TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 24, No. 6, December 2026, pp. 1928~1941
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v24i6.28031

Web-based pressure ulcer stage classification using convolutional neural network and transfer learning

Suci Imani Putri[1], Ananda Prastuti Sutrisno[2]
[1]Electromedical Engineering Study Program, Akademi Teknik Elektromedik Andakara, Jakarta, Indonesia
[2]Nursing Study Program, Akademi Keperawatan Andakara, Jakarta, Indonesia

Article Info

Article history:

Received May 11, 2026
Revised Aug 12, 2026
Accepted Aug 28, 2026

Keywords:

Clinical decision support
Convolutional neural network
Medical image classification
Pressure ulcer staging
Transfer learning
Web-based application

ABSTRACT

Accurate pressure ulcer staging is essential for wound assessment and treatment planning; however, visual evaluation remains subjective and prone to inter-observer variability. This study presents a web-based pressure ulcer stage classification system using convolutional neural networks (CNNs) and transfer learning to support automated clinical decision-making. A publicly available dataset containing 1,091 pressure ulcer images was categorized into four clinical stages and divided into training and validation sets. Comparative experiments were performed using a baseline CNN, a proposed CNN, MobileNetV2, and ResNet50. The proposed framework incorporated image preprocessing, data augmentation, batch normalization, dropout regularization, and transfer learning to improve classification performance and model stability. Experimental results showed that MobileNetV2 achieved the best performance, with an accuracy of 76.50%, precision of 0.77, recall of 0.76, and an F1-score of 0.76, while the proposed CNN outperformed the baseline CNN, achieving 69.12% accuracy. Confusion matrix analysis indicated that most misclassifications occurred between adjacent pressure ulcer stages because of similar visual characteristics. Preliminary expert validation showed agreement between model predictions and expert annotations in three of four representative cases. These findings demonstrate the potential of the proposed lightweight web-based framework to support preliminary pressure ulcer assessment and real-time clinical decision-support applications.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Ananda Prastuti Sutrisno
Nursing Study Program, Akademi Keperawatan Andakara
Jakarta, Indonesia
Email: anandaprastutis@gmail.com

1. INTRODUCTION

Pressure ulcers, also referred to as bedsores or pressure injuries, are localized damage to the skin and underlying soft tissue, usually occurring over bony prominences as a result of prolonged pressure or pressure combined with shear [1]. These wounds are commonly found in patients with limited mobility, older adults, and individuals requiring long-term hospitalization or bed rest. If not properly assessed and managed, pressure ulcers may lead to infection, tissue necrosis, prolonged hospitalization, and reduced quality of life [2], [3]. Therefore, early and accurate assessment of pressure ulcer severity is essential to support appropriate wound management, improve clinical decision-making, and reduce the risk of serious complications.

Pressure ulcer staging is generally performed through visual assessment by healthcare professionals based on wound appearance, tissue involvement, color changes, and the depth of tissue damage. The classification usually ranges from stage 1 to stage 4, with higher stages indicating more severe tissue

involvement [1], [4]. However, manual wound assessment remains subjective and highly dependent on clinical expertise. Variability in interpretation among clinicians, particularly when distinguishing adjacent stages with similar visual characteristics, may lead to inconsistent diagnoses and inappropriate treatment planning. Consequently, automated pressure ulcer classification systems are needed to provide more objective, consistent, and reliable wound assessment.

Recent advances in artificial intelligence have significantly improved medical image analysis. In particular, convolutional neural networks (CNNs) have demonstrated remarkable capability in automatically extracting discriminative image features, making them highly effective for various healthcare applications, including blood cell classification, radiological image interpretation, and wound analysis [5]–[7]. Numerous studies have applied CNN-based approaches to pressure ulcer assessment. A mobile-assisted pressure ulcer categorization system using Faster region-based CNN (Faster R-CNN) was developed to support community healthcare practice [8]. A YOLOv5-based framework was subsequently proposed for real-time pressure ulcer detection and classification [9]. More recently, a vision transformer (ViT) model was implemented within an internet of things (IoT)-enabled healthcare environment [10]. At the same time, several CNN architectures, including AlexNet, VGG16, ResNet18, and DenseNet121, were comparatively evaluated to determine their effectiveness in pressure ulcer classification [11]. Furthermore, a multimodal deep learning framework integrating wound images with clinical information was introduced to improve the fine-grained classification of pressure ulcers and incontinence-associated dermatitis [12]. Collectively, these studies demonstrate the rapid development of deep learning techniques for pressure ulcer assessment while highlighting continuous efforts to improve classification accuracy and clinical applicability.

Besides conventional CNN architectures, transfer learning has become an increasingly popular strategy for medical image classification because pretrained models can learn generalized visual representations from large-scale datasets [13]. Among the most widely adopted transfer learning models, MobileNetV2 offers computational efficiency through lightweight depth wise separable convolutions, making it suitable for resource-constrained and real-time applications [14]. In contrast, ResNet50 employs residual learning mechanisms that facilitate the training of deeper neural networks while alleviating vanishing gradient problems, resulting in robust classification performance [15], [16].

Despite these advances, several limitations remain in existing deep learning-based pressure ulcer classification systems. First, classification accuracy is still affected by the high visual similarity between adjacent pressure ulcer stages, resulting in frequent misclassification. Second, many studies rely on relatively limited or institution-specific datasets, which may reduce model generalizability across diverse clinical settings. Third, although recent studies have demonstrated promising classification performance, most primarily focus on offline model evaluation with limited attention to lightweight web-based deployment for real-time clinical decision support. These limitations highlight the need for an accessible classification framework that combines comparative evaluation of deep learning models with practical web-based implementation to support automated pressure ulcer stage assessment. Recent cloud-based healthcare systems have also emphasized the importance of scalable and efficient web-based deployment for real-time medical applications [17]. Web-based healthcare information systems have shown potential to improve accessibility, usability, and practical clinical support in medical environments [18].

To improve classification robustness, several optimization techniques have been widely adopted in deep learning. Data augmentation increases dataset diversity and helps reduce overfitting, particularly when training datasets are limited [19]. Batch normalization stabilizes network training by normalizing intermediate feature distributions, whereas dropout improves model generalization by randomly deactivating neurons during training [19]–[21]. These techniques have been shown to improve the robustness and reliability of CNN-based medical image classification systems.

Based on the identified research gap, this study proposes an automated web-based pressure ulcer stage classification system using CNN and transfer learning approaches. Unlike previous studies that primarily focused on model development or offline image classification, the proposed framework integrates transfer learning within a lightweight web-based clinical decision-support application, enabling real-time pressure ulcer stage prediction through an accessible web interface. The proposed framework further incorporates image preprocessing, data augmentation, batch normalization, dropout, and an optimized learning rate to improve classification performance and model stability. Comparative experiments were conducted using a baseline CNN, an enhanced CNN architecture, MobileNetV2, and ResNet50 to evaluate classification accuracy and computational efficiency. Furthermore, the best-performing model was deployed as a lightweight web-based application that enables real-time pressure ulcer stage prediction from uploaded wound images. Preliminary expert validation was also conducted to assess the clinical relevance of the proposed system. The main contributions of this study are threefold: i) a comparative evaluation of CNN and transfer learning architectures for pressure ulcer stage classification; ii) the development of a lightweight web-based classification framework for real-time deployment; and iii) the implementation and preliminary clinical validation of an accessible decision-support prototype for pressure ulcer assessment.

2. METHOD

2.1. Research design

This study adopted a deep learning-based approach using a CNN to classify pressure ulcer images into four stages: stage 1, stage 2, stage 3, and stage 4. The research procedure consisted of several main stages, including data acquisition, preprocessing, model design, training, and evaluation. This workflow was designed to enable the model to learn relevant visual features from medical images and produce stage classification results. The overall research procedure is illustrated in Figure 1.

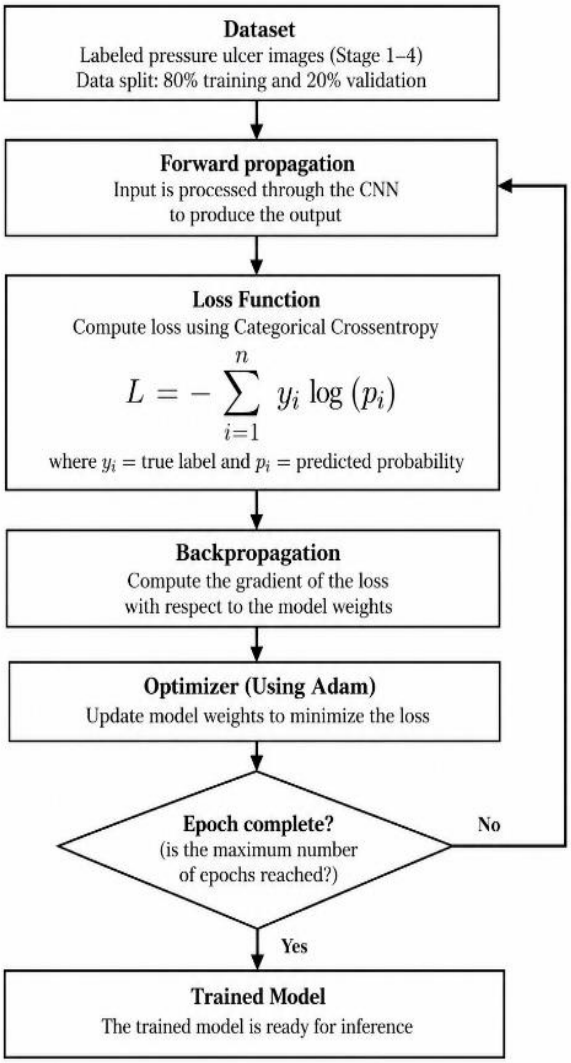


Figure 1. Research procedure

2.2. Dataset and data acquisition

The dataset used in this study was obtained from the publicly available Pressure Ulcers Stages dataset on Kaggle [22]. It consisted of 1,091 pressure ulcer images categorized into four clinical stages: stage 1, stage 2, stage 3, and stage 4. The class distribution was relatively balanced across all categories, although a slight imbalance was observed in stage 2. Because this imbalance was considered minimal, no additional balancing techniques, such as oversampling, under sampling, or class weighting, were applied during training. The dataset was divided into 80% training data and 20% validation data to support model learning and performance evaluation. To improve data diversity and enhance model generalization, data augmentation techniques, including rotation, zooming, horizontal flipping, and brightness adjustment, were applied during training. Model performance was subsequently evaluated using confusion matrices, along with class-wise precision, recall, and F1-score to provide a comprehensive assessment of classification performance across

all pressure ulcer stages. The detailed distribution of images across the four pressure ulcer stages, including the allocation of training and validation samples, is presented in Table 1.

Table 1. Dataset distribution

Class	Total	Training	Validation
1	230	184	46
2	313	250	63
3	275	220	55
4	273	218	55
Total	1091	872	219

2.3. Data preprocessing

Data preprocessing was performed to improve image quality and ensure compatibility with the CNN model. The preprocessing steps included resizing all images to 224×224 pixels, converting images into numerical arrays, normalizing pixel values to a range of 0–1, and adjusting the input dimensions. In addition, data augmentation techniques, including rotation, zooming, shifting, and horizontal flipping, were applied to increase data variability and reduce overfitting. The preprocessing and feature extraction process is illustrated in Figure 2.

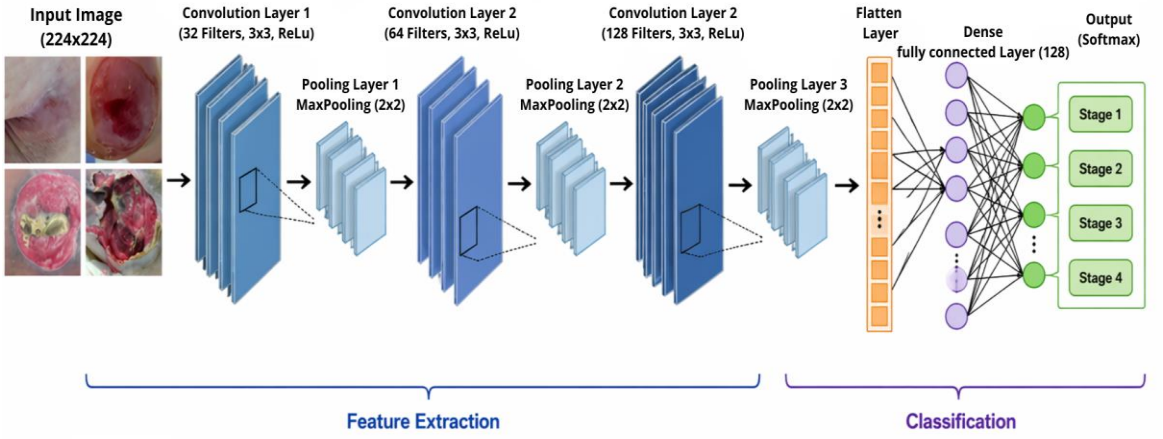


Figure 2. The preprocessing and feature extraction process

2.4. CNN model architecture

The CNN architecture used in this study consisted of multiple convolutional layers followed by batch normalization and max-pooling layers for feature extraction [23], [24]. The extracted feature maps were then flattened and processed through fully connected layers. A dropout layer with a rate of 0.4 was applied to reduce overfitting during training. The final output layer used a SoftMax activation function to generate class probabilities for the four pressure ulcer stages.

2.5. Transfer learning models

In addition to the proposed CNN model, this study evaluated two transfer learning architectures, namely MobileNetV2 and ResNet50. Transfer learning uses pretrained models to extract general visual features from large-scale datasets and adapt them to new image classification tasks with limited data [13]. MobileNetV2 was selected. It is a lightweight architecture suitable for efficient and real-time applications, whereas ResNet50 was used because its residual connections support deeper feature learning [14]-[16].

Both models were initialized using pretrained ImageNet weights. The original classification layers were replaced with new layers for four-class pressure ulcer stage classification, consisting of stage 1, stage 2, stage 3, and stage 4. The input image size was set to 224×224 pixels, and the final layer used SoftMax activation to generate class probabilities [25]. The models were trained using the same 80% training and 20% validation split, with the Adam optimizer and categorical cross-entropy loss function [26]. Their performance was compared with the baseline CNN and the proposed CNN using accuracy, precision, recall, and F1-score.

2.6. Training procedure

The training process was performed iteratively to optimize model performance. Each training iteration consisted of forward propagation, loss calculation, backpropagation, and weight updating. During training, each model was trained independently using the same training and validation datasets, preprocessing pipeline, data augmentation strategy, optimizer configuration, and evaluation protocol to ensure a fair comparison among the evaluated architectures. The loss function used in this study was categorical cross-entropy, defined as:

$$L = -\sum_{i=1}^{c} y_i log(p_i) \quad (1)$$

where c represents the number of classes, y_i denotes the true label for class i, and p_i denotes the predicted probability for class i. Model performance on the validation dataset was monitored after each epoch, and the final evaluation was performed using the trained model that achieved the best validation performance.

2.7. Training parameters

The training parameters were selected to balance model performance, training stability, and computational efficiency. The input image size was set to 224×224 pixels to match the default input requirements of MobileNetV2 and ResNet50 architectures while preserving sufficient visual detail for pressure ulcer classification. A batch size of 8 was selected to accommodate hardware memory constraints and maintain stable gradient updates during training.

The hyperparameters used in this study, including the learning rate, batch size, dropout rate, and number of training epochs, were determined through preliminary empirical experiments. Several parameter combinations were evaluated based on validation accuracy and training stability. The final configuration (learning rate = 0.00005, batch size = 8, dropout = 0.4, and 15 epochs) provided the best balance between classification performance, convergence stability, and computational efficiency.

The Adam optimizer was adopted because of its effectiveness in deep learning optimization tasks [25]. The selected hyperparameter configuration ensured stable convergence during training, particularly for transfer learning models initialized with pretrained weights. Categorical cross-entropy was used as the loss function because the task involved multi-class classification [25]. The dataset was divided into 80% training data and 20% validation data to evaluate the model's generalization capability on unseen images. The proposed models were implemented using Python (version 3.12.10) with the TensorFlow–Keras deep learning framework. At the same time, while the web-based prototype was developed using the Flask framework to support real-time pressure ulcer stage classification. To improve experimental reproducibility, fixed random seeds were applied using NumPy and TensorFlow during model initialization and data processing.

2.8. Model inference and prediction

After the training process was completed, the developed CNN model was used to perform inference on unseen pressure ulcer images. During prediction, each input image was processed by the trained network to generate probability values for each pressure ulcer stage. The inference process aimed to identify the most probable class based on the visual features learned during training. The model output was represented as a probability vector:

$$p = [p_1, p_2, \mathrm{p}_3, \mathrm{p}_4] \quad (2)$$

where p_i represents the probability of the image belonging to the i-th pressure ulcer stage. The probability values satisfy the following condition:

$$\sum_{i=1}^{4} p_i = 1 \quad (3)$$

The probability distribution was generated using the SoftMax activation function at the output layer [27]. The SoftMax function converts the linear output values into normalized probabilities ranging from 0 to 1:

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^{4} e^{z_j}} \quad (4)$$

where z_i denotes the output value of the neuron i before activation, and p_i represents the probability value of class i. The SoftMax function ensures that all output probabilities sum to one, allowing the model to

determine the most likely pressure ulcer stage. The predicted class was determined using the argmax operation, which selects the class with the highest probability value:

$$\hat{y} = arg\ max(p_i) \quad (5)$$

For example, if the model produces an output probability vector of [0.10, 0.20, 0.60, 0.10], the image is classified as stage 3 because the third probability value is the highest among all classes. In addition to the predicted class label, the system also provides a confidence score to indicate the certainty level of the prediction. The confidence score is calculated from the maximum probability value obtained from the SoftMax output:

$$Confidence = \max(p_1) \times 100\% \quad (6)$$

Higher confidence values indicate stronger model confidence in the predicted pressure ulcer stage. The confidence score was displayed together with the predicted label in the developed web-based system, enabling users to interpret the prediction results more easily. The developed system produced one of four classification outputs, namely stage 1, stage 2, stage 3, or stage 4. The prediction results were generated automatically after the uploaded image passed through the preprocessing and inference stages, making the proposed system suitable for web-based pressure ulcer image classification.

2.9. Model evaluation metrics

The performance of the developed CNN model was evaluated using the validation dataset, which consisted of 20% of the total dataset and was not used during the training process. The evaluation aimed to measure the model's ability in generalize when classifying unseen pressure ulcer images. Several evaluation metrics were used in this study, including accuracy, precision, recall, and F1-score [28]. Accuracy represents the proportion of correctly classified samples relative to the total number of evaluated samples and is defined as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

where TP denotes true positive, TN denotes true negative, FP denotes false positive, and FN denotes false negative predictions. Higher accuracy values indicate better overall classification performance.

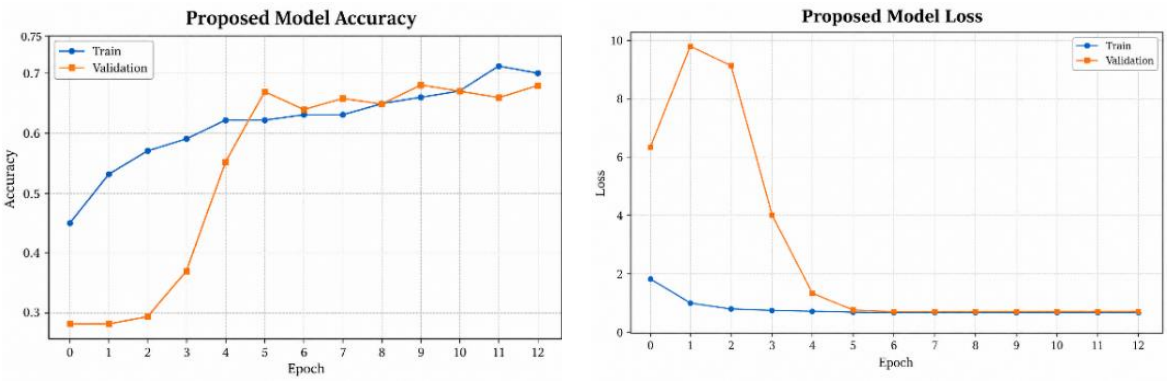
In addition to accuracy, precision, recall, and F1-score were calculated to provide a more comprehensive evaluation of model performance across all pressure ulcer stages. Precision measures the proportion of correctly predicted positive samples, recall measures the ability of the model to identify relevant samples, and the F1-score represents the harmonic mean of precision and recall. Furthermore, confusion matrix analysis was performed to evaluate classification performance for each pressure ulcer stage and to identify potential misclassification patterns between visually similar classes.

3. RESULTS AND DISCUSSION

3.1. Training performance

The training performance of the proposed CNN model was evaluated using accuracy and loss values during the training process. The results are presented in Figure 3, which show the changes in accuracy and loss across training epochs. As shown in Figure 3(a), the accuracy curve indicates a gradual improvement during training. At the initial epochs, the validation accuracy was relatively low, suggesting that the model had not yet learned discriminative visual features from the pressure ulcer images. As training progressed, the validation accuracy increased and stabilized at approximately 0.67–0.69. This trend indicates that the model improved its classification capability on unseen validation data. The training accuracy also followed a similar pattern and slightly exceeded the validation accuracy, which is commonly observed in deep learning models. The relatively small gap between training and validation accuracy suggests that the proposed model did not experience severe overfitting.

Figure 3(b) shows the training and validation loss curves. The loss decreased rapidly during the early epochs and gradually stabilized in the later epochs, indicating that the model converged during training. The validation loss trend was comparable to the training loss, further suggesting stable learning behavior and no clear indication of divergence. Overall, the training results show that the proposed CNN model learned relevant visual features from the dataset and achieved stable convergence. The use of batch normalization, dropout, and a low learning rate may have contributed to improving training stability and reducing overfitting.



(a) (b)

Figure 3. Training and validation performance of the proposed CNN model: (a) accuracy and (b) loss

3.2. Confusion matrix analysis

The confusion matrix was used to examine the classification performance of the proposed CNN model in more detail by comparing the predicted labels with the true labels. The confusion matrix results are presented in Figure 4. As shown in the Figure 4, the proposed model achieved correct classifications across all four pressure ulcer stages. For stage 1, the model produced a relatively high number of correct predictions, indicating that early-stage pressure ulcers had more distinguishable visual characteristics. Stage 2 also showed acceptable classification performance, although several samples were misclassified into neighboring classes. In contrast, stage 3 appeared to be the most challenging class, as indicated by the lower number of correct predictions and the higher number of misclassifications.

A considerable proportion of stage 3 samples was incorrectly classified as stage 4, suggesting that these two stages may share similar visual features, particularly in terms of wound depth and tissue damage. Similarly, some stage 4 samples were misclassified as stage 3, further indicating that distinguishing advanced stages remains difficult due to subtle differences in visual patterns. Overall, the main source of classification error occurred between adjacent stages, particularly between stage 2 and stage 3, as well as between stage 3 and stage 4. This pattern is reasonable in medical image classification tasks, where class boundaries are not always visually distinct, and feature patterns may overlap. Despite these challenges, the confusion matrix results indicate that the proposed CNN model was able to capture relevant features for pressure ulcer stage classification. Further improvements, such as increasing dataset diversity, using higher-quality clinical images, or applying more advanced model architectures, may help reduce misclassification between visually similar stages.

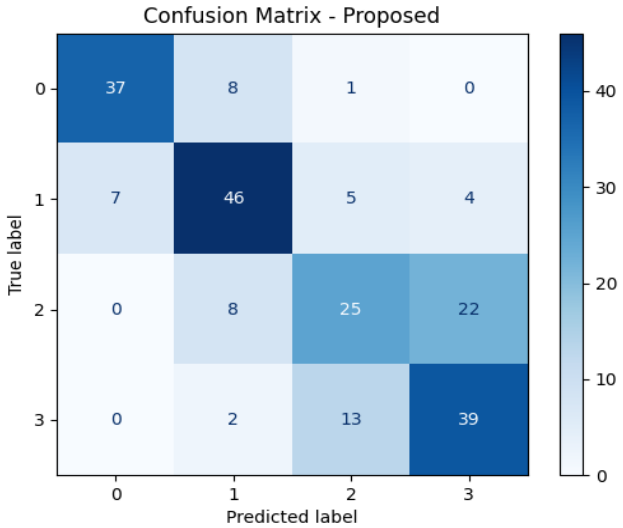


Figure 4. Confusion matrix of the proposed CNN model

3.3. Model evaluation

The performance of the proposed CNN model was evaluated using accuracy, precision, recall, and F1-score to provide a comprehensive assessment of its classification capability across the four pressure ulcer stages. The proposed CNN achieved an overall accuracy of 69.12%, while the corresponding precision, recall, and F1-score were all approximately 0.69, indicating relatively balanced classification performance without a strong bias toward any particular pressure ulcer stage. To further evaluate the effectiveness of the proposed approach, comparative experiments were conducted using the baseline CNN, MobileNetV2, and ResNet50 architectures. The overall performance comparison of all evaluated models is summarized in Table 2.

Table 2. Model comparison

Model	Accuracy (%)	Precision	Recall	F1-score
Baseline CNN	62	0.63	0.62	0.61
Proposed CNN	69.12	0.69	0.69	0.69
MobileNetV2	76.50	0.77	0.76	0.76
ResNet50	44.24	0.38	0.44	0.38

As shown in Table 2, MobileNetV2 achieved the best overall performance among the evaluated models, demonstrating the effectiveness of transfer learning in extracting representative visual features from pressure ulcer images. The proposed CNN also showed a substantial improvement over the baseline CNN while maintaining a simpler architecture suitable for lightweight web-based implementation. In contrast, the lower performance of ResNet50 suggests that deeper architectures generally require larger training datasets to achieve better generalization. These findings indicate that transfer learning provides a practical advantage for pressure ulcer classification when only a limited number of annotated medical images are available. To provide a more detailed evaluation, the performance of the proposed CNN was further analyzed at the class level using precision, recall, and F1-score for each pressure ulcer stage, as presented in Table 3.

Table 3. Classification report per class

Class	Precision	Recall	F1-score	Support
Stage 1	0.77	0.78	0.77	46
Stage 2	0.68	0.66	0.67	63
Stage 3	0.68	0.58	0.63	55
Stage 4	0.65	0.76	0.70	55
Average	0.69	0.69	0.69	219

The class-level evaluation shows that the proposed CNN achieved the highest performance for stage 1, whereas stage 3 remained the most challenging class. This finding is consistent with the confusion matrix analysis, indicating that intermediate pressure ulcer stages exhibit greater visual similarity than early or advanced stages. Overall, the results demonstrate that the proposed CNN provides relatively balanced classification performance across all pressure ulcer stages and is suitable for preliminary automated pressure ulcer image classification.

3.3.1. Expert validation

To provide an additional assessment of the clinical relevance of the proposed classification system, a preliminary expert validation was conducted after the model evaluation phase. The validation was conducted by an anonymous certified wound care expert with extensive clinical experience in pressure ulcer assessment. The expert holds a master of nursing science degree and is both an academic nurse and a certified wound care practitioner who routinely manages patients with chronic wounds in clinical practice. Four representative pressure ulcer images were intentionally selected to provide a preliminary proof-of-concept clinical validation covering different pressure ulcer stages. These images were independent of the public Kaggle dataset and were not used during model training or validation.

The expert independently reviewed each image and assigned a pressure ulcer stage according to established clinical wound assessment criteria without access to the model predictions. The expert-assigned labels were subsequently compared with the corresponding model predictions and confidence scores generated by the proposed web-based classification system to evaluate the agreement between expert clinical judgment and the model predictions. This comparison also provided an initial assessment of the potential clinical applicability of the proposed system. The comparison results are summarized in Table 4.

The results showed that the model predictions agreed with the expert annotations in three of the four cases (75%). The only disagreement occurred for a stage 3 image, which the model classified as stage 4. This

finding is consistent with the confusion matrix analysis, indicating that most misclassifications occurred between adjacent pressure ulcer stages with similar visual characteristics. Although the expert validation involved only a limited number of representative samples and was intended as a preliminary proof-of-concept assessment, the findings provide initial evidence that the proposed system can generate clinically meaningful predictions. Future studies should include larger clinically annotated datasets, involve multiple wound care experts, and perform formal inter-rater agreement analyses to strengthen further the clinical validity and generalizability of the proposed system.

Table 4. Comparison between model prediction and expert validation

No	Sample	Expert label	Model prediction	Confidence (%)	Agreement
1	Sample 1	Stage 4	Stage 4	92.83	Yes
2	Sample 2	Stage 3	Stage 4	66.16	No
3	Sample 3	Stage 2	Stage 2	68.75	Yes
4	Sample 4	Stage 1	Stage 1	84.18	Yes

3.4. Web-based system implementation

To demonstrate the practical application of the proposed model, a web-based prototype was developed using the trained CNN model. The system allows users to upload pressure ulcer images and obtain classification results along with confidence scores. The web-based system was designed to evaluate the applicability of the proposed model in a real-world scenario by providing a simple and efficient classification interface. The prototype is available at: https://pressure-ulcer-predictor-production.up.railway.app/. Figure 5 presents the web-based prototype developed for pressure ulcer stage classification. The interface allows users to upload wound images and obtain predicted stage labels together with confidence scores generated by the trained model.

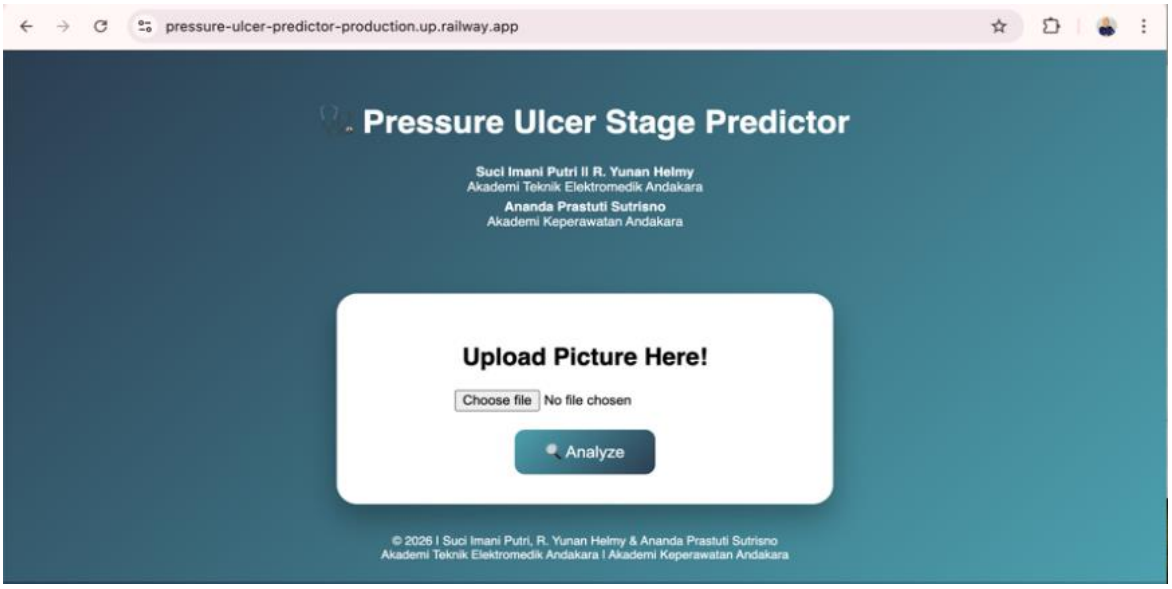


Figure 5. Web-based interface for pressure ulcer classification

The system processes input images using the same preprocessing steps applied during training and performs classification using the trained model. The output includes the predicted pressure ulcer stage and the corresponding confidence value, allowing users to interpret the results more easily. Figure 6 presents representative examples of classification results generated by the web-based system. Specifically, Figure 6(a) shows a stage 4 prediction with a confidence score of 92.83%, Figure 6(b) shows a stage 4 prediction with a confidence score of 66.16%, Figure 6(c) illustrates a stage 2 prediction with a confidence score of 68.75%, and Figure 6(d) shows a stage 1 prediction with a confidence score of 84.18%. These examples demonstrate how the system provides both stage predictions and confidence information for different pressure ulcer images.

The results indicate that the system can provide prediction outputs in real time and present them in a format that is easy to understand. The web-based prototype successfully displays the predicted stage and confidence score for each uploaded image. These findings show that the proposed system can be implemented as a preliminary decision-support tool for pressure ulcer stage classification. Nevertheless, the system should not be used as a replacement for clinical assessment, and further validation with larger and more diverse clinical datasets is necessary before use in clinical practice.

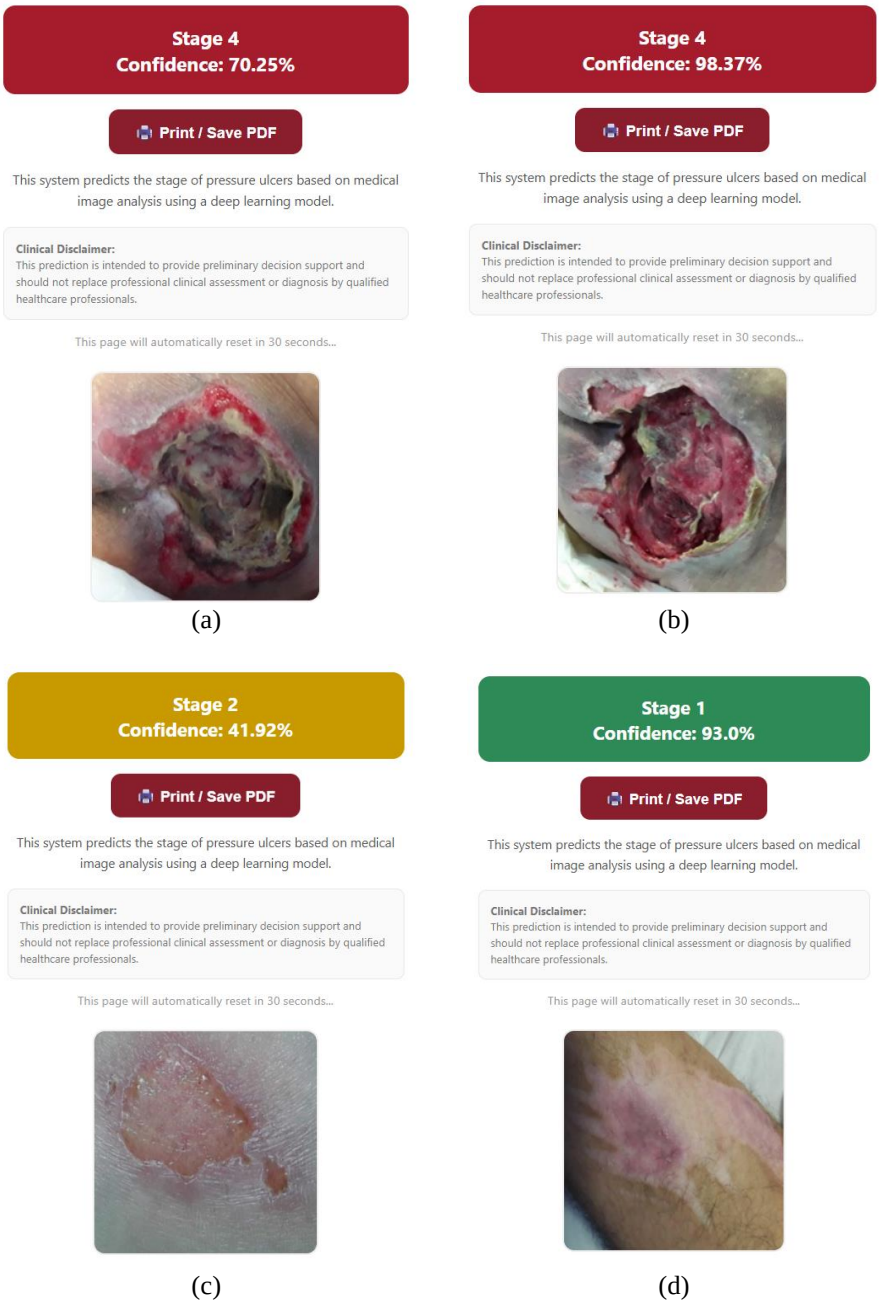


Figure 6. Representative classification results of the proposed model: (a) stage 4 (confidence: 92.8%), (b) stage 4 (confidence: 64.1%), (c) stage 2 (confidence: 67.5%), and (d) stage 1 (confidence: 84.1%)

3.5. Discussion

The experimental results demonstrate that the evaluated deep learning models were able to classify pressure ulcer stages with varying levels of performance. Among the evaluated models, MobileNetV2 consistently achieved the highest classification performance, with an accuracy of 76.50%, precision of 0.77,

recall of 0.76, and F1-score of 0.76. This finding indicates that transfer learning can effectively extract representative visual features from pressure ulcer images, particularly when only a limited number of training samples are available [13]. The superior performance of MobileNetV2 may be attributed to its lightweight architecture and the use of depth wise separable convolutions, which reduce the number of trainable parameters while maintaining effective feature extraction capability [14]. Consequently, the model is less prone to overfitting than deeper architectures when trained on relatively small datasets while still achieving competitive classification performance. These findings are consistent with previous studies reporting that transfer learning approaches improve pressure ulcer image classification by enhancing feature representation and classification consistency [11], [29]–[32].

The proposed CNN model achieved an accuracy of 69.12%, with precision, recall, and F1-score values of approximately 0.69, outperforming the baseline CNN, which achieved an accuracy of 62%. This improvement indicates that the proposed CNN architecture learned more representative visual features than the conventional baseline model. The consistent trends observed in the training and validation curves further suggest stable learning without clear evidence of severe overfitting. This improvement is likely associated with the combined use of data augmentation, batch normalization, dropout regularization, and a low learning rate, which enhanced model stability and improved generalization performance during training [20], [21], [26]. In contrast, although ResNet50 has demonstrated excellent performance in large-scale image recognition tasks, its deeper architecture contains substantially more trainable parameters and generally requires larger datasets to achieve optimal generalization. Consequently, the relatively limited dataset used in this study may have restricted its learning capability, resulting in lower classification performance than MobileNetV2.

As summarized in Table 5, recent studies have employed different deep learning architectures and reported promising classification performance using various evaluation metrics. Studies reporting higher performance generally utilized larger clinically annotated datasets or more advanced deep learning architectures, such as ViT and DenseNet-based models. In contrast, the present study emphasizes the comparative evaluation of conventional CNN and transfer learning models within a lightweight web-based clinical decision-support framework while incorporating preliminary expert validation. Rather than focusing solely on maximizing classification accuracy, the proposed framework provides a practical balance between classification performance, computational efficiency, deployment feasibility, and clinical applicability for real-time pressure ulcer assessment [8]-[11].

Table 5. Comparison with recent deep learning-based pressure ulcer classification studies

Study	Deep learning model	Best reported performance (%)	Main contribution	Web-based system	Expert validation
Fergus *et al*. [8]	Faster R-CNN	F1 = 67.86	Mobile-assisted pressure ulcer assessment	✓	✓
Aldughayfiq *et al*. [9]	YOLOv5	mAP = 76.9	Real-time pressure ulcer detection	✗	✗
Cho and Yoo [10]	ViT	Accuracy = 97.76	IoT-based pressure ulcer staging	✓	✗
Lei *et al*. [11]	AlexNet, VGG16, ResNet18, DenseNet121	Accuracy = 93.71	Comparative evaluation of CNN architectures	✗	✓
This study	Baseline CNN, Proposed CNN, MobileNetV2, ResNet50	Accuracy = 76.50	Comparative CNN evaluation with lightweight web-based decision support	✓	✓

Despite the encouraging results, several misclassifications were observed, particularly between adjacent stages such as stage 2 and stage 3, as well as stage 3 and stage 4. These errors are likely attributable to the high visual similarity between neighboring pressure ulcer stages, including similarities in wound color, tissue appearance, lesion depth, and wound boundaries. Similar challenges have also been reported in previous pressure ulcer classification studies because adjacent stages often exhibit overlapping visual characteristics in two-dimensional wound images [11], [30], [32]. Furthermore, clinical pressure ulcer staging relies not only on surface appearance but also on tissue involvement, wound depth, and clinical examination [1], [4]. Therefore, image-based classification alone may remain limited when complementary clinical information is unavailable.

The preliminary expert validation demonstrated that three out of four model predictions were consistent with expert annotations, indicating the potential of the proposed framework as a preliminary decision-support tool for pressure ulcer stage assessment. However, because only four representative cases were evaluated, these findings should be regarded as an initial assessment rather than comprehensive clinical

validation. The developed web-based implementation further demonstrates the practical feasibility of deploying the trained model through an accessible interface that enables users to upload wound images and obtain automated stage predictions with confidence scores. Although the achieved classification accuracy is not intended to replace clinical judgment, it demonstrates that the proposed framework can provide reliable preliminary decision support by assisting healthcare professionals in pressure ulcer stage assessment, particularly in resource-constrained healthcare settings where wound care specialists may not always be readily available. From an engineering perspective, the lightweight MobileNetV2 architecture requires relatively low computational resources, making it suitable for resource-constrained healthcare environments. In addition, the web-based framework supports scalable deployment through centralized model management, remote accessibility, and simplified software maintenance. These characteristics also facilitate future integration with hospital information systems (HIS) via standardized application programming interfaces (APIs), enabling automated pressure ulcer assessment results to be incorporated into electronic medical records and longitudinal patient monitoring [31], [33].

Nevertheless, the proposed framework should be considered a preliminary engineering prototype. The study primarily focused on evaluating different deep learning architectures and did not compare the proposed system with existing commercial wound assessment systems or established clinical decision-support platforms. Therefore, future research should include multicenter clinical validation involving multiple wound-care specialists and larger clinically annotated datasets. Comparative evaluation with existing clinical assessment platforms, standardized image acquisition protocols, and more advanced deep learning approaches, including explainable artificial intelligence (XAI), may further improve the reliability, scalability, interpretability, and clinical applicability of the proposed framework.

4. CONCLUSION

This study presented a web-based pressure ulcer stage classification framework that integrates comparative evaluation of conventional CNN and transfer learning models with an accessible clinical decision-support application for automated wound assessment. The proposed framework demonstrates the practical feasibility of combining lightweight deep learning models with web-based deployment to support real-time pressure ulcer stage prediction, thereby enhancing the accessibility and potential clinical applicability of artificial intelligence-assisted wound assessment.

Despite the encouraging performance, several limitations should be acknowledged. First, the study relied on a relatively small publicly available dataset, which may limit the generalizability of the developed models. In addition, the class distribution was not perfectly balanced across all pressure ulcer stages, which may have influenced the model's learning performance for minority classes. Furthermore, variations in image quality, illumination, viewing angle, and wound appearance within the dataset may have affected feature extraction and classification accuracy. Second, external clinical validation was limited to a preliminary expert assessment involving only four representative images and was therefore not intended as a comprehensive clinical evaluation. Finally, the classification models were developed using image data alone without incorporating additional clinical information that may further improve diagnostic accuracy.

Future research should focus on larger multicenter clinically annotated datasets, broader external clinical validation involving multiple wound-care specialists, real-time deployment and evaluation in routine healthcare environments, comparative evaluation with existing clinical assessment platforms, and the investigation of more advanced deep learning approaches and XAI techniques to improve further the reliability, scalability, interpretability, and clinical applicability of the proposed framework.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous certified wound care expert for assisting in the preliminary validation of the pressure ulcer stage classification results in this study. The authors also acknowledge the research support provided by Akademi Teknik Elektromedik Andakara and Akademi Keperawatan Andakara, DKI Jakarta, Indonesia, under Research Contract No. 007/KP/UPPM-ATEM/IV/2026.

FUNDING INFORMATION

This research was supported by Akademi Teknik Elektromedik Andakara and Akademi Keperawatan Andakara, DK Jakarta, Indonesia.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Suci Imani Putri	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓			
Ananda Prastuti Sutrisno	✓	✓		✓		✓	✓	✓	✓	✓			✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

DATA AVAILABILITY

The dataset used for model training and validation is openly available in the Pressure Ulcers Stages dataset on Kaggle at https://www.kaggle.com/datasets/sinemgokoz/pressure-ulcers-stages, reference number [22]. The four images used for preliminary expert validation were private author-provided samples and are not publicly available due to privacy considerations.

REFERENCES

[1] National Pressure Injury Advisory Panel, "NPIAP Pressure Injury Stages." https://cdn.ymaws.com/npiap.com/resource/resmgr/online_store/npiap_pressure_injury_stages.pdf (accessed Apr. 27, 2026).

[2] E. Roussou *et al*., "Quality of life of patients with pressure ulcers: a systematic review," *Medicine and Pharmacy Reports*, vol. 96, no. 2, pp. 123–130, 2023, doi: 10.15386/mpr-2531.

[3] S. R. H. Zaidi, S. El Khatib, and S. Sharma, "Pressure Ulcer", Treasure Island: StatPearls Publishing, 2024. https://www.ncbi.nlm.nih.gov/books/NBK553107/ (accessed Apr. 25, 2026).

[4] European Pressure Ulcer Advisory Panel, National Pressure Injury Advisory Panel, and Pan Pacific Pressure Injury Alliance, "*Prevention and treatment of pressure ulcers/injuries: Clinical practice guideline*", E. Haesler, ed.: EPUAP/NPIAP/PPPIA, 2019. https://www.internationalguideline.com/2019 (accessed Apr. 27, 2026).

[5] R. H. A. Al-Mallah, "Optimizing blood cell classification: evaluating feature dimensionality and validation strategies," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 24, no. 1, p. 359, 2026, doi: 10.12928/telkomnika.v24i1.27269.

[6] J. Florian, H. Agarwal, N. Singh, and B. R. Goldsmith, "Why halides enhance heterogeneous metal ion charge transfer reactions," *Chemical Science*, vol. 12, no. 38, pp. 12704–12710, 2021, doi: 10.1039/d1sc03642d.

[7] Y. Patel *et al.*, "Integrated image and location analysis for wound classification: a deep learning approach," *Scientific Reports*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-56626-w.

[8] P. Fergus, C. Chalmers, W. Henderson, D. Roberts, and A. Waraich, "Pressure Ulcer Categorisation using Deep Learning: A Clinical Trial to Evaluate Model Performance," 2022, *arXiv*. doi: 10.48550/ARXIV.2203.06248.

[9] B. Aldughayfiq, F. Ashfaq, N. Z. Jhanjhi, and M. Humayun, "YOLO-Based Deep Learning Model for Pressure Ulcer Detection and Classification," *Healthcare (Switzerland)*, vol. 11, no. 9, 2023, doi: 10.3390/healthcare11091222.

[10] Y. B. Cho and H. Yoo, "Development of a pressure ulcer stage determination system for community healthcare providers using a vision transformer deep learning model," *Medicine (United States)*, vol. 104, no. 7, p. e41530, 2025, doi: 10.1097/MD.0000000000041530.

[11] 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 Medical Informatics*, vol. 13, 2025, doi: 10.2196/62774.

[12] A. Brehmer *et al.*, "Fine-Grained Classification of Pressure Ulcers and Incontinence-Associated Dermatitis Using Multimodal Deep Learning: Algorithm Development and Validation Study," *JMIR AI*, vol. 4, no. 1, pp. e67356–e67356, May 2025, doi: 10.2196/67356.

[13] X. Yu, J. Wang, Q. Q. Hong, R. Teku, S. H. Wang, and Y. D. Zhang, "Transfer learning for medical images analyses: A survey," *Neurocomputing*, vol. 489, pp. 230–254, 2022, doi: 10.1016/j.neucom.2021.08.159.

[14] T. B. Nguyen-Tat, T. Q. Hung, P. T. Nam, and V. M. Ngo, "Evaluating pre-processing and deep learning methods in medical imaging: Combined effectiveness across multiple modalities," *Alexandria Engineering Journal*, vol. 119, pp. 558–586, 2025, doi: 10.1016/j.aej.2025.01.090.

[15] W. Xu, Y.-L. Fu, and D. Zhu, "ResNet and its application to medical image processing: Research progress and challenges," *Computer Methods and Programs in Biomedicine*, vol. 240, no. 1, Oct. 2023, doi: 10.1016/j.cmpb.2023.107660.

[16] H. Hong and S. Huang, "Survey of Residual Network in Image Processing," in *Proceedings of 2025 6th International Conference on Computer Information and Big Data Applications, CIBDA 2025*, 2025, pp. 1411–1416. doi: 10.1145/3746709.3746948.

[17] M. B. Irfansyah, B. Waheed, I. Winarno, and A. Alimudin, "Performance evaluation of serverless cloud-native API deployment: a

case study on a mobile health application," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 24, no. 1, p. 34, 2026, doi: 10.12928/telkomnika.v24i1.27261.
[18] R. van de Wetering, "IT-Enabled Clinical Decision Support: An Empirical Study on Antecedents and Mechanisms," *Journal of Healthcare Engineering*, vol. 2018, pp. 1–10, Dec. 2018, doi: 10.1155/2018/6945498.
[19] T. Islam, M. S. Hafiz, J. R. Jim, M. M. Kabir, and M. F. Mridha, "A systematic review of deep learning data augmentation in medical imaging: Recent advances and future research directions," *Healthcare Analytics*, vol. 5, 2024, doi: 10.1016/j.health.2024.100340.
[20] D. Rivoir, I. Funke, and S. Speidel, "On the pitfalls of Batch Normalization for end-to-end video learning: A study on surgical workflow analysis," *Medical Image Analysis*, vol. 94, 2024, doi: 10.1016/j.media.2024.103126.
[21] I. Salehin and D.-K. Kang, "A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain," *Electronics*, vol. 12, no. 14, Jul. 2023, doi: 10.3390/electronics12143106.
[22] S. Gokoz, "Pressure-ulcer-stages dataset." https://www.kaggle.com/datasets/sinemgokoz/pressure-ulcers-stages (accessed Apr. 12, 2026).
[23] F. Chollet, *Deep learning with python*. Shelter Island: Manning Publications Co., 2018.
[24] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge: The MIT Press, 2016.
[25] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed. Sebastopol: O'Reilly Media, Inc, 2019.
[26] M. Reyad, A. M. Sarhan, and M. Arafa, "A modified Adam algorithm for deep neural network optimization," *Neural Computing and Applications*, vol. 35, no. 23, pp. 17095–17112, 2023, doi: 10.1007/s00521-023-08568-z.
[27] C. M. Bishop, *Pattern recognition and machine learning*. New York: Springer, 2006.
[28] S. A. Hicks *et al.*, "On evaluation metrics for medical applications of artificial intelligence," *Scientific Reports*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-09954-8.
[29] B. Ay, B. Tasar, Z. Utlu, K. Ay, and G. Aydin, "Deep transfer learning-based visual classification of pressure injuries stages," *Neural Computing and Applications*, vol. 34, no. 18, pp. 16157–16168, 2022, doi: 10.1007/s00521-022-07274-6.
[30] J. Kim *et al.*, "Augmented Decision-Making in wound Care: Evaluating the clinical utility of a Deep-Learning model for pressure injury staging," *International Journal of Medical Informatics*, vol. 180, 2023, doi: 10.1016/j.ijmedinf.2023.105266.
[31] T. J. Liu *et al.*, "A pressure ulcers assessment system for diagnosis and decision making using convolutional neural networks," *Journal of the Formosan Medical Association*, vol. 121, no. 11, pp. 2227–2236, 2022, doi: 10.1016/j.jfma.2022.04.010.
[32] S. Seo *et al.*, "Visual classification of pressure injury stages for nurses: A deep learning model applying modern convolutional neural networks," *Journal of Advanced Nursing*, vol. 79, no. 8, pp. 3047–3056, 2023, doi: 10.1111/jan.15584.
[33] O. Y. Dweekat, S. S. Lam, and L. McGrath, "Machine Learning Techniques, Applications, and Potential Future Opportunities in Pressure Injuries (Bedsores) Management: A Systematic Review," *International Journal of Environmental Research and Public Health*, vol. 20, no. 1, 2023, doi: 10.3390/ijerph20010796.

BIOGRAPHIES OF AUTHORS



Suci Imani Putri received the Bachelor's degree in Electrical Engineering from Universitas Brawijaya, Indonesia, in 2011, and the Master's degree in Electrical Engineering from Universitas Brawijaya, Indonesia, in 2014. She is currently a lecturer at Akademi Teknik Elektromedik Andakara, Indonesia. Her research interests include artificial intelligence, medical instrumentation, biomedical signal processing, fuzzy logic systems, embedded systems, medical image analysis, and health information systems. She has published several scientific papers in the fields of medical technology, control systems, deep learning, and healthcare applications. She can be contacted at email: s.imaniputri@gmail.com.



Ananda Prastuti Sutrisno received the Bachelor of Nursing degree from Andalas University, Indonesia, in 2020, the Professional Nurse degree from Andalas University, Indonesia, in 2022, and the Master of Nursing Science degree from Universitas Gadjah Mada, Indonesia, in 2024. She is currently a lecturer in the Nursing Study Program at Akademi Keperawatan Andakara, Indonesia. Her academic and research interests include psychiatric nursing, nursing education, clinical nursing, community health nursing, and the application of health technology in nursing practice. She has been actively involved in research and community service activities related to nursing care, healthcare innovation, and nursing education. In this study, she contributed to the conceptualization of the pressure ulcer classification problem, clinical interpretation, data resources, validation coordination, and manuscript preparation. She can be contacted at email: anandaprastutis@gmail.com.