

PAPER

Pressure Ulcer Detection and Classification Using Convolutional Neural Networks: Development of a Mobile Clinical Support Tool

Ana Huamani-Huaracca¹ ,
 Brigitte Torres-Pinedo¹ ,
 Juan Morales² ,
 Sebastian Ramos-Cosi¹ 

¹Image Processing Research Laboratory (INTI-Lab),
 Universidad de Ciencias y Humanidades, Lima, Peru

²Department of E-Health Research, Universidad de Ciencias y Humanidades,
 Lima, Peru

ahuamani@uch.edu.pe

ABSTRACT

The aim of this study is to develop a mobile application for the automated detection and classification of pressure ulcers (PU) using convolutional neural networks (CNNs). The methodology of this study was developed using a deep detection approach for the identification and classification of PU from clinical imaging. CNN methods were used to classify PU images according to their degree of damage. To evaluate its effectiveness, various CNNs architectures were compared and the model with the best classification accuracy was selected. The results showed that YOLOv11-CLS managed to achieve the highest overall comparative accuracy at 97.1%. Regarding the PU detection task, the model obtained a recall of 96.1%, a precision of 98.3%, an F1-score of 97.2%, and an accuracy of 98.1%. Likewise, for the multiclass severity classification, it managed to achieve a recall of 95.7%, a precision of 97.6%, an F1-score of 96.9%, and an accuracy of 97.9%, which guarantees a high efficiency of the application. In conclusion, the proposed mobile system represents a promising tool for the early detection of PU and can be used as a complement to improve the effectiveness of initial medical follow-up.

KEYWORDS

convolutional neural networks (CNNs), medical image classification, mobile clinical support tool, pressure ulcer (PU) detection and classification

1 INTRODUCTION

Pressure injury, also known as pressure ulcer (PU) is a persistent and high-impact health problem [1]. Globally, PU cases have doubled since 1990, exceeding 2.4 million in 2021 [2]. PU affects a high proportion of hospitalized patients, with a prevalence of 12.8% (95% CI: 11.8–13.9%) and an incidence of 5.4 (95% CI: 3.4–7.8) per 10,000

Huamani-Huaracca, A., Torres-Pinedo, B., Morales, J., Ramos-Cosi, S. (2026). Pressure Ulcer Detection and Classification Using Convolutional Neural Networks: Development of a Mobile Clinical Support Tool. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(10), pp. 222–241. <https://doi.org/10.3991/ijoe.v22i10.62606>

Article submitted 2026-06-10. Revision uploaded 2026-07-23. Final acceptance 2026-07-23.

© 2026 by the authors of this article. Published under CC-BY.

patient-days [3]. In intensive care units, the prevalence of PUs can reach up to 83% (95% CI: 71.2–90.7) [4].

Developed countries have a higher incidence of PUs, but lower mortality, while regions with low socioeconomic development bear the highest burden of disability and mortality [2, 5]. The PU acquired in the hospital represents a major source of burden for the health system. The national cost of PUs purchased from U.S. hospitals is estimated at \$26.8 billion annually and 60% of these costs are attributable to stage 3 and 4 injuries [6].

The origin of PUs depends on many factors such as the physiology of the patient and the time of exposure to mechanical forces [7]. In adults, key risk factors for the development of PUs include the mechanical conditions of the limits, e.g., increased interface pressure or BMI < 18.5, as well as factors affecting interindividual susceptibility (male sex, advanced age, anemia, hypoalbuminemia, diabetes, hypotension, low physical activity, existing pressure injuries) and treatment-related aspects, such as the length of stay in intensive care units [8]. In hospital emergency departments, most documented PUs is stage I. The most problematic anatomical locations were the sacrococcygeal region and heels [9].

PUs represents an important challenge in health care, both in the in-hospital and out-of-hospital settings. Its detection and timely classification depend largely on clinical observation, which can generate diagnostic variability and influence therapeutic decision-making [10]. In this process, the support of technology represents an opportunity for non-specialized professionals, students in the health area or for users with limited access to health services.

With the constant advancement of technology and computing, the use of digital tools and artificial intelligence (AI) plays an important role in the healthcare field, with applications that encompass the diagnosis, therapeutic management, monitoring, and prognosis of patients [11, 12]. Within this field, deep learning models applied to image recognition have gained special relevance, highlighting convolutional neural networks (CNN), a type of artificial neural network oriented to image processing, for the detection of patterns and the extraction of significant features, which have shown extraordinary potential in the analysis of medical images [13, 14].

In the case of PUs, the use of CNN has shown promising results in the detection and classification of a mobile tool based on CNN for the automated detection and classification of PUs, as a support for the clinical diagnosis and follow-up of patients in healthcare settings. Automated classification of lesions, achieving levels of accuracy comparable to and even higher than those obtained by specialized health professionals [15, 16, 17]. These findings support the potential of CNN-based tools as diagnostic support and follow-up in the management of patients with pressure ulcer.

Existing studies on PU analysis still have several limitations. For example, some of them focus only on a single task, such as binary lesion detection, wound segmentation, or classification of a certain limited number of stages, and employ small datasets or do not perform a comparison of multiple modern architectures under the same experimental conditions. Also, only a few studies integrate the selected model into a mobile application designed for practical use by medical personnel. Therefore, the motivation for this study lies in the need to develop a complete and accessible solution capable of not only automatically detecting PU but also classifying them according to their degree of severity. The main contribution of this study is the benchmarking of the YOLOv11-CLS, RT-DETR, ResNet50

and YOLOv8 models in a single clinical dataset, the selection of a more effective architecture and its integration into a mobile application to support clinical decision-making.

In this way, CNNs offer an innovative alternative to automate the analysis of lesions and support clinical judgment. However, the practical application of these technologies is still limited, and requires development, validation and performance evaluation. Therefore, the present study aims to develop and evaluate a mobile tool based on CNN for the automated detection and classification of PUs, as a support for the clinical diagnosis and follow-up of patients in health-care contexts.

2 RELATED WORK

Currently, CNN model-oriented research in the identification and classification of dermatological diseases, especially PUs, has made considerable progress both internationally and nationally, as evidenced in different recent studies. In Taiwan [17], deep learning algorithms were applied to design a mobile application to detect PUs and support clinical decision-making. To this end, Inception-ResNet-v2, AlexNet, VGG19 and ResNet50 were used, using 327 images of lesions initially classified as “erythema” or “non-erythema” and, subsequently, in “extensive necrosis”, “moderate necrosis” or “limited necrosis.” The superiority of Inception-ResNet-v2 was evidenced by obtaining a recall of 88.9%, accuracy of 100%, accuracy of 99.4% and F1 of 94%. In the classification of necrotic tissue, it achieved a recall of 94.7%, accuracy of 92.4%, accuracy of 94.8% and F1 of 93.5%, developing on this basis a mobile application capable of processing clinical images and providing recommendations to health personnel.

In Spain [18], a web application was developed for the diagnosis of PUs by analyzing 11030 clinical images, compared in Unet, FPN and PSPNet. The results indicated that Unet was the most efficient model, reaching an IoU (Intersection Over Union) of 75.6% and managing to recognize all classes, while FPN obtained an IoU of 71.48% and PSPNet of 75%, although both showed limitations by not identifying ulcers with necrotic tissue. Based on this model, the application was implemented with Flask, CSS, and HTML to articulate the interaction between client and server.

In Colombia [19], a system was proposed to detect and classify the different dermatological diseases through CNNs, using 15,500 images and the ResNet50 architecture. As part of the proposal, a website was implemented capable of uploading images, performing the corresponding classification and offering basic recommendations to health personnel. The results indicated that the system generated an approximate accuracy of 60.95% and an F1-score of 60.87%, in addition to a marked variability in performance by class, since some pathologies reached an F1 of up to 83.9%, while others did not exceed 50%. The evaluation of the system by health professionals showed a comprehensibility of 4/5, documentation 3/5, portability 5/5, interoperability 5/5 and analyzability and capacity for change 4/5, suggesting the need for improvements in these indicators.

In Peru [20], a CNN model has been developed for the diagnosis and classification of PUs from 11,030 images using the ResNet50v2 architecture. This resulted in an accuracy of 75% in qualifying, accuracy of 87%, recall of 69% and an F1 of 77%. In addition to an area under the curve of 83%, which supports its discriminative

capacity in clinical contexts. Likewise, [21] a study was proposed with the aim of identifying and predicting various skin diseases more accurately through the use of CNN from 10,000 images. A comparison was made between VGG-19, DenseNet201, ResNet50, Inception V3 and EfficientNet. The findings confirmed that DenseNet201 obtained the best performance, with a weighted accuracy of 90.9%, a recall of 91.3% and an F1 of 98.3%, outperforming VGG-19 which reached slightly lower values. In contrast, the other architectures presented heterogeneous and less consistent performances.

Although previous studies have demonstrated the feasibility of CNN-based systems for analyzing PU, several limitations remain. For example, some of the studies used relatively small datasets, focusing only on binary classification or necrotic tissue identification, and some of them even implemented web rather than mobile solutions. In addition, direct comparisons between recent architectures of both classification and detection under identical training conditions are still limited. Unlike previous approaches, the present study compares YOLOv11-CLS, YOLOv8, ResNet50, and RT-DETR using the same clinical dataset and hyperparameters, and integrates the best-performing model into a functional mobile application. Table 1 summarizes the main characteristics, performance metrics, platforms, and limitations of the related studies.

Table 1. Comparative analysis of related studies on pressure ulcer detection and classification

Study	Dataset	Task	Model	Main Metric	Platform	Main Limitation
Liu et al. [17]	327 images	Detection/classification	Inception-ResNet-v2	F1 94%	Mobile	Small dataset
Aguilar [18]	11,030 images	Segmentation	U-Net	IoU 75.6%	Web	Segmentation only
Santis et al. [19]	15,500 images	Skin disease classification	ResNet50	Accuracy 60.95%	Web	Multiple pathologies
Proposed study	2,013 original/ 7,621 augmented	Detection and grading	YOLOv11-CLS	Accuracy 97.9%	Mobile	External validation pending

3 METHODS AND MATERIALS

The methodology of this study is developed based on a deep detection approach for the identification and classification of PU from clinical images. The process is structured in several interconnected phases, from initial preprocessing and identification of the presence of ulcers, to multi-categorical classification by severity. In this sense, it was used to use a set of labeled data for the training and validation of different CNNs, including various computer vision models, in order to evaluate and select the architecture that offers the best data accuracy. This approach allows the development of a robust system capable of optimizing the diagnosis and clinical management of these lesions, culminating in the design of an application prototype for its own validation.

3.1 Dataset

The images used were retrospectively extracted from clinical records from three representative hospital centers located in different districts of Lima, collecting a total of 5,760 images between 2020 and 2025, such data come from people who went through an informed consent before providing any information. These selected hospital centers correspond to the centers with which the University of Sciences and Humanities (UCH) has an agreement for pre-professional internships of its students of health careers. Likewise, this study has the approval of the ethics committee of said university (code-049-2025).

After a filtering process, those images that presented distortions, low resolution, blur, duplicity or that did not correspond to PU, such as those of diabetic origin, were discarded in order to reduce biases due to generalization, leaving a total of 2,013 images for inclusion. Subsequently, all the images were converted to JPG format and compressed, with the aim of standardizing their storage and facilitating their computational treatment.

The 2,013 images selected were divided into 92% for training ($n = 1,852$), 5% for testing ($n = 101$), and 3% for validation ($n = 60$). This distribution was chosen due to the limited amount of clinical data and the need to provide the model with as many images as possible so that the model can learn the visual characteristics of the various stages of PU, keeping the sets separate for adjustment and final performance assessment. Likewise, the images were included in one of the sets, which excluded the simultaneous presence of related images in the training, validation and testing sets, as well as reducing the risk of data leakage. During preprocessing, auto-orientation was applied and the resolution was adjusted to 640×640 pixels. Subsequently, data augmentation techniques were implemented, including horizontal and vertical flipping, rotation in a range of $\pm 15^\circ$, deformations in the horizontal and vertical axes of $\pm 15^\circ$, and the addition of noise up to 5% of the pixels. As a result, the dataset was expanded to a total of 7621 images, which were used in the training phase of the model.

3.2 PHASE 1: CNN classifier selection

An exhaustive search was carried out for the selection of the most suitable CNN model for the analysis of PUs. This stage is fundamental and includes the breakdown of several activities, including; the analysis of data requirements and labeling to standardize the figures, the comparison of CNN models and the evaluation of metrics for the dataset. This seeks to establish a technical basis for selecting the model that demonstrates greater efficacy and accuracy in the classification of ulcers.

Analysis of data and label requirements. In this section, we sought to describe the phase on image identification, which seeks to discern whether a figure inserted as data contains PU or not. Figure 1 shows the identification process using two classification values: whether it is a PU or not a PUs. The diagram details the workflow, which begins with the input of a clinical image of the patient, towards a detection of characteristic parameters. This filtering stage is crucial to ensure that only relevant images make it to the next stage of classification, in order to improve the efficiency of subsequent analysis.

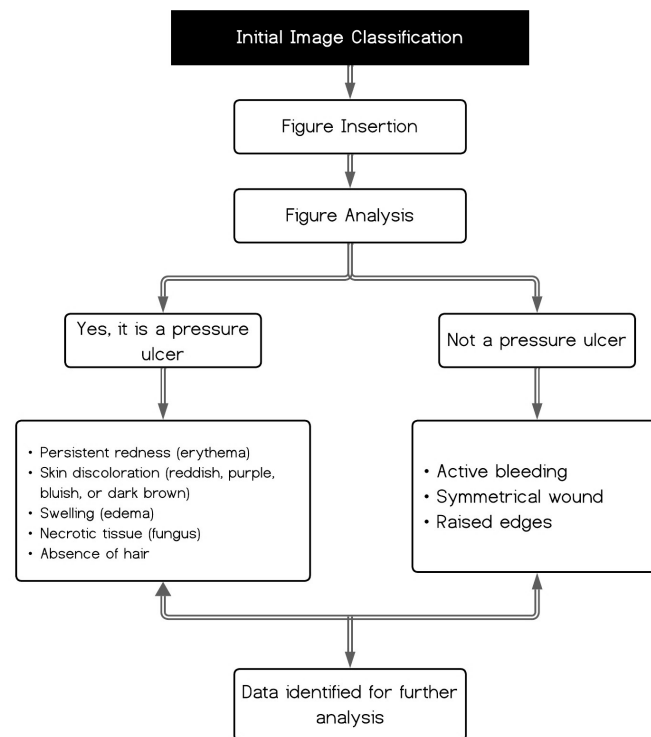


Fig. 1. Initial image classification diagram

Likewise, the images identified as PUs are classified multicategorically, where clinical characteristics, considered as hyperparameters, are used in order to train the model in the correct categorization of degrees of severity of the lesion according to the classification of the European Pressure Ulcer Advisory Panel (EPUAP) in its 2019 edition and the ICD-11 WHO in 2018, the most recent editions [2]. Figure 2 shows the classification details for each grade, considering their characteristics to be found or identified in that grade.

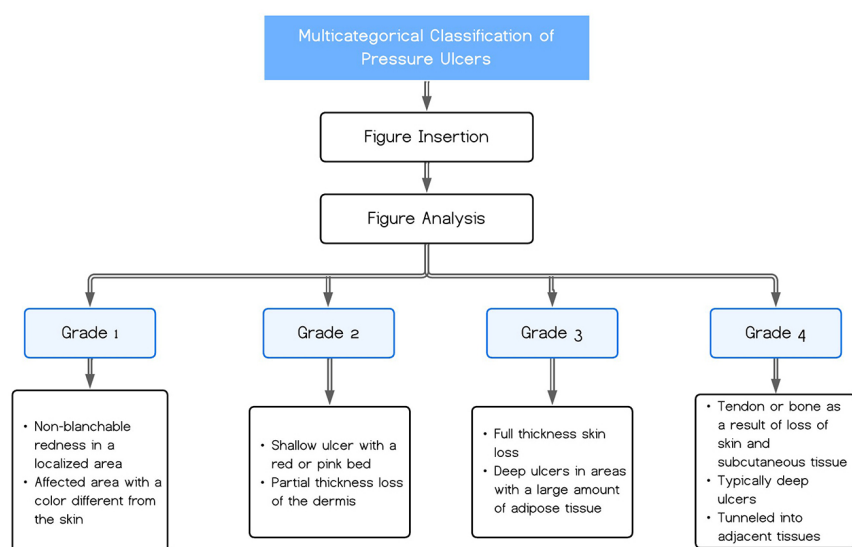


Fig. 2. Multi-categorical classification of PUs

Comparison of CNN models. The YOLOv11-CLS, ResNet50, YOLOv8 and real-time detection transformer (RT-DETR) models were selected because they are recognized in recent literature as state-of-the-art architectures in the field of CNNs applied to the detection and classification of skin lesions and ulcers of hospital origin. These models stand out for their ability to recognize complex patterns, and were chosen considering key criteria such as accuracy, speed of inference and portability, evaluating their respective strengths and limitations in clinical contexts.

YOLO11-CLS. It is a CNN that represents the evolution of the YOLO model family, optimized for image classification. Its modular architecture contains an efficiency for feature extraction, using multiple convolutional layers, normalization layers, and activation functions.

ResNet50. It is an advanced CNN that uses residual connections to mitigate the problem of gradient fading in very deep models. Its architecture allows information to flow through the convolution and clustering layers, ensuring that key characteristics of ulcer images are maintained during processing, even at great depth.

YOLOv8. It is one of the latest models in the YOLO model family, known for its balance between inference speed and high accuracy. Its modular architecture, designed to be highly efficient, incorporates significant improvements to its detection head, allowing for robust feature extraction and rapid image analysis.

RT-DETR. It is a hybrid architecture that combines the efficiency of a CNN with the contextual processing power of a Transformer model. This technology shows a balance between speed and accuracy by using an optimized encoder-decoder design.

The 7621 images enlarged for training were used by applying 25 epochs, 3×3 filters and a weight decay of 0.0005 to avoid overfitting. This process was executed on Google Colaboratory Pro+ using NVIDIA Tesla T4 GPU and 128 GB of system RAM. The four CNN models were evaluated under the same conditions and with identical hyperparameters, which allowed a fair comparison of their performance in terms of accuracy and robustness.

Validating metrics for the dataset. The performance of the models was validated using the precision, recall, and F1-score metrics for the selection of the highest scores, which indicated greater effectiveness. In this sense, accuracy is the percentage of predicted positives that are correctly predicted, which is interpreted as the division of correct classifications (PTs) by the sum of PTs and false positives (FPs), as expressed in the following formula:

$$Precision = \frac{TP}{TP + FP}$$

The recall parameter is the percentage of actual positives that are correctly predicted, which is interpreted as the division of the TPs by the sum of TPs and false negatives (FNs), as follows:

$$Recall = \frac{TP}{TP + FN}$$

The F1-score value is a measure that evaluates the performance of the predictive model when the size of the classes differs considerably, for this precision and retrieval are used and expressed through the following formula:

$$F1\ score = \frac{2}{Recall^{-1} + Precision^{-1}} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

In addition, to determine the proper classification of the models, a manual validation was carried out by a group of three doctors specializing in general surgery and plastic surgery, who carefully evaluated the labels granted by the models to determine if they are adequate. With this, the model with the best overall performance and greatest agreement with the validation of the specialists is selected, and the best and last files corresponding to the selected CNN model are preserved.

It should be mentioned that the performance of the model was evaluated in two complementary tasks, firstly, it consisted of determining the binary detection to see if an input image contained a PU. Second, it consisted of multiclass classification to assign them to the PU detected according to one of the four degrees of severity. In addition, precision, recall, F1 score, and accuracy were calculated separately for both tasks. On the other hand, the general comparative accuracy of each architecture was used to perform an analysis and thus the model with the best performance under the same experimental conditions can be selected.

3.3 PHASE 2: Multiplatform mobile app deployment

This phase focuses on the development and deployment of the cross-platform mobile application interface, turning the optimal model found into a functional and useful clinical tool. This stage is divided into several key components that allow you to obtain the necessary tools for a successful deployment.

Design and prototyping. The design of the application focuses on the creation of an interactive prototype in Figma to visualize the user flow and interface. The application is oriented towards a shortcut to speed up the experience, eliminating the need for a login, but starting with a shortcut to the camera for the capture of the image to be analyzed. This simplicity is an effective privacy and usability consideration. The app includes a popup-type rating so that users can provide feedback on the accuracy of the model after each analysis, which will improve the quality of the app in the long run. Table 2 shows the functional and non-functional requirements required for the development of the application system.

Table 2. Functional and non-functional requirements

Type of Requirement	Description
Functional	
RF01: Accessing the Camera	The app should allow users to take photos of the ulcers.
RF02: Uploading Images	The application must process the captured or uploaded images.
RF03: PU Identification	The application must identify the presence or absence of PUs in the image.
RF04: PU Classification	The application should classify the pressure ulcer of the image according to the grade to which it belongs.
RF05: Feedback	The app must display a rating popup to the user to assess the accuracy of the diagnosis.
Non-functional	
RNF01: Usability	The user interface should be intuitive and easy to use.
RNF02: Performance	The response time of the classification must not exceed 5 seconds.
RNF03: Compatibility	The app must be compatible with Android and iOS operating systems.

CNN model integration. It focuses on adapting the selected model to work optimally in the mobile environment. This involves using lightweight libraries such as TensorFlow Lite or PyTorch Mobile, designed specifically for devices with limited resources, such as the Cortex processors found in most current cell phones. This is intended to optimize the analysis of the model from the experimentation phase to the practical experience of the end user.

Application development. It seeks the implementation of the prototype and integration of its main components. The app makes use of the device's camera as the main input tool, allowing healthcare personnel to capture images of PUs in real-time. The user interface, previously designed in Figma, is developed in an intuitive and functional environment, in order to guide the user through the process of capturing, analyzing, visualizing results and evaluating the analysis by the user.

4 RESULTS

4.1 PHASE 1: CNN classifier selection results

In this first phase, the selection of the CNN classifier for the identification of PUs was carried out, this section allows the selection of the model that demonstrated greater efficiency and accuracy in the classification of ulcers. For this, the labeling of the clinical images used during the training of the model was first carried out. Therefore, Figures 3 and 4 illustrate the labeling process for standardization of figures, which was carried out prior to the classification of the most appropriate CNN model.

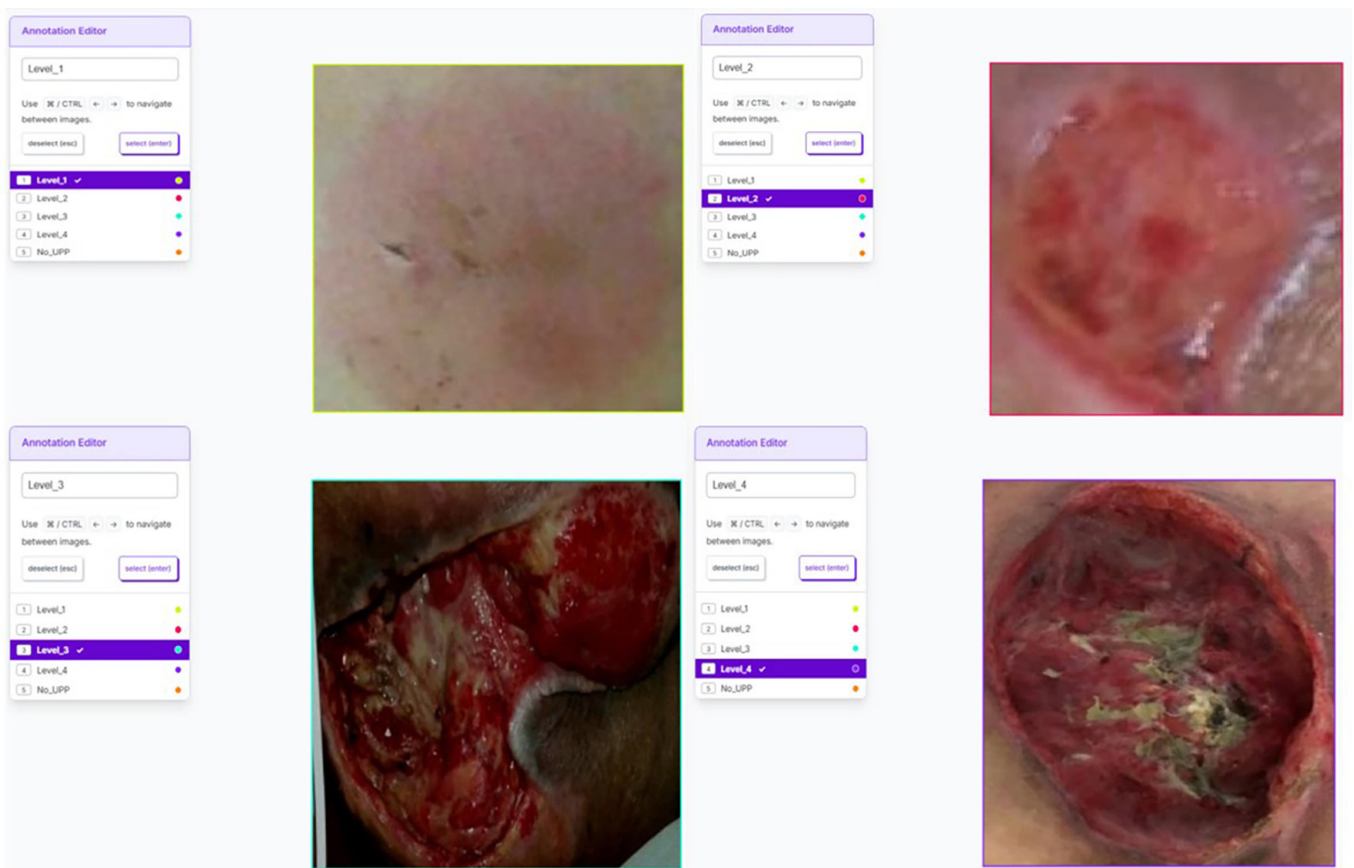


Fig. 3. Process of labeling and classification of clinical images for the training of the 4 degrees of UP

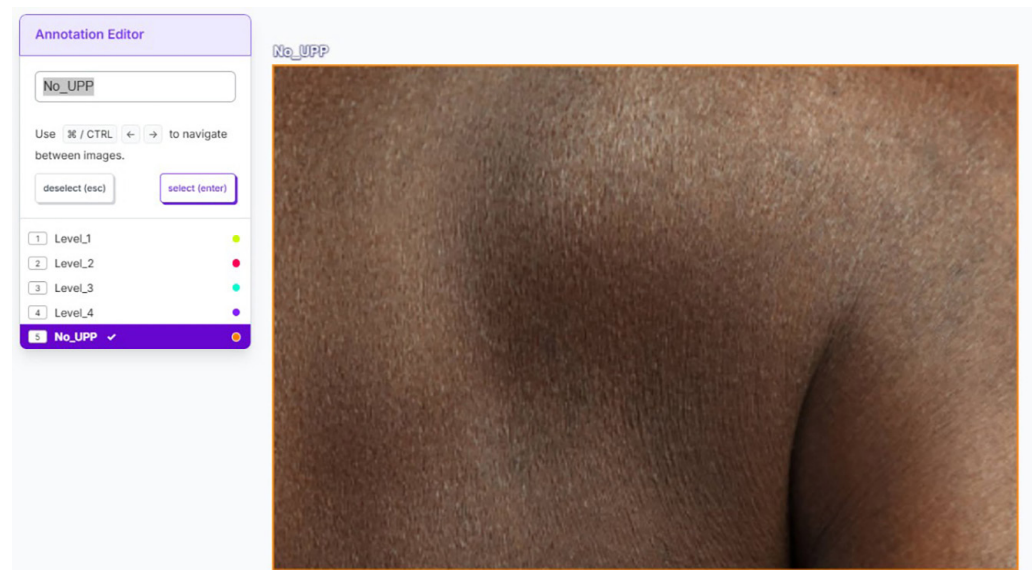


Fig. 4. Clinical image labeling and classification process for healthy skin training

In addition, before the figure labeling process, and based on the result obtained, the most appropriate CNN classifier was selected. Figure 5 shows the average classification accuracy of PU grades achieved by each of the included CNNs. Likewise, after a comparative evaluation, it was shown that YOLOv11-CLS achieved the highest overall accuracy of 97.1%, followed by RT-DETR with 95.3%, ResNet50 with 94.8% and YOLOv8 with 92.3%. Thus, after having obtained the results of the models, the YOLOv11-CLS model was selected as the most optimal architecture to integrate into the mobile application.

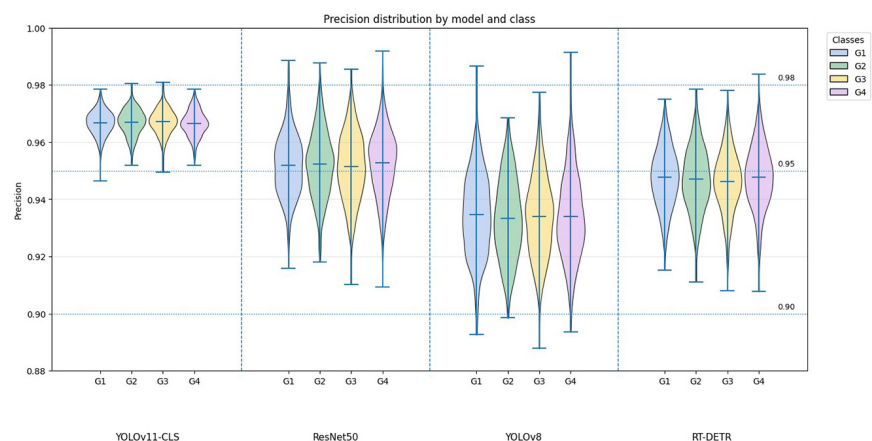


Fig. 5. Accuracy distribution by CNN and UPP grades

Detailed performance of YOLOv11-CLS. After having selected YOLOv11-CLS as the architecture with the best performance, an analysis was performed to see the performance separately for PU detection and the multiclass severity classification, which is detailed in Table 3. Where it was possible to achieve in the binary detection task, a recall of 96.1%, an accuracy of 98.3%, an F1 score of 97.2% and an accuracy of 98.1%. Therefore, these results could be interpreted as the model demonstrated a high capacity for the identification of PU images, maintaining a low proportion of FP. On the other hand, in the multiclass classification task, YOLOv11-CLS managed to obtain a recall of 95.7%, an accuracy of 97.6%, an F1 score of 96.9%, and an accuracy

of 97.9%. However, although these values were slightly lower than those obtained in binary detection, the model maintained a balanced performance by being able to distinguish between the four grades of pressure ulcers.

Table 3. Detailed performance of YOLOv11-CLS

Evaluation Task	Precision	Recall	F1-Score	Accuracy
Pressure ulcer detection	98.3%	96.1%	97.2%	98.1%
Pressure ulcer grade classification	97.6%	95.7%	96.9%	97.9%

Confusion matrix analysis

		Predicted class					
		Level_3	Level_2	Level_1	Level_4	No_UPP	False negatives
Actual class	Level_3	180	4	0	2	0	7
	Level_2	2	162	3	0	0	5
	Level_1	1	1	168	0	0	3
	Level_4	2	0	0	176	0	5
	No_UPP	0	0	0	0	15	0
	False positives	0	0	0	0	0	0

Fig. 6. Confusion matrix obtained by YOLOv11-CLS for pressure ulcer severity classification

Figure 6 presents the confusion matrix obtained by YOLOv11-CLS for PU severity classification. The rows represent the actual classes, while the columns indicate the classes predicted by the model. Likewise, values that are on the main diagonal are those that correspond to the correct predictions, while values outside the diagonal represent erroneous classifications between severity levels.

On the other hand, the model correctly classified 180 Level_3 images, 162 Level_2 images, 168 Level_1 image, and 176 Level_4 images. It can be seen that, among the categories mentioned, Level_1 showed the highest rate of class-specific recognition, with 168 correctly classified images out of 173 cases evaluated. This is followed by Level_4, with 176 correct predictions out of 183 cases; Level_2, with 162 out of 172; and Level_3, with 180 out of 193. Also, the 15 images corresponding to class No_UPP were correctly identified, although this result must be interpreted considering the smaller number of samples available in this category.

It should be mentioned that most of the classification errors occurred between adjacent or visually related severity levels. For example, in the Level_3, four images were classified as Level_2 and two as Nivel_4. For the Nivel_2, two cases were assigned to the Nivel_3 and three to the Nivel_1. The Nivel_1 showed only two interclass errors, one classified as Nivel_3 and one as Nivel_2. In the Nivel_4, two images were incorrectly classified as Nivel_3. Therefore, these results indicate that the model maintained a high ability for it to distinguish the four levels of PU, with limited confusion concentrated in which it was mainly focused between classes that share similar visual characteristics.

Finally, the FN column showed seven undetected cases for Nivel_3, five for Nivel_2, three for Nivel_1 and five for Nivel_4. No false-positive background assignments

were recorded in this evaluation. Therefore, it can be interpreted in a general way that the concentration of values along the main diagonal and the low number of errors outside the diagonal demonstrate the stable and consistent performance of the YOLOv11-CLS model in the classification of pressure ulcers.

4.2 PHASE 2: Results of the multiplatform mobile application deployment

In the second phase, corresponding to the mobile application oriented to the visualization of results of PU according to classification, a separation of results by sections is shown, which are related to the design, integration and development of the application. Each result is evidenced in the following sections.

Design and prototyping. In the design and prototyping section of the mobile app, a prototype of pre-classification is displayed. In this sense, Figure 7 consists of 4 parts, first illustrating the section of taking a photograph for the identification of the figure to be inserted, it is the first module to appear, since this application does not contain a login with credentials for greater ease of access. Secondly, it shows the quick preview option of the figure, in which the options to download photo and show identified information are shown. Thirdly, it shows the option to show identified information, which consists of the data inserted in the source code, which allows identifying a type of ulcer degree according to the inserted image and providing the option to show information about said discovery. Fourth, it shows the option in an information popup, where the user can give information related to user satisfaction, illustrated on a scale of 1 to 5 in quality.

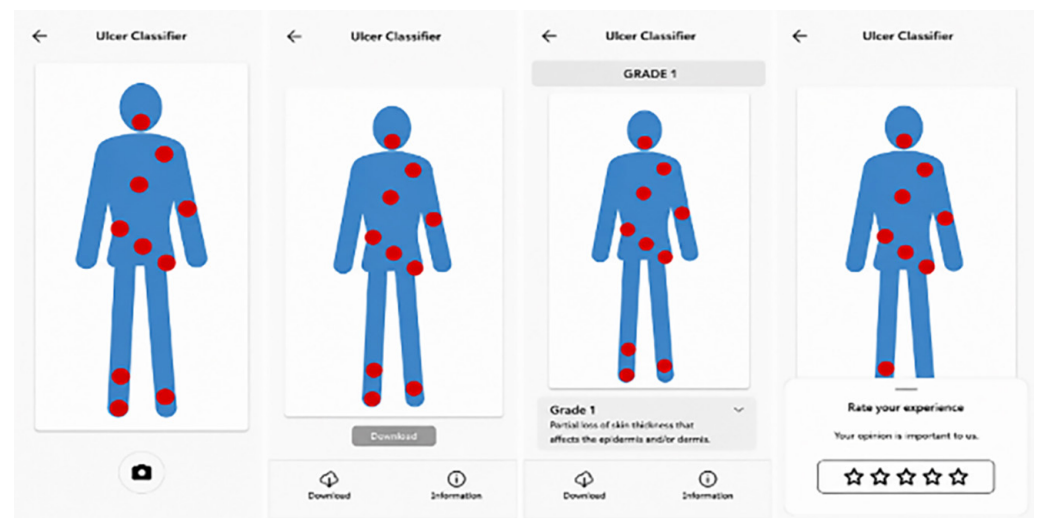


Fig. 7. Mobile application interface for pressure ulcer location and preliminary classification

CNN model integration. Before starting with the development, the planning of the operating flow and the technologies to be used for the application was carried out. Therefore, the architecture of the system is presented, which can be visualized as the interaction between the user interface, local data processing and an AI system based on CNN. As shown in Figure 8, the system starts with the application's user interface, where it accesses the main modules of the application: the home screen, camera, image upload, processing screen, result display, and previous analysis history. On the other hand, the interface of the application was made using

the language of Kotlin and Jetpack Compose, in order to provide a modern and intuitive user experience.

After selecting or capturing an image of the affected skin area, the app performs local data processing on the mobile device. This includes validating the uploaded image, converting the file to Base64 format, managing URIs, and preparing the information for later transmission to the inference server. In addition, the local storage of images and analysis results was implemented through the HistoryStorage module, which allows the user to review previous studies, i.e. history, and thus track them. The next step is interaction with the API service. To do this, the Retrofit, OkHttp and Gson libraries were used for data exchange, thus facilitating secure HTTP communication, data serialization and the reception of responses from the AI model in JSON format. The image is sent to an external inference service, where it is processed using a pre-trained CNN model.

In the AI layer, the app employed a CNN model trained on a clinical dataset of images of PU of varying severity, as well as healthy skin. After receiving the image, the model automatically performs inference to determine the expected severity or degree of the PU, along with the confidence level of the prediction. Finally, in the final stage, the results are returned to the mobile app and viewed by the user on the PU classification information, the model's confidence percentage, and the date and time of the analysis. Moreover, all results can be saved locally for later review and monitoring of the patient's condition.

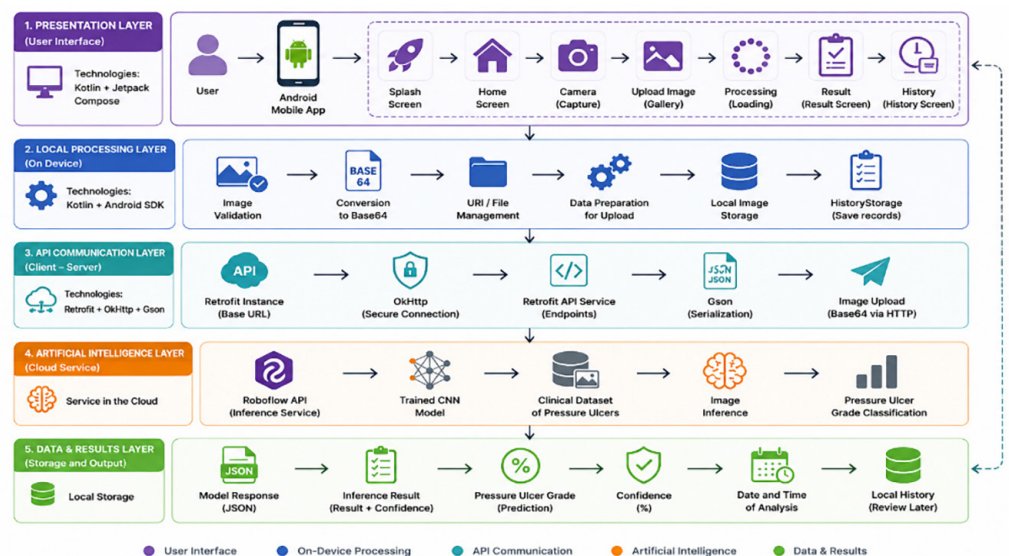


Fig. 8. UlcScan system architecture

App development results. This section introduces the main interfaces of the mobile application for PU detection. Next, the most relevant modules of the system and the functionality in user interaction with the application will be shown.

Main interface and navigation. The main navigation is shown in Figure 9. The first interface can be visualized to display the home loading screen, displaying the UlcScan logo and a brief description of the software's purpose. After a three-second loading screen, the user is automatically redirected to the main interface. This interface is the system's navigation module and provides access to three sections of the application. The first option, "Take Photo," allows users to capture a

clinical image in real time using the device camera. The next access is the “Upload image” button, which makes it easy for the user to upload an image as a second option from the device’s gallery for later analysis. Next, there is the “History” button, where the user can access previously saved analysis results, including images, model confidence level, and date of the analysis. Each element of the interface was designed so that the user can access the application in a simple way, without having any technical knowledge for its use, minimizing the user’s interactions with the system. Moreover, the design is adaptable for any mobile device of different sizes.



Fig. 9. UlcerScan mobile app main navigation

Capturing and uploading images. In this phase of the system’s operation, it was implemented as previously mentioned that the user will have two ways to send an image to analyze it, the first is through the device’s camera and the second through the selection of an image through the device’s gallery. An example of this can be seen in Figure 10, where the first photo is focusing on the affected area that you want to capture in real time and the second interface, shows that the user has already selected a photo from their gallery of their device to send it for later analysis, in that same section the user also has the option to change the image in case it is not the correct one blue “Change image” button. As you can see, the app shows you a preview, allowing the user to visually verify the correctness of the selected clinical image before starting the analysis.

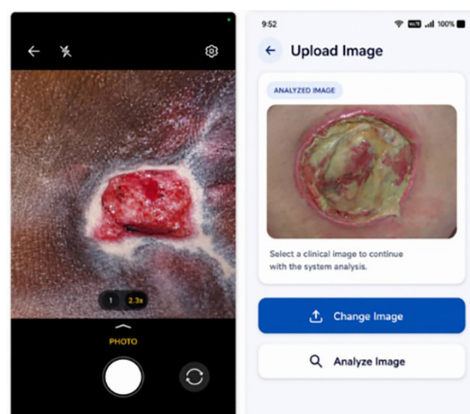


Fig. 10. Clinical image capture and selection interface

Processing and inference. After the clinical image is sent, the application automatically proceeds to the processing and inference stage, as shown in Figure 11. This module is a key component of UlcerScan, because it is the stage where the interaction between the application and the CNN-based AI model occurs. The first interface evidences a standby state, informing the user that the image has been transferred to the intelligent analysis system and is being processed. After the analysis is complete, it sends you to the results interface, which can be displayed on the second interface. This screen shows the user the uploaded clinical image, the estimated degree of PU determined by the CNN model, and the confidence level of the result obtained. The app also generates a brief description of the analysis, which provides the user with a clear interpretation of the ranking result. For example, this test shows that the submitted image obtained Grade III with a confidence level of 100%. On the other hand, it shows you a button so that you can start a new analysis or consult the history of previously saved results.

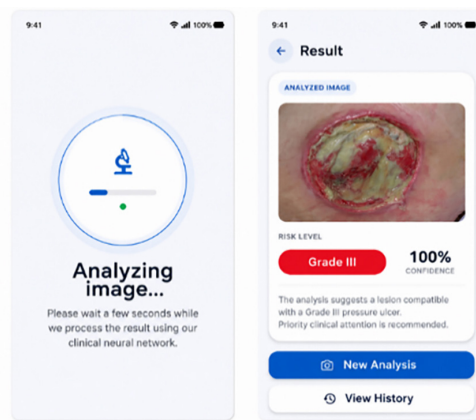


Fig. 11. Processing interfaces and result of the analysis of the CNN model

Analysis history. This implemented section is a test history module, shown in Figure 12. This screen is designed to store and view previous results made within the app. Each saved result includes the clinical image, the degree of PU determined by the app, the confidence level of the result, as well as the date and time of the test performed. In this way, the user can follow up on the tests carried out. In addition, the user has the option to delete individual records or completely clear the test history.

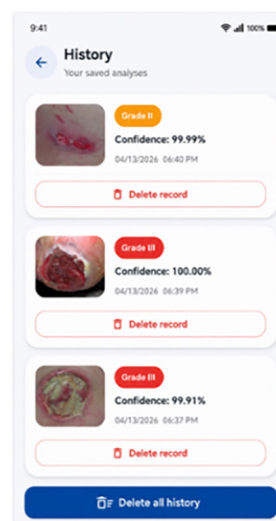


Fig. 12. Results history interface

5 DISCUSSION

The findings obtained from the study demonstrate that the use of CNN is an effective approach to automatically classify PU from clinical images. In this study, the YOLOv11-CLS model demonstrated the highest accuracy, reaching 97.1%, ensuring its ability to correctly distinguish between different grades of pressure ulcers.

On the other hand, the comparison with previous studies shows that the results obtained are competitive and, in some cases, exceed those presented in the literature. For example, in the study [17], the Inception-ResNet-v2 model demonstrated high accuracy and F1 scores when analyzing necrotic tissue; however, classification was performed with a smaller set of images and focused primarily on supporting clinical decision-making. In addition, the present study implemented a fully functional mobile environment for automated image analysis and user visualization of results. Compared to the study [18], which used image segmentation models and achieved a maximum IoU of 75.6%, the proposed CNN model demonstrated greater stability in the classification of PU grades. This may be due to the process of prior standardization of images and separation of data by levels of PU severity prior to training.

The detailed evaluation of YOLOv11-CLS managed to demonstrate in the results a good balanced performance in the different metrics. Also, in the detection of PU, it obtained an accuracy of 98.3%, indicating that most of the images identified as positive were correctly classified, while the recall of 96.1% shows that the model detected a high proportion of real cases of PU. In addition, the F1 score of 97.2% confirms an appropriate balance between accuracy and recall.

On the other hand, for the multiclass severity classification, the model managed to achieve an accuracy of 97.6%, recall of 95.7%, an F1 score of 96.9%, and an accuracy of 97.9%. Although less recall was obtained in the multiclass classification it may be due to the fact that the images in some of the cases were visually similar between adjacent grades of PU, particularly when tissue depth, necrosis, or wound boundaries are difficult to distinguish from a single clinical image.

In addition, the results of this study achieved higher reported classification accuracy than those obtained in the study [19], where the average classification accuracy was 60.95%. Despite using the ResNet50 architecture and a large number of images, the authors observed instability in the results between the different classes of diseases. In the present study, we achieved a more consistent classification among PU categories, which positively impacted the overall accuracy of the model. Similarly, compared to the study [20], where the classification accuracy was 75%, the developed model demonstrated a higher performance in recognizing PU grades. These results indicate that the choice of CNN architecture and image preprocessing play a key role in improving the quality of classification.

Finally, the results are consistent with the findings of [21], which also highlights the high effectiveness of CNNs in the analysis of skin diseases. Despite the differences in model architectures, both studies confirm that deep learning has significant potential for application in image-based medical diagnosis. Overall, this study confirms that the integration of CNNs into a mobile app not only improves the accuracy of PU classification, but also facilitates a faster and more convenient preliminary skin lesion analysis process in clinical practice.

Overall, the findings obtained in the present study suggest that the integration of CNN-based models into mobile applications can facilitate accessible image analysis, visualization of results, and preliminary clinical support. Likewise, previous studies mentioned that intelligent mobile healthcare systems have also highlighted the value of combining AI, mobile computing, and data analytics to improve accessibility

and delivery of health services [22, 23]. However, the proposed application should be considered as a complementary tool, not a substitute for professional clinical assessment.

Finally, the present study has several limitations: first, the dataset used was based only on images obtained in only three hospitals in Lima, so the generalizability of the results to other clinical institutions and patient groups is limited. It should be mentioned that no external and prospective validation of the system was carried out with independent data. Likewise, the quality of the classification can also be affected by different conditions, whether they are lighting, camera angle and distance, skin tone or visual similarity of PU with other types of lesions. On the other hand, the system was not evaluated for its clinical sensitivity, safety, or impact on medical decision-making. Therefore, future studies should perform external clinical validation, including a wider range of differential imaging and diagnostics, evaluate usability and safety, and implement local model processing on a mobile device. The app is designed solely as a clinical support tool and should not replace diagnostic, staging, or treatment decisions made by qualified healthcare professionals.

6 CONCLUSIONS

The study achieved its main objective, as a mobile application was developed for the automated classification of PU using CNNs. The results confirmed that the use of this model allows for effective clinical image analysis and accurate determination of the degree of PU, facilitating a faster preliminary assessment of skin lesions. In addition, the developed mobile application demonstrated the feasibility of integrating AI into an intuitive and accessible system for visual image analysis. The features implemented for photo capture, image upload, data processing, and analysis history storage created a more intuitive and convenient user experience.

A comparison of various CNN architectures revealed that the YOLOv11-CLS model obtained the best classification performance among the evaluated options, obtaining a higher overall accuracy of 97.1%, offering the most consistent results in the recognition of the different degrees of PU. It also demonstrated a solid performance related to the detection of PU, with an F1 score of 97.2% and an accuracy of 98.1%, and in the classification of multiclass severity, with an F1 score of 96.9% and an accuracy of 97.9%. This confirms the importance of correctly selecting the model architecture and pre-standardizing the images before training. Despite the positive results, the study had some limitations. Specifically, the analysis was performed with a limited set of clinical images, and the system focuses primarily on classifying PU grades without extensive clinical context of the patient. Therefore, future research could focus on using broader and more diverse medical datasets, as well as integrating additional parameters such as patient age, tissue type, clinical history, and healing dynamics.

Likewise, a promising line of research is the implementation of local inference directly on a mobile device without the need for external services, which would improve the processing speed and availability of the system in environments with limited internet connectivity. We also recommend exploring the use of other deep learning architectures and image segmentation methods to improve the accuracy of recognition of complex UP cases. In summary, the results of the study confirm that the use of deep learning technologies in combination with mobile platforms represents a promising approach to support preliminary clinical assessment and digitalization of medical processes.

7 ACKNOWLEDGMENT

The authors of this study thank the Director of Research and the Image Processing Laboratory (INTI-Lab) of the University of Sciences and Humanities (UCH) for the support they provided throughout the study process.

8 REFERENCES

- [1] B. C. Velozo, M. V. Hong, L. C. Bernardo, M. C. N. E. Castro, J. Contreras-Ruiz, and L. P. F. Abbade, "Pressure injury: update on general concepts, clinical aspects, and laboratory findings – Part I," *An Bras Dermatol*, vol. 100, p. 501187, 2025. <https://doi.org/10.1016/j.abd.2025.501187>
- [2] X. Lan *et al.*, "Global, regional, and national burden of pressure ulcers from 1990 to 2021 and projections over the next decade: Results from the 2021 GBD study," *Wound Repair Regen*, vol. 33, no. 4, p. e70064, 2025. <https://doi.org/10.1111/wrr.70064>
- [3] Z. Li, F. Lin, L. Thalib, and W. Chaboyer, "Global prevalence and incidence of pressure injuries in hospitalised adult patients: A systematic review and meta-analysis," *Int. J Nurs. Stud.*, vol. 105, p. 103546, 2020. <https://doi.org/10.1016/j.ijnurstu.2020.103546>
- [4] P. Isfahani *et al.*, "Prevalence of hospital-acquired pressure injuries in intensive care units of the eastern mediterranean region: A systematic review and meta-analysis," *Patient Saf Surg*, vol. 18, Art. no. 1, 2024. <https://doi.org/10.1186/s13037-023-00383-8>
- [5] C. Siotos *et al.*, "Burden of pressure injuries: Findings from the global burden of disease study," *Eplasty*, vol. 22, p. e19, 2022.
- [6] W. V. Padula and B. A. Delarmente, "The national cost of hospital-acquired pressure injuries in the United States," *Int Wound J*, vol. 16, no. 3, pp. 634–640, 2019. <https://doi.org/10.1111/iwj.13071>
- [7] T. V. Boyko, M. T. Longaker, and G. P. Yang, "Review of the current management of pressure ulcers," *Adv. Wound Care (New Rochelle)*, vol. 7, pp. 57–67, 2018. <https://doi.org/10.1089/wound.2016.0697>
- [8] M. L. Chung *et al.*, "Risk factors for pressure injuries in adult patients: A narrative synthesis," *Int J Environ Res Public Health*, vol. 19, p. 761, 2022. <https://doi.org/10.3390/ijerph19020761>
- [9] P. M. G. Sardo, J. P. F. Teixeira, A. M. S. F. Machado, B. F. Oliveira, and I. M. Alves, "A systematic review of prevalence and incidence of pressure ulcers/injuries in hospital emergency services," *J Tissue Viability*, vol. 32, pp. 179–187, 2023. <https://doi.org/10.1016/j.jtv.2023.02.001>
- [10] T. Samah and M. Samar, "Investigating the key trends in applying artificial intelligence to health technologies: A scoping review," *PLoS One*, vol. 20, no. 5, p. e0322197, 2025. <https://doi.org/10.1371/journal.pone.0322197>
- [11] A. M. G. Alhejaily, "Artificial intelligence in healthcare (Review)," *Biomed Rep*, vol. 22, no. 1, p. 11, 2024. <https://doi.org/10.3892/br.2024.1889>
- [12] G. Joseph, N. Bhatti, R. Mittal, and A. Bhatti, "Current application and future prospects of artificial intelligence in healthcare and medical education: A review of literature," *Cureus*, vol. 17, no. 1, p. e77313, 2025. <https://doi.org/10.7759/cureus.77313>
- [13] C. Chen, N. A. Mat Isa, and X. Liu, "A review of convolutional neural network based methods for medical image classification," *Comput. Biol. Med.*, vol. 185, p. 109507, 2025. <https://doi.org/10.1016/j.compbiomed.2024.109507>
- [14] I. D. Mienye, T. G. Swart, G. Obaido, M. Jordan, and P. Ilono, "Deep convolutional neural networks in medical image analysis: A review," *Information*, vol. 16, p. 195, 2025. <https://doi.org/10.3390/info16030195>

- [15] C. Lei, Y. Jiang, K. Xu, S. Liu, H. Cao, and C. Wang, "Convolutional neural network models for visual classification of pressure ulcer stages: Cross-sectional study," *JMIR Med Inform*, vol. 13, p. e62774, 2025. <https://doi.org/10.2196/62774>
- [16] B. Aldughayfiq, F. Ashfaq, N. Z. Jhanjhi, and M. Humayun, "YOLO-based deep learning model for pressure ulcer detection and classification," *Healthcare (Basel)*, vol. 11, no. 9, p. 1222, 2023. <https://doi.org/10.3390/healthcare11091222>
- [17] T. J. Liu *et al.*, "A pressure ulcers assessment system for diagnosis and decision making using convolutional neural networks," *J Formos. Med. Assoc.*, vol. 121, no. 11, pp. 2227–2236, 2022. <https://doi.org/10.1016/j.jfma.2022.04.010>
- [18] P. Aguilar, "Segmentation of pressure ulcer images using deep learning," Final Degree Project, University of Malaga, Malaga, Spain, 2022. <https://hdl.handle.net/10630/25143>
- [19] E. R. Santis, D. F. Sierra, and J. A. Vergara, "Intelligent system for the detection of skin diseases using convolutional neural networks," Academic Work, Universidad del Norte, Barranquilla, Colombia, 2025. https://manglar.uninorte.edu.co/bitstream/handle/10584/13390/Paper_Proyecto_Final.pdf?sequence=1&isAllowed=y
- [20] A. M. Yuen, D. Jauregui, and D. Subauste, "Evaluation of pressure ulcers using machine learning techniques," in *Memoirs of the Fifteenth Ibero-American Conference on Complexity, Informatics and Cybernetics (CICIC 2025)*, 2025. <https://www.iiis.org/CDs2025/CD2025Spring/papers/CB807RW.pdf>
- [21] O. Iparraguirre-Villanueva and M. Cabanillas-Carbonell, "Application of convolutional neural networks in skin disease prediction: Accuracy and efficiency in dermatological image analysis," *Int J Online Biomed Eng*, vol. 21, no. 2, pp. 18–37, 2025. <https://doi.org/10.3991/ijoe.v21i02.52871>
- [22] R. Jeya, H. A. Karim, and S. B. Mansor, "Artificial intelligence and mobile apps support intelligent healthcare systems for mental health services," *International Journal of Interactive Mobile Technologies (ijIM)*, vol. 18, no. 20, pp. 157–168, 2024. <https://doi.org/10.3991/ijim.v18i20.50743>
- [23] B. Muruganantham, H. A. Karim, and S. B. Mansor, "Mobile health service adoption: Reducing the digital gap and age's moderating influence," *International Journal of Interactive Mobile Technologies (ijIM)*, vol. 18, no. 17, pp. 139–153, 2024. <https://doi.org/10.3991/ijim.v18i17.50741>

9 AUTHORS

Ana Huamani-Huaracca has a degree in Nursing from the Image Processing Research Laboratory (INTI-Lab), University of Sciences and Humanities, Lima, Peru. She also holds a Master's degree in Health Services Management and a Master's degree in Public Health. She possesses a strong academic background and a comprehensive approach to the field of health. Her areas of interest include scientific and experimental research in public health with clinical applications, health promotion, and epidemiology. She is also deeply interested in the management of health systems and services, focusing on the implementation of technology-based strategies to optimize the quality of health services (E-mail: ahuamani@uch.edu.pe).

Brigitte Torres-Pinedo is a Systems and Computer Engineering from the Image Processing Research Laboratory (INTI-Lab), University of Sciences and Humanities, with an interest in research-oriented technological development. She possesses programming skills in languages such as Java, Python, and PHP, as well as basic experience with frameworks such as ASP.NET. She is proficient in database management using SQL, with a special emphasis on the efficient organization and administration of information. She also has a particular interest in research areas related to the

use of emerging technologies for solving health problems (E-mail: britorrespin@uch.edu.pe).

Juan Morales is with the E-Health Research Center, Universidad de Ciencias y Humanidades, Lima, Peru. He is a physician and specialist in Family and Community Medicine. He holds a PhD in Medicine, a Master's degree in Project and Social Program Management, and a Master's degree in Digital Health. He has experience in clinical practice, medical education, and public health research (E-mail: jmoralesq@uch.edu.pe).

Sebastian Ramos-Cosi is with the Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, Lima, Peru. He is an Electronic Engineer specializing in Bioengineering, Artificial Intelligence, and emerging technologies. He has extensive experience in writing and publishing articles, and is certified in intellectual property. He has experience in the field of medical technologies, having developed monitoring and diagnostic devices that have a positive impact on the healthcare sector. Furthermore, in the field of artificial intelligence and data analysis, he has applied advanced techniques in predictive modeling and process optimization, implementing these techniques to improve expert systems (E-mail: sramos@uch.edu.pe).