 generally provide sufficient sensitivity for the detection of medium-sized objects.

Performance of the model in detecting large objects: IoU values greater than 0.75 generally provide sufficient sensitivity for the detection of large objects [45].

For example, if the average accuracy values for all IoU values between 0.5 and 0.95 are 0.8, 0.9 and 0.7, the mAP@ 0.5:0.95 value would be (0.8 + 0.9 + 0.7) / 3 = 0.83. In general, high mAP@0.5 and mAP@0.5:0.95 values for a YOLOv7 model indicate that the model performs well [44].

Fig. 5 shows the box loss, object loss, and class loss values at each epoch of the training process. Box loss is the difference between the predicted and actual bounding box coordinates, while object loss represents the confidence score for each detected object in an image. Class loss represents the probability that each detected object belongs to a particular class. These results show that the YOLOv7 model is a successful tool for the accurate detection and classification of burn wounds and can positively contribute to medical diagnosis and intervention processes (Fig. 5). In artificial intelligence training, iteration is a process where data is used repeatedly to improve the performance of a model. In each iteration, the model is trained on the data and its performance is measured. If the performance is not satisfactory, the model is updated and trained on new data. This process continues until the performance of the model reaches an acceptable level. In the graph in Fig. 5, the x-axis represents the iteration.

The loss value measures how much the model's predictions are wrong or how much error it makes. This value shows how much the model's predictions deviate from the actual values.

The loss curve in the graph shows how the model's loss changes during the training process. Ideally, this curve should decrease during training. As the loss decreases during training, the model starts to perform better (BOX, objectness, classification, val Box, val objectness, val classification) (Fig. 5).

The Accuracy value is the percentage of correct predictions of the model. This shows how successful the model is on the given data.

The accuracy curve in the graph shows how the accuracy of the model changes during training. As the accuracy increases during training, the model starts to perform better ((Precision, recall, mAP@0.5, mAP@0.5:0.95) (Fig. 5).

The curve in Fig. 6 illustrates how precision changes for different confidence levels. In general, the graph shows that the higher the confidence level, the higher the precision. This

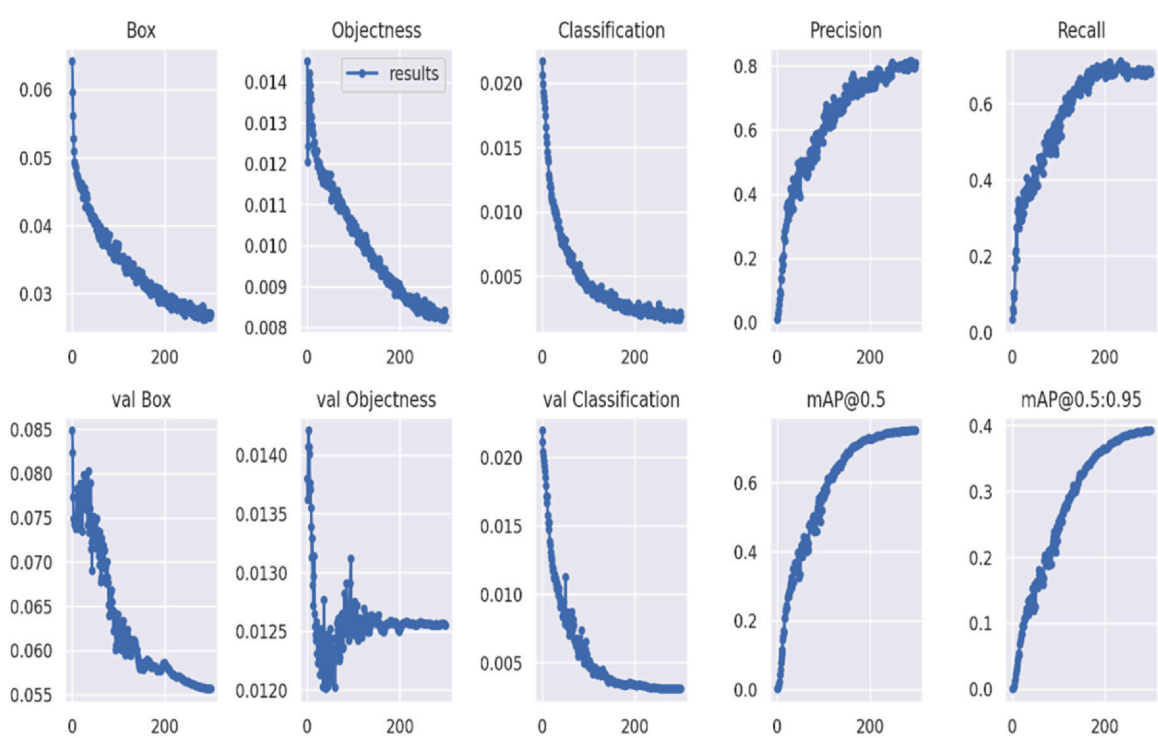

Fig. 5 – Detection of burn wounds according to the YOLOv7 model.

implies that the model produces more reliable results with forecasts with high confidence levels. A threshold, called the confidence-cut-off point (threshold), is chosen. This threshold determines the confidence level at which the model is acceptable, and forecasts that exceed this confidence level are considered positive (Fig. 6). According to Fig. 6, 0: represents 1st Degree, 1: 2nd Degree, 2: 3rd Degree. In YOLOv7 object detection models, the threshold selection is made to maximise the mAP@0.5:0.95 value of the model. mAP@0.5:0.95 value is the average accuracy calculated for all IoU values between 0.5 and 0.95. Threshold is a criterion used to determine whether a prediction box represents a real object [44]. The higher the threshold value, the more intersections a prediction box must have to represent a real object. The threshold value should be chosen depending on the test data and the requirements of the application. In this study, the threshold value is set to 0.30, as the detection of small objects is particularly important in the application.

Significant success has been achieved in the detection and classification of 2nd-degree and 3rd-degree burn wounds. Our proposed model shows high levels of accuracy in identifying and localizing burn wounds in images, achieving mAP@0.5 (average precision) values of 0.808 and 0.801, respectively. This shows that the model is able to accurately detect a significant number of burn wounds with high accuracy. Overall, the results of our YOLOv7 model show that it is successful in accurately detecting and classifying burn wounds in the images used for training and validation. These results highlight the potential of the YOLOv7 model to detect and classify burn wounds in medical images. The performance evaluation of the burn wound detection model at different levels is shown in Table 3.

Additional evaluation metrics were created to further assess the performance of our YOLOv7 model. These metrics helped us to delve deeper into the model's ability to accurately detect and classify burn wounds in images. The precision confidence curve, recall confidence curve, precision-recall (PR) curves, F1 score curves, and confusion matrices are shown in Figs. 7–10, respectively. These additional evaluation metrics include PR curves and F1 score curves, which show the balance between precision and recall for different decision thresholds, and confusion matrices, which show the number of true positives, true negatives, false positives, and false negatives for each class, providing a more comprehensive analysis (Figs. 7–10).

Table 3 – Performance evaluation of burn wound detection model at different degrees.

Classes	P	R	mAP@50	mAP@95
All	0.887	0.679	0.751	0.391
1. degree	0.774	0.547	0.639	0.314
2. degree	0.816	0.743	0.808	0.417
3. degree	0.832	0.748	0.801	0.443

3.2. *Prediction with mobile application*

In the second stage of the study, the accuracy of the model created with the mobile application was determined. The application was implemented and rigorously tested on an Android mobile device. In this work, we take an important step forward in the detection and classification of skin burn wounds. The mobile application we have developed has been tested with real-world hospital experiences and has proven its effectiveness in accurately detecting burn wounds. Our application offers a new perspective to the health sector and contributes to the effective management of burn wounds.

The Android application consists of two main parts, each offering separate functionality.

As part of our study, we first conducted a study on 50 subjects to test the app, all of whom had skin burns. The burn images were photographed and recorded, then verified by the mobile app, and the accuracy of the mobile app was

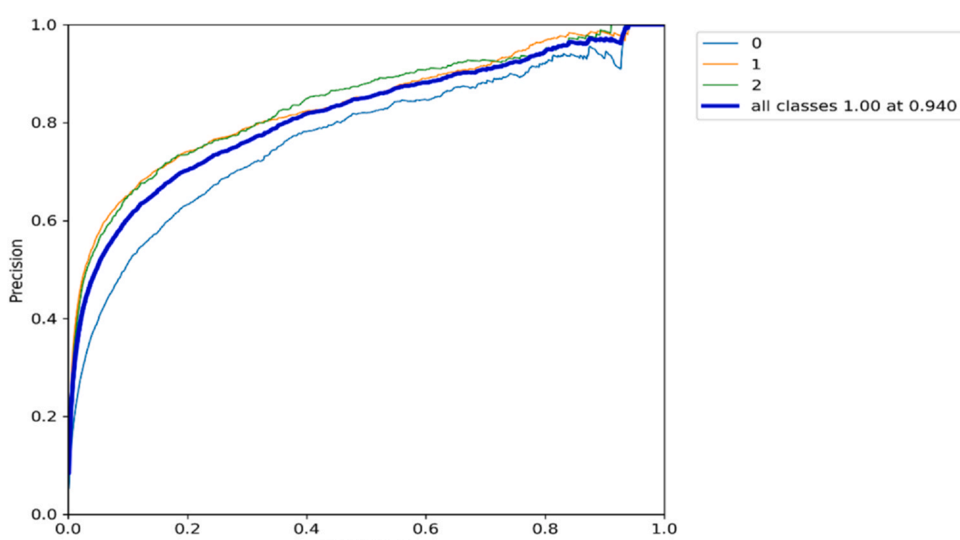

Fig. 6 – Precision confidence curve (0: 1.Degree, 1: 2.Degree, 2: 3.Degree).

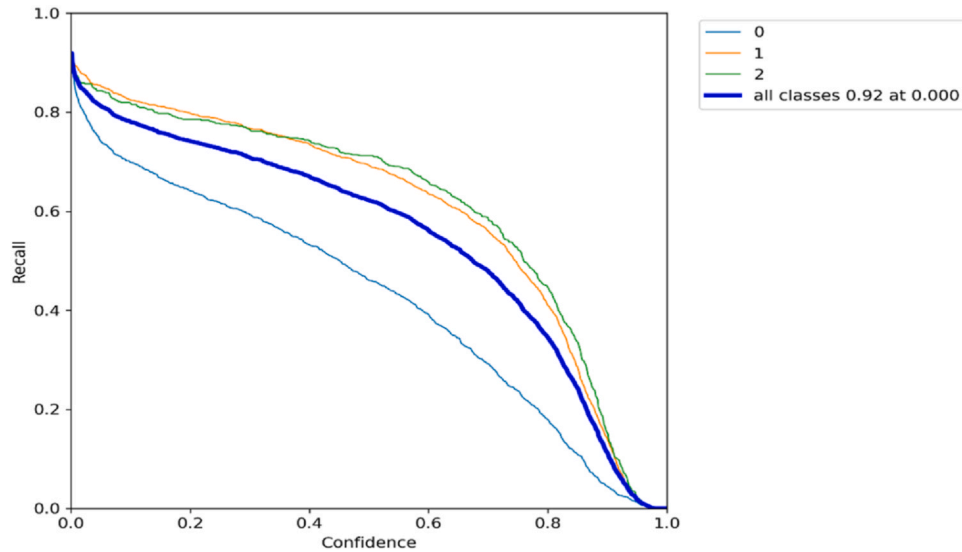

Fig. 7 – Recall confidence curve.

confirmed by healthcare professionals working in the emergency department. The burns were caused by hot water and were categorised into different degrees. Of the burns, 14 were 1st degree burns, 31 were 2nd degree burns and 5 were 3rd degree burns. The age range of the burned individuals was between 3–42 years. 40 of the data from 50 patients were correctly estimated. The success rate of our mobile application for 50 data was 80%. Efforts are underway to run the mobile application with IOS and to reach the peripheral regions.

3.2.1. *Part one: introduction and photo upload*

This section welcomes the user and asks them to upload a photo. The user can get this photo immediately by taking a picture of the wound on the patient's surface with the camera of his mobile phone, or he can upload a previously taken photo by selecting it from his gallery.

3.2.2. *Part two: automatic detection and results*

Once the photo is uploaded, the app automatically performs the detection process and redirects the user to the results page. On

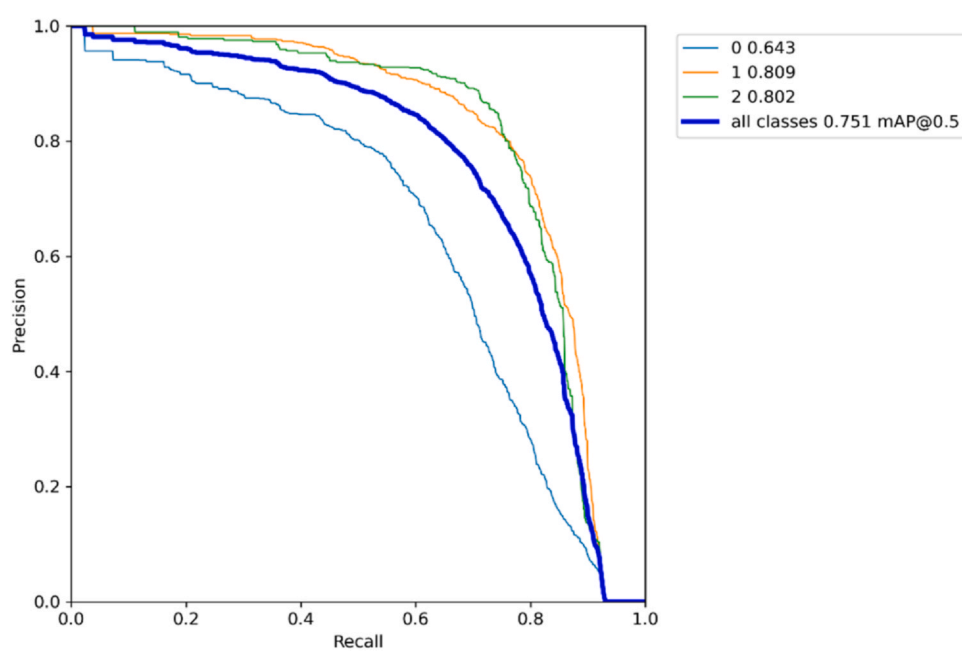

Fig. 8 – Precision–recall curve.

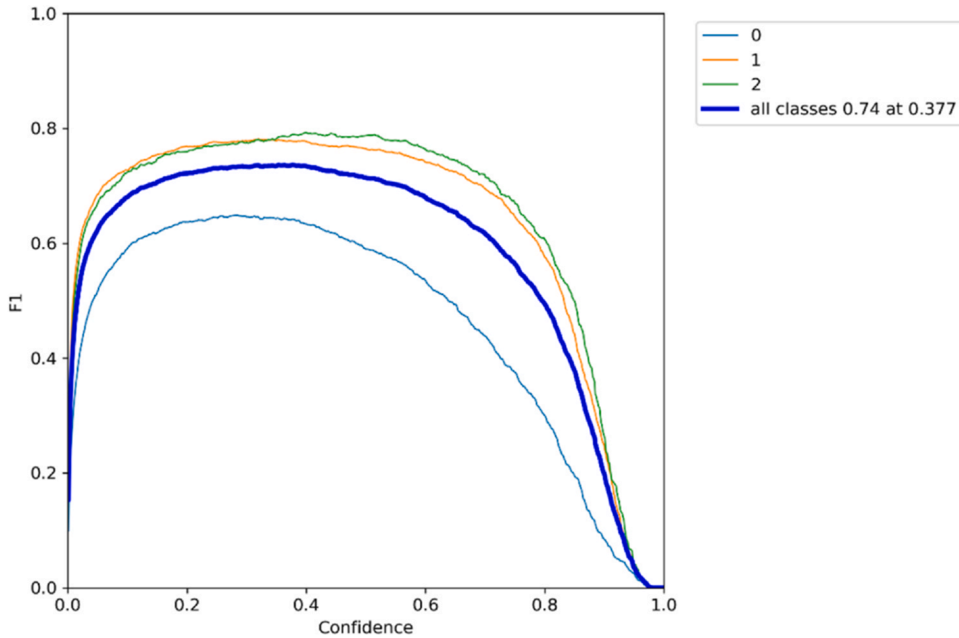

Fig. 9 – F1 score curve.

this page, the class and grade of the wound are displayed in a rectangular box. In addition, a reliability score is shown, indicating the degree of accuracy of the wound. This reliability score is an important measure of accuracy, indicating how accurately the model detected the wound on the surface.

3.2.3. Additional content and information

The app also provides the user with important information, such as which treatment methods to use and what to look for based on the severity of the wound detected. This helps the user to be more informed and knowledgeable when caring for patients.

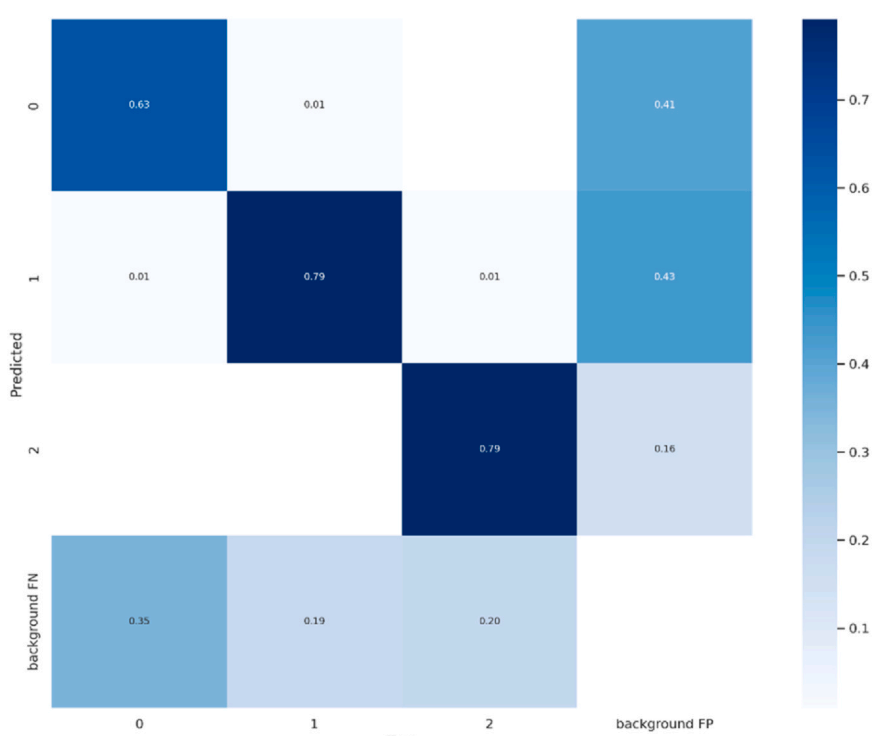

Fig. 10 – Confusion metrics.

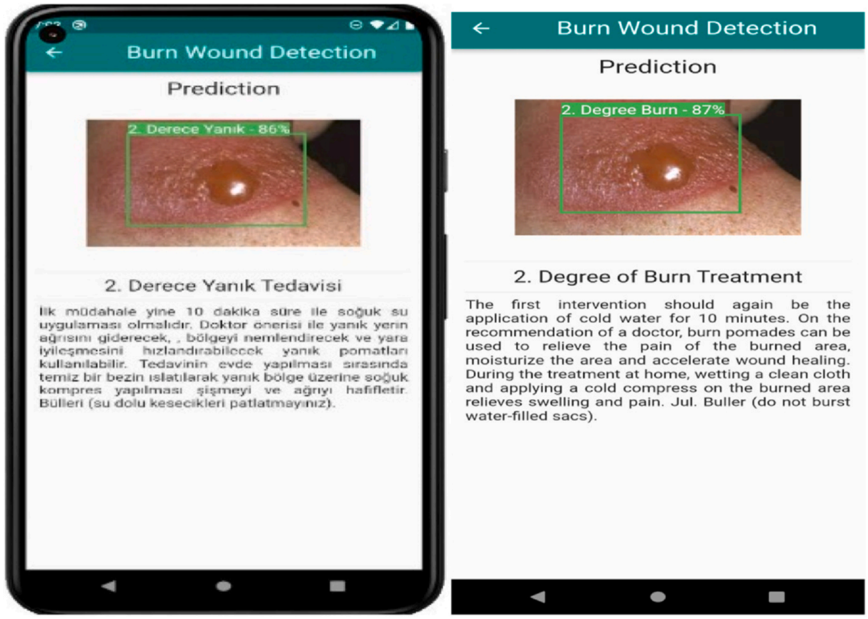

Fig. 11 – Mobile application test examples.

In this way, the app provides the user with both medical diagnostics and treatment recommendations and information, making it an important tool to improve patient care and achieve more effective outcomes. This shows that it is reliable and valid similar to the artificial intelligence model. Some mobile application output results are also given with sample figures. Looking at Fig. 11; the burn area is indicated by a frame. The degree of the burn and the percentage of correct prediction were determined. In addition, it is correctly reported which intervention should be done to the estimated degree (Fig. 11).

4. Discussion

Computer-based systems are widely used in medical image analysis. However, in many cases, the performance of these systems depends on the quality of the protocols followed. While burn image classification is a popular area of research, studies on the discrimination of burned skin from healthy skin are limited [46–48]. In this study, 75.12% of skin burn images were successfully detected. To achieve higher accuracy in our study, the quality of the data, the acquisition site, the number of data, and the class distribution are crucial. In some studies in the literature [49,50], the dataset used contains low-quality images, some of these images were collected from web pages, some were scanned from books, and some were obtained from hospitals.

Although there are many studies in the literature on the detection and classification of burn images using deep learning, this is the first study in this field with the YOLOv7 model. Our study is also the first in the detection and classification of burn wounds and the development of a mobile application. We also report a mAP@0.5 value of 0.75.12, which indicates that our model performs well in detecting and classifying burn images into different stages. Our model outperforms the benchmark models in terms of the amount and diversity of classes in the dataset used. The work of Wantanajittikul et al. [48] is another method that uses SVM to classify second and third-degree burns. The results of this method are compared with the results of Bayesian and k-nearest neighbor classifiers.

Deepak et al. [49] and Suvarna et al. [51] used k-nearest neighbor classifier, SVM, and template matching methods to classify burns into superficial, partial thickness, and full thickness burns [51]. Unlike Suvarna and Niranjan [51], in [52] ANN is used for the same purpose.

Badea et al. [46] handled pixel-based skin/burn area detection methods. They took 32×32 samples from sections of images marked as burned by expert surgeons and used these samples in the training phase of deep learning Convolutional Neural Networks (CNN). The study by Badea et al. presents the use of color imaging as a starting point for burn wound assessment by differentiating between healthy skin and burn wounds. Skin/burn area delineation is performed on a pixel-by-pixel basis based on the characteristics of the entire surrounding patch. Classification is learned under a scenario supervised by a deep-learned convolutional neural network, according to ground truth identified by expert surgeons from a large database of pediatric cases. In our study, both classification and grading are performed differently and a mobile application has been developed.

In another study by Badea et al. [53], the LeNet architecture is used for skin burn segmentation. Badea et al. used the LeNet classification model to classify 32×32 images for the detection of burn wounds. This study used the LeNet model to determine which class the images belong to. LeNet is a type of deep learning network, which has been particularly successful in recognition tasks with small-sized images, such as MNIST. Badea et al. used this model to classify burn wounds on images of size 32×32. In the work of Tran et al. [54], CNN is used to classify burns into four degrees. The local binary pattern operator was used as the input to the CNN model. Sabeena and Kumar [55] proposed a model based on an image-based supervised learning approach and a multi-scale super-pixel computer-based SVM classification with parallel propagation. Sabeena et al. conducted a classification study based on the color distribution of wounds using image processing techniques. In this study, a method was developed to determine which class the wounds belong to by using the information obtained from the color distributions of the wounds. The color characteristics of the wounds were examined through image processing techniques and the data obtained from these color distributions were used for classification. Chang et al. also analyzed burn images using deep learning [56]. Changa et al. utilized segmentation models to classify burn images and obtained successful results. In this study, burn images were classified and at the same time regionally delimited. In other words, both the classes to which the images belong and the regional locations of these classes on the image were determined. This method provides a detailed and comprehensive approach to classifying burn wounds and at the same time marking their specific regions.

The work by Changa et al. is closer to our work, but unlike this work, our work emphasizes speed and accuracy performance in real-time applications. While other studies generally focus on classification, our system is able to determine the regional locations of burn images by framing them in addition to class detection. In this way, if there are multiple degrees of burn regions in a single image, they can be easily separated by framing them separately. This creates a comprehensive classification system while at the same time enabling detailed analysis by identifying and marking individual burn areas. This can provide a more precise method of identifying and monitoring the severity of different areas of burn wounds.

When other studies similar to our study are analysed, it is determined that Karthik et al. worked with 104 images and classified them in 3 degrees. Our study contains more data than the study of Karthik et al. and the fact that the mobile application has been developed makes our study more powerful [57]. Liu et al. also found an accuracy similar to our study [34].

In our study, unlike the previous studies, we developed a model that classifies the degree of burn as well as the segmentation and detection of the burn.

In this work, we introduce an important innovation in the methods of detecting the grading of skin burn wounds. An in-depth review of the literature revealed that YOLOv7, an up-to-date object detection architecture for the classification and grading of skin burn wounds, has not yet been applied in this field. At this point, our work has been developed to demonstrate the usability of YOLOv7 for the detection of skin burn wounds. In particular, one of the most important contributions of this work is that we use this state-of-the-art object detection model for the classification and grading of skin burn wounds. Our model is capable of accurately classifying skin burn wounds into different grades. This model can help healthcare professionals to quickly and accurately assess the severity of burn wounds. However, this work is not limited to the development of a model. In addition, a mobile application was developed and this application has an interface that can be easily used by both hospital staff and the general public. This user-friendly application has great potential for grading burn wounds and sharing important health information. In conclusion, this study makes a great contribution to the health sector and society by presenting a YOLOv7-based model for rapid and effective assessment of skin burn wounds and a user-friendly mobile application. In addition, for the first time in Turkey, a mobile application was developed for the public to detect skin burn images with a mobile application.

Although the application developed by Wallis et al. is similar to our mobile application, our mobile application has an easier interface and provides faster feedback. In the app developed by Wallis et al. the user can point to specific injured body surfaces through a touchscreen interface and an integrated calculator estimates the total body surface area affected by the burn injury. Predefined standardized care recommendations, including total fluid requirement, are provided instantly by the software and case data is transmitted to a cloud server. A text message is automatically sent to the burn specialist during the call, who can then access the cloud server via smartphone app or web browser, review the case and images, and respond to the healthcare professional with both structured and personalized recommendations. In our study, the rapid response without waiting for the expert's feedback to the message shows that it is a better practice [58]. Goldberg et al. also developed a mobile application for burn area detection. In this application, only the burn area was detected and no grading was done. It is important that our study is more powerful than Goldberg et al.'s study in that it performs both surface area detection and grading in an ergonomic way [59].

This study has many original aspects. Firstly, preprocessing and data organisation techniques were applied to open access burn images. These applications made the data quantitatively and qualitatively unique. Instead of training the data directly with a model, a special approach was developed to process the data. These specialised processing methods have contributed greatly to the better understanding and utilisation of difficult and restricted access data in the healthcare field. These methods, which include cleaning the data, filling in missing values, feature engineering and other specialised data processing tasks, have improved the reliability of the results. When developing the model, we prepared the data in a special way using our unique data processing methods. This approach helped our model to perform better and achieve high accuracy. Another aspect that strengthens the originality of the study is the YOLOv7 architecture, which is an architecture used for the first time in the field of health, has both increased the accuracy of the study and pioneered new studies to be carried out. One of the most important and unique aspects of this study is its widespread impact on the field of health. In the literature, there are a limited number of studies on the use of artificial intelligence for public use. There is no similar study in the country where the study was conducted. In addition, the development of the mobile application makes it possible for individuals living in the periphery to intervene early in burns. The study has paved the way for both the prevention of malpractice in the interventions of students receiving health education for the degree of burns, the rapid processing of health personnel, and the public to access at any time and save money and time. In addition, early intervention has paved the way for a public health service that will reduce the complications of the patient and reduce the economic burden of both the individual and the hospital.

The burn mobile application we have developed makes a significant contribution to the public, health trainees and health personnel. Firstly, this application offers a great advantage to the public. Quick and accurate grading of burn wounds allows the public to understand the severity of the burn wound and intervene faster in an emergency. This can improve the treatment process of the burn wound and ultimately increase recovery rates. For health trainees, the practice can strengthen the teaching and learning processes. Detection and grading of burn wounds can help health trainees to improve their clinical skills. Furthermore, this practice provides an opportunity for prospective healthcare personnel to better understand the basics of burn wound treatment. For healthcare staff, this app offers quick access to up-to-date information and references on the classification of burn wounds. This can support treatment decisions and optimise patient management. It can also help healthcare professionals to keep their training up to date. In conclusion, the burn mobile application we have developed makes an important contribution to public health, health education and the practice of health professionals. By providing access to faster, accurate and accessible information on the detection and classification of burn wounds, this application can contribute to better outcomes in the healthcare sector and support the effective management of burn wounds.

5. Limitations

If the segmented region contains both healthy and burned skin, the extracted features from this region may cause the classifiers to be trained incorrectly.

Most segmentation methods can accurately cluster the pixels in an image when the number of clusters is specified correctly, but choosing the optimal number of clusters is a difficult process. In our study, having a light or dark skin color limits successful prediction when determining the degree of burn. The fact that the dataset contains low-quality images, these images are taken from web pages, and the number of images in the dataset is small is a limitation in the correct prediction of the model. A limitation of the study is that it was not able to detect close or small wounds at the desired level.

6. Conclusion

In this study, burn images were segmented and classified, and objects were detected using artificial intelligence, and a mobile application was developed. YOLO V7 model was used for the first time in burn images. For the first time in Turkey, a mobile application called Burn Wound Detection was developed to detect burn images. This study provides a new perspective and application to the literature in this field. First aid is crucial in burn cases, and it is a very important development for public health that people living in the periphery can quickly determine the degree of burn through the mobile application and provide first aid according to the instructions of the mobile application.

Patient consent statement

Written informed consent was obtained from the patients.

Consent/ethical approval

Approval for the study was obtained from the Scientific Research and Publication Ethics Committee of a university (Date and Number: 26.10.2021–293).

Patient consent statement

Patient consent form was obtained.

Research funding

No funding was received from any institution or organization during the conduct of the research.

CRediT authorship contribution statement

Conceptualization, Methodology, Software: **Metin Yıldız, Mehmet Okuyar**. Data curation, Writing– original draft: All authors. Visualization, Investigation: **Metin Yıldız, Mehmet Okuyar, Mehmet Yıldız**. Supervision: **Metin Yıldız, Mehmet Okuyar, Mehmet Yıldız**. Software, Validation: **Mehmet Okuyar, Mehmet Yıldız**. Writing – review & editing: All authors.

Declaration of Competing Interest

The author(s) declared no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Acknowledgement

We thank all participants who participated in the study. This study is dedicated to Cantürk ÇAPIK.

REFERENCES

[1] Feizkhah A, Mobayen M, Habibiroudkenar P, Toolaroud PB, Bejarpasi ZP, Mirmasoudi SS, et al. The importance of considering biomechanical properties in skin graft: are we missing something? Burns 2022;48:1768–9. https://doi.org/10.1016/j.burns.2022.07.005

[2] Mehrabi A, Falakdami A, Mollaei A, Takasi P, Ghorbani Vajargah P, Jafari H, et al. A systematic review of self-esteem and related factors among burns patients. Ann Med Surg 2022;84:104811https://doi.org/10.1016/j.amsu.2022.104811
[3] Greenhalgh DG. Management of burns. N Engl J Med 2019;380:2349–59. https://doi.org/10.1056/NEJMra1807442
[4] World Health Organization. Burns. 2018. Retrieved 01.09. 2023 Available from: 〈https://www.who.int/news-room/fact-sheets/detail/burns〉. [Accessed 1 September 2023].
[5] Bennett CV, Maguire S, Nuttall D, Lindberg DM, Moulton S, Bajaj L, et al. First aid for children's burns in the US and UK: an urgent call to establish and promote international standards. Burns 2019;45:440–9. https://doi.org/10.1016/j.burns.2018.09.003
[6] Moehrlen T, Landolt MA, Meuli M, Moehrlen U. Non intentional burns in children: analyzing prevention and acute treatment in a highly developed country. Burns 2019;45:1908–17. https://doi.org/10.1016/j.burns.2019.05.018
[7] Mengistu ND, Obsa MS, Gemeda LA. Burn pain management at burn unit of Yekatit 12 Hospitals, Addis Ababa. Pain Res Treat 2018;2018:1–7. https://doi.org/10.1155/2018/1092650
[8] Zemedie B, Sultan M, Zewdie A. Acute poisoning cases presented to the Addis Ababa Burn, Emergency, and Trauma Hospital Emergency Department, Addis Ababa, Ethiopia: a cross-sectional study. Emerg Med Int 2021;2021:6028123. https://doi.org/10.1155/2021/6028123
[9] Devaney L, Rowell KS. Improving surgical wound classification-why it matters. AORN J 2004;80:208–23. https://doi.org/10.1016/S0001-2092(06)60559-0
[10] Irfan-Maqsood M. Classification of wounds: know before research and clinical practice. J Genes Cells 2018;4:1–4. https://doi.org/10.15562/gnc.61
[11] Crouzet C, Nguyen JQ, Ponticorvo A, Bernal NP, Durkin AJ, Choi B. Acute discrimination between superficial-partial and deep-partial thickness burns in a preclinical model with laser speckle imaging. Burns 2015;41:1058–63. https://doi.org/10.1016/j.burns.2014.11.018
[12] Resch TR, Drake RM, Helmer SD, Jost GD, Osland JS. Estimation of burn depth at burn centers in the United States: a survey. J Burn Care Res 2014;35:491–7. https://doi.org/10.1097/bcr.0000000000000031
[13] Wearn C, Lee KC, Hardwicke J, Allouni A, Bamford A, Nightingale P, et al. Prospective comparative evaluation study of Laser Doppler Imaging and thermal imaging in the assessment of burn depth. Burns 2018;44:124–33. https://doi.org/10.1016/j.burns.2017.08.004
[14] Sen CK, Ghatak S, Gnyawali SC, Roy S, Gordillo GM. Cutaneous imaging technologies in acute burn and chronic wound care. Plast Reconstr Surg 2016;138:119s–128ss. https://doi.org/10.1097/prs.0000000000002654
[15] Ganapathy P, Tamminedi T, Qin Y, Nanney L, Cardwell N, Pollins A, et al. Dual-imaging system for burn depth diagnosis. Burns 2014;40:67–81. https://doi.org/10.1016/j.burns.2013.05.004
[16] Burmeister DM, Cerna C, Becerra SC, Sloan M, Wilmink G, Christy RJ. Noninvasive techniques for the determination of burn severity in real time. J Burn Care Res 2017;38:e180–91. https://doi.org/10.1097/bcr.0000000000000338
[17] Taib BG, Karwath A, Wensley K, Minku L, Gkoutos GV, Moiemen N. Artificial intelligence in the management and treatment of burns: a systematic review and meta-analyses. J Plast, Reconstr Aesthetic Surg 2023;77:133–61. https://doi.org/10.1016/j.bjps.2022.11.049
[18] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 1943;5:115–33.
[19] Ackley DH, Hinton GE, Sejnowski TJ. A learning algorithm for Boltzmann machines. Cogn Sci 1985;9:147–69. https://doi.org/10.1016/S0364-0213(85)80012-4
[20] Kelley HJ. Gradient theory of optimal flight paths. ARS J 1960;30:947–54. https://doi.org/10.2514/8.5282
[21] Dreyfus S. The numerical solution of variational problems. J Math Anal Appl 1962;5:30–45.
[22] LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86:2278–324. https://doi.org/10.1109/5.726791
[23] Hinton GE, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets. Neural Comput 2006;18:1527–54. https://doi.org/10.1162/neco.2006.18.7.1527
[24] Hossain S, Lee D-J. Autonomous-driving vehicle learning environments using unity real-time engine and end-to-end CNN approach. J Korea Robot Soc 2019;14:122–30.
[25] Kocić J, Jovičić N, Drndarević V. An end-to-end deep neural network for autonomous driving designed for embedded automotive platforms. Sensors 2019;19:2064. https://doi.org/10.3390/s19092064
[26] Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. Nat Med 2019;25:24–9. https://doi.org/10.1038/s41591-018-0316-z
[27] Miotto R, Wang F, Wang S, Jiang X, Dudley JT. Deep learning for healthcare: review, opportunities and challenges. Brief Bioinform 2017;19:1236–46. https://doi.org/10.1093/bib/bbx044
[28] Mousavi SM, Zhu W, Sheng Y, Beroza GC. CRED: a deep residual network of convolutional and recurrent units for earthquake signal detection. Sci Rep 2019;9:10267. https://doi.org/10.1038/s41598-019-45748-1
[29] Heaton JB, Polson NG, Witte JH. Deep learning for finance: deep portfolios. Appl Stoch Models Bus Ind 2017;33:3–12. https://doi.org/10.1002/asmb.2209
[30] Mohanty SP, Hughes DP, Salathé M. Using deep learning for image-based plant disease detection. Front Plant Sci 2016;7:1419. https://doi.org/10.3389/fpls.2016.01419
[31] Sladojevic S, Arsenovic M, Anderla A, Culibrk D, Stefanovic D. Deep neural networks based recognition of plant diseases by leaf image classification. Comput Intell Neurosci 2016;2016:1–11. https://doi.org/10.1155/2016/3289801
[32] Antony J., McGuinness K., Connor N.E.O., Moran K. Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. In: Proceedings of the twenty third international conference on pattern recognition (ICPR)2016. p. 1195–200.
[33] Seo H, Badiei Khuzani M, Vasudevan V, Huang C, Ren H, Xiao R, et al. Machine learning techniques for biomedical image segmentation: an overview of technical aspects and introduction to state-of-art applications. Med Phys 2020;47:e148–67. https://doi.org/10.1002/mp.13649
[34] Liu H, Yue K, Cheng S, Li W, Fu Z. A framework for automatic burn image segmentation and burn depth diagnosis using deep learning. Comput Math Methods Med 2021;2021:5514224. https://doi.org/10.1155/2021/5514224
[35] Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014;12:573–6. https://doi.org/10.1370/afm.1713
[36] Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. Pharm Ther 2014;39:356–64.
[37] Godwin Z, Tan J, Bockhold J, Ma J, Tran NK. Development and evaluation of a novel smart device-based application for burn assessment and management. Burns 2015;41:754–60. https://doi.org/10.1016/j.burns.2014.10.006
[38] Tocco-Tussardi I, Presman B, Huss F. Want correct percentage of TBSA Burned? Let a layman do the assessment. J Burn Care Res 2018;39:295–301. https://doi.org/10.1097/bcr.0000000000000613
[39] Wallace S, Clark M, White J. It's on my iPhone': attitudes to the use of mobile computing devices in medical education, a

mixed-methods study. BMJ Open 2012;2:e001099. https://doi.org/10.1136/bmjopen-2012-001099
[40] Boissin C. Clinical decision-support for acute burn referral and triage at specialized centres–Contribution from routine and digital health tools. Glob Health Action 2022;15:2067389. https://doi.org/10.1080/16549716.2022.2067389
[41] Skin burn degree classification. 2023. Available from: ⟨https://universe.roboflow.com/binus-if3z9/skin-burn-degree-classification⟩. [Accessed 10September 2023].
[42] Shafiee M.J., Chywl B., Li F., Wong A. Fast YOLO: A fast you only look once system for real-time embedded object detection in video. arXiv preprint arXiv:170905943. 2017. Available from: https://doi.org/10.48550/arXiv.1709.05943.
[43] Shinde S, Kothari A, Gupta V. YOLO based human action recognition and localization. Procedia Comput Sci 2018;133:831–8. https://doi.org/10.1016/j.procs.2018.07.112
[44] Padilla R, Passos WL, Dias TL, Netto SL, Da Silva EA. A comparative analysis of object detection metrics with a companion open-source toolkit. Electronics 2021;10:279. https://doi.org/10.3390/electronics10030279
[45] Dai J., He K., Sun J. Instance-aware semantic segmentation via multi-task network cascades. In: Proceedings of the IEEE conference on computer vision and pattern recognition;2016. p. 3150–8.
[46] Badea M.S., Vertan C., Florea C., Florea L., Bădoiu S. Automatic burn area identification in color images. In: Proceedings of the International Conference on Communications (COMM). 2016:Bucharest, Romania, pp. 65–8. Available from: https://doi.org/10.1109/ICComm.2016.7528325.
[47] Despo O, Yeung S, Jopling J, Pridgen B, Sheckter C, Silberstein S, et al. BURNED: towards efficient and accurate burn prognosis using deep learning; 2017. p. 1–8.
[48] Wantanajittikul K., Auephanwiriyakul S., Theera-Umpon N., Koanantakool T. Automatic segmentation and degree identification in burn color images. In: Proceedings of the fourth 2011 biomedical engineering international conference. 2012:Chiang Mai, Tayland, pp.169–73. Available from: https://doi.org/10.1109/BMEiCon.2012.6172044.
[49] Deepak L., Antony J., Niranjan U. Hardware co-simulation of skin burn image analysis. In: Proceedings of the ninteenth IEEE International Conference in High Performance Computing: Student Research Symposium Pune, India; 2012.
[50] Suvarna M, Venkategowda N. Performance measure and efficiency of chemical skin burn classification using KNN method. Procedia Comput Sci 2015;70:48–54.
[51] Suvarna M, Niranjan U. Classification methods of skin burn images. Int J Comput Sci Inf Technol 2013;5:109–18.
[52] Suvarna M, Kumar K, Niranjan U. Diagnosis of burn images using template matching, k-nearest neighbor and artificial neural network. Int J Image Process 2013;7:191–202.
[53] Badea M.-S., Felea I.-I., Florea L.M., Vertan C. The use of deep learning in image segmentation, classification and detection. arXiv preprint arXiv:160509612. 2016:1–5. Available from: https://doi.org/10.48550/arXiv.1605.09612.
[54] Tran HS, Le TH, Nguyen TT. The degree of skin burns images recognition using convolutional neural network. Indian J Sci Technol 2016;9:1–6. https://doi.org/10.17485/ijst/2016/v9i45/106772
[55] Beena B, Kumar P. Diagnosis and detection of automatic skin burn area color images identification of burn area depth in color images. Int J Eng Trends Technol 2017;48:48–54.
[56] Chang CW, Ho CY, Lai F, Christian M, Huang SC, Chang DH, et al. Application of multiple deep learning models for automatic burn wound assessment. Burns 2023;49:1039–51. https://doi.org/10.1016/j.burns.2022.07.006
[57] Karthik J., Nath G.S., Veena A. Deep learning-based approach for skin burn detection with multi-level classification. Advances in computing and network communications: proceedings of CoCoNet 2020, Volume 2: Springer; 2021. p. 31–40.
[58] Wallis LA, Fleming J, Hasselberg M, Laflamme L, Lundin J. A smartphone app and cloud-based consultation system for burn injury emergency care. PLoS One 2016;11:e0147253. https://doi.org/10.1371/journal.pone.0147253
[59] Goldberg H, Klaff J, Spjut A, Milner S. A mobile app for measuring the surface area of a burn in three dimensions: comparison to the Lund and Browder assessment. J Burn Care Res 2014;35:480–3. https://doi.org/10.1097/BCR.0000000000000037