Original Article
http://mjiri.iums.ac.ir
Medical Journal of the Islamic Republic of Iran (MJIRI)

Med J Islam Repub Iran. 2026 (16 Feb);40.16. https://doi.org/10.47176/mjiri.40.16

Image Processing for Diagnosing Psoriasis: A Machine Learning Approach to Classify Skin Lesions into Psoriasis Subtypes

Hoorie Masoorian[1], Marsa Gholamzadeh[1], Alireza Firooz[2], Reza Safdari[1]*

Received: 7 May 2025 Accepted: 3 Sep 2025 Published: 16 Feb 2026

Abstract

Background: Psoriasis is a chronic autoimmune skin condition that affects 2-3% of the global population and manifests in various subtypes, including plaque, guttate, inverse, pustular, and erythrodermic psoriasis. Accurate subtype differentiation is crucial for effective treatment, but traditional diagnostic methods are time-consuming and prone to observer variability. This study aims to develop a machine learning model that classifies psoriasis lesions into the five primary subtypes using convolutional neural networks (CNNs) and transfer learning, offering a scalable tool to assist clinicians in diagnosing psoriasis and making informed treatment decisions.

Methods: This is a methodological–developmental study that develops and evaluates a deep learning model for psoriasis subtype classification. The dataset was obtained from from Kaggle, applying image augmentation techniques (rotation, translation, shearing, flipping, zoom) to enhance dataset diversity. A pre-trained Visual Geometry Group 16-layer architecture (VGG16) model was used for feature extraction, with a custom classification head added, incorporating ReLU-activated dense layers and dropout regularization to mitigate overfitting. The model was trained and evaluated using accuracy and loss metrics, with early stopping and model checkpointing for optimization.

Results: The model achieved 96% accuracy on the training dataset and 90% on the test dataset, demonstrating strong generalization. A confusion matrix analysis confirmed accurate differentiation between the five subtypes.

Conclusion: This study developed a deep learning model that accurately classifies psoriasis subtypes, utilizing CNNs and transfer learning. The model was integrated into a web-based tool, providing real-time diagnostic assistance for clinicians. This AI-driven system has the potential to enhance diagnostic accuracy, improve clinical workflows, and offer scalable solutions for psoriasis management, particularly in areas with limited access to dermatologists.

Keywords: Psoriasis, Machine Learning, Convolutional Neural Networks (CNN), Image Augmentation

Cite this article as: Masoorian H, Gholamzadeh M, Firooz A, Safdari R. Image Processing for Diagnosing Psoriasis: A Machine Learning Approach to Classify Skin Lesions into Psoriasis Subtypes. *Med J Islam Repub Iran.* 2026 (16 Feb);40:16. https://doi.org/10.47176/mjiri.40.16

Introduction

Psoriasis is a long-lasting inflammatory skin condition driven by the immune system, impacting about 2–3% of people worldwide (1). It is marked by red, flaky patches that can appear on various parts of the body, leading to considerable physical and emotional stress for those impacted (2). While psoriasis is generally not life-threatening, it is associated with significant physical discomfort, emotional distress, and an increased risk of comorbidities, such as cardiovascular diseases, diabetes, and mental health disorders (3-6).

Psoriasis manifests in various subtypes, each differing in severity, symptoms, and treatment responses. The primary psoriasis subtypes include plaque, guttate, inverse, pustular, erythrodermic, and nail psoriasis (5). Accurately differentiating between these subtypes is essential for effective treatment, as certain subtypes respond better to

Corresponding author: Dr Reza Safdari, rsafdari@tums.ac.ir

1. Health Information Management and Medical informatics Department, School of Allied Medical Science, Tehran University of Medical Sciences, Tehran, Iran
2. Center for Research and Training in Skin Diseases and Leprosy, Tehran University of Medical Sciences, Tehran, Iran

↑What is "already known" in this topic:
Effective treatment of psoriasis depends on a precise diagnosis of its subtypes. Traditional diagnosis relies on biopsy and clinical knowledge, both of which can be laborious and subjective. Convolutional neural networks (CNNs), a type of artificial intelligence, have shown promise in classifying dermatological images.

→What this article adds:
A VGG16-based deep learning model that has 90% test accuracy in classifying five psoriasis subtypes is presented in this study. It highlights the model's viability in real-time clinical settings, particularly in places with limited access to dermatology, and presents a web-based diagnostic tool.

[DOI: 10.47176/mjiri.40.16]

specific therapies than others. For example, plaque psoriasis, the most common subtype, often responds well to topical treatments, while pustular or erythrodermic psoriasis may require systemic or biologic therapies (6, 7).

In order to identify subtypes and plan treatments, visual information—especially lesion images—is essential (8).

Traditional psoriasis diagnosis often relies on clinical examination and, in some cases, biopsy, which can be time-consuming and subject to inter-observer variability (9). Consequently, researchers have explored the use of machine learning and deep learning methods for the automated classification of psoriasis subtypes (10).

In terms of the use of artificial intelligence (AI), especially in image-based diagnostics, dermatology has become a pioneering field. This is especially advantageous for conditions like psoriasis, which can be readily documented through photographs taken by patients and healthcare providers. The diagnostic process is streamlined by automated lesion detection and subtype classification made possible by AI-powered image analysis (11-13).

There are several benefits to using AI-based image processing for psoriasis diagnosis and treatment. Automated analysis of psoriasis photographs can improve diagnostic accuracy, reduce inter-observer variability, and enable more efficient clinical workflows. In addition, the creation of individualized treatment programs based on subtype characteristics is facilitated by extensive image evaluation. In the end, incorporating AI into dermatological procedures could revolutionize the clinical management of this complex chronic skin condition (14-16).

This study aims to develop a machine learning model that can accurately categorize photographs of psoriasis lesions into the five primary subtypes: plaque, guttate, inverse, pustular, and erythrodermic psoriasis. The objective was to create a reliable and scalable diagnostic tool to aid in clinical decision-making by utilizing convolutional neural networks (CNNs) and transfer learning.

Methods

ata Preparation and Augmentation

Step 1: Dataset description

This research is conducted as a methodological–developmental study, where a deep learning model is constructed and tested for the classification of psoriasis subtypes. This Dataset contains multiple images regarding various dermatological disorders. Our dataset was obtained from the Kaggle platform, a widely recognized online community and repository for data science and machine learning resources. Kaggle, owned by Google LLC, hosts a variety of datasets, competitions, and collaborative projects, making it a valuable tool for data scientists and researchers alike (17).

Medical photos of different dermatological conditions, especially psoriasis lesions, make up the dataset used in this investigation. Because they showed a wide range of skin conditions, these photos were essential for both training and testing our model because they helped it differentiate between various skin diseases. 307 photos in all, divided into five psoriasis subtypes (plaque, guttate, inverse, pustular, and erythrodermic) made up the original dataset. Plaque psoriasis had the most images (82 samples, 26.7%) and pustular psoriasis had the fewest (43 samples, 14.0%), indicating an uneven class distribution. Plaque (82), Inverse (67), Guttate (60), Erythrodermic (55), and Pustular (43) images were specifically included in the raw dataset. Biased learning and limited generalization were risks associated with this class imbalance.

We used specific data augmentation techniques, such as rotation, flipping, shearing, zooming, and translation, to artificially expand and balance the dataset in order to solve this problem .As indicated in Table 1, the final training dataset consisted of 1030 images after augmentation, with an additional 250 images for validation and another 250 for testing. Each class was then expanded to roughly 500–550 images. This method improved the model's resistance to overfitting and guaranteed a more evenly distributed class distribution.

By utilizing the Kaggle dataset, we ensured that our model was trained and tested on high-quality, real-world data—a crucial factor in developing a reliable and effective diagnostic system. Additionally, Kaggle's platform provided valuable metadata, standardized labeling, and community discussions, which informed our data preprocessing and modeling decisions.

Step 2: Image Augmentation

The distribution of samples across the disease classes in our dataset was not uniform. Deep networks may become overfitted, biased, and less accurate as a result of this imbalance. We used image augmentation techniques, which produce altered versions of the original data in order to increase the quantity and balance of images, to allay this worry. This procedure improves overall reliability and stabilizes model training(18), Data augmentation's primary objective is to increase machine learning models' accuracy and resilience so they can function dependably even in situations where the available data is limited or incompletely representative.(19).

Several augmentation operations were used to increase the dataset's robustness and balance, including::

Rotation: Each image was rotated along its axis by a randomly selected angle from a predefined distribution (0°–15°, 30°, 60°, or 90°) (20). This simulates varied orientations and improves model robustness to angular variations.

Translation: Images were shifted vertically and horizontally using a shift vector to apply translation, preserving the relative pixel arrangement and minimizing positional bias (21).

Shearing: A parallelogram-like shape was created by shearing, which caused angular deformation by moving one of the image's edges along the vertical or horizontal axis. A predetermined shear angle dictated the extent of this transformation (18).

Horizontal Flipping: Images were mirrored along the horizontal axis (preferred over vertical flipping, as the top and bottom of medical images are often anatomically distinct) (22).

[DOI: 10.47176/mjiri.40.16]

Random Zoom: Variations in zoom were applied to simulate differences in image scale.

These techniques collectively expanded the dataset's variability while maintaining biological plausibility.

Model Architecture

A custom classifier is used after a pre-trained VGG16 model for feature extraction in this transfer learning study. The following is an outline of the methodology's main steps:

Base Model: VGG16

We used the VGG16 architecture, a popular CNN that was first trained on ImageNet, for feature extraction. By retaining its convolutional layers and discarding the dense top layers, the model could transfer general visual representations to our domain-specific psoriasis dataset. (23). The model was adapted through the following modifications:

Top-layer removal: The fully connected (Dense) layers were discarded (include_top=False), retaining only the convolutional base for feature extraction.

Weight freezing: All pre-trained weights were locked (trainable=False) to preserve the model's learned representations and prevent further training of the VGG16 layers.

Input adjustment: The input dimensions were standardized to 180×180 pixels to match the resolution of the study's image dataset.

Custom Architecture

A new classification head was implemented atop the VGG16 feature extractor to adapt the model to the target task. The architecture consisted of the following sequential layers:

Flatten layer: Transformed the 3D feature maps (output from VGG16) into a 1D feature vector to enable processing by dense layers.

Dense layer (256 units, ReLU activation): Learned non-linear combinations of extracted features, with ReLU activation introducing sparsity and mitigating vanishing gradients.

Dropout layer (rate=0.5): Randomly deactivated 50% of neurons during training to prevent co-adaptation and reduce overfitting.

Based on initial tuning experiments and standard procedures in comparable CNN architectures, the 256-unit dense layer was chosen. According to well-established deep learning best practices (as recommended by (24, 25)) the 0.5 dropout rate was selected in order to successfully minimize overfitting while maintaining model capacity.

Output layer (Dense, softmax activation): Generated class probabilities with num_classes units, where softmax ensured probabilistic interpretation of outputs (for multi-class classification).

Hidden Layers

Non-linear combinations of the extracted features are learned by the custom model's single hidden dense layer, which has 256 neurons and ReLU activation.

Furthermore, as a regularization technique to avoid overfitting, a dropout layer with a rate of 0.5 is applied after the dense layer, randomly deactivating neurons during training(25).

In addition to improving generalization, this structure lowers the possibility of neuronal co-adaptation

Additional Details Regarding Model Architecture

A: Model Compilation

- Optimizer: The model is compiled using the Adam optimizer, which is well-suited for training deep learning models due to its adaptive learning rate properties(26).
- Loss Function: The categorical cross-entropy loss function is used, which is appropriate for multi-class classification tasks where the output is a probability distribution across multiple classes(27).
- Metrics: Accuracy is used as a metric to evaluate the model's performance during training and validation(28).

B: Model Checkpoints and Early Stopping

- Model Checkpoints: The Model Checkpoint callback is used to save the model's weights during training whenever the validation accuracy improves. This ensures that the best-performing model is retained(29).

We used some methods to prevent overfitting:

- Dropout: The dropout method is a deep learning technique that involves randomly setting each neuron to zero during training and maintaining the same scale during inference. It is a fundamental building block and helps prevent overfitting in neural architectures and applications(24).
- By locking VGG16 weights
- Data augmentation: Data augmentation involves artificially inflating the training dataset size through data warping or oversampling to extract more information from the original dataset(21).
- Splitting data appropriately
- Early-stopping: The Early Stopping callback monitors the validation loss and stops training if it does not improve for a specified number of epochs (patience), restoring the best weights. This helps prevent overfitting and saves training time (30). Early-stopping is a strategy to stop the search at the point of overfitting, achieved through cross-validation analysis on training data (35). It includes tracking the model's performance on a validation set throughout the training process and halting the training when performance begins to decline. This approach helps avoid overfitting to the training data by constraining the extent of training (36).

Model Training and tuning

In machine learning, it is essential to divide the dataset into training and testing subsets to ensure that the model can generalize well to new, unseen data. To ensure class balance across all subsets, the dataset was split into training (80%), validation (10%), and test (10%) sets using random stratified sampling. To give an objective assessment of model performance during tuning, the validation set was kept apart from the training and testing data. This process helps prevent overfitting, where a model performs well on training data but poorly on real-world data. In this

[DOI: 10.47176/mjiri.40.16]

step we split data into training and test sets, allowing a portion of the data (commonly 70–80%) to be used for training the model, while the remaining portion (usually 20–30%) is reserved for testing its performance. This step is crucial in evaluating how well a machine learning model is likely to perform in practice (31).

Both models were trained on the training dataset with validation on the validation set to monitor performance. To prevent overfitting and ensure the selection of the best-performing model, the following techniques were employed:

Using the Adam optimizer(26), the model was trained for a maximum of 50 epochs with a batch size of 32 and a learning rate of 0.0001. Model checkpointing was utilized to preserve the weights of the top-performing model, and early stopping was implemented with a patience value of 5.

- Early Stopping: Training was halted if the validation loss did not improve over a specified number of epochs, conserving computational resources and avoiding overfitting.
- Model Checkpointing: The best model weights, based on validation performance, were saved during training to facilitate optimal evaluation.

Model Evaluation

The trained models were evaluated on the validation and test datasets using metrics such as accuracy and loss. Accuracy and Loss are the two most well-known and discussed metrics in machine learning (32). The difference between the true and predicted class probabilities is measured by the categorical cross-entropy loss function. It guides learning through gradient descent by giving the model a quantitative goal to minimize during training (27). One of the most widely used metrics for assessing classification performance is accuracy. It shows the percentage of instances in which the model's predicted label corresponds to the actual label and is typically expressed as a percentage (33). A loss function, sometimes referred to as a cost function, measures the gap between predicted outcomes and the actual values while considering the associated probabilities or uncertainty. This provides a more detailed perspective on the model's performance than accuracy alone (34).

Model Implementation

Finally, our model was embedded in a psoriasis decision-making system to aid diagnose the type of psoriasis in patients with Golang programming language due to its high performance and optimal use of resources.

Results

Training, testing, and validation were the three main directories into which the dataset was arranged. Our workflow was streamlined by the meticulous labeling of each image. As illustrated in Figure 1, data was divided into subfolders for each of the five subtypes of psoriasis: erythrodermic, guttate, inverse, plaque, and pustular as showed in Figure 2. The Google Colab environment was used for both model development and training, making use of its computational resources to implement the models effectively.

Initially, our total data was 307, which we increased the number of photos with the augmentation method and distributed them into the test, training, and validation groups in the Table 1.

Summary of the Custom Architecture

The architecture of the developed deep learning model for dermatological disease classification was explained in previous section. The model accepts input images of shape (180*180), 3), corresponding to RGB images of specified dimensions. At its core, the architecture utilizes a pre-trained VGG16 model as the base network, with its layers frozen to retain the learned features from a large-scale dataset, thereby enabling efficient feature extraction without the need for extensive retraining. Following the VGG16 base, a flatten layer is employed to convert the 3D

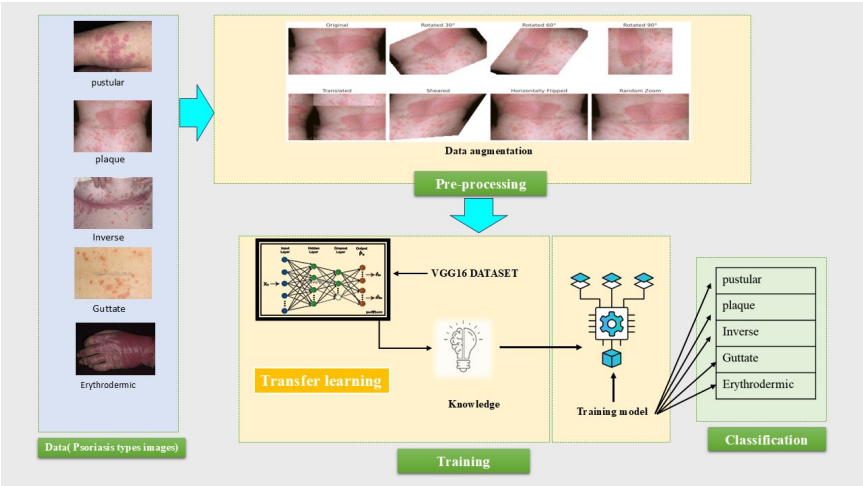

Figure 1. The process of performing image processing. designed by the author. Skin lesion images obtained from DermNet (www.dermnetnz.org).

[DOI: 10.47176/mjiri.40.16]

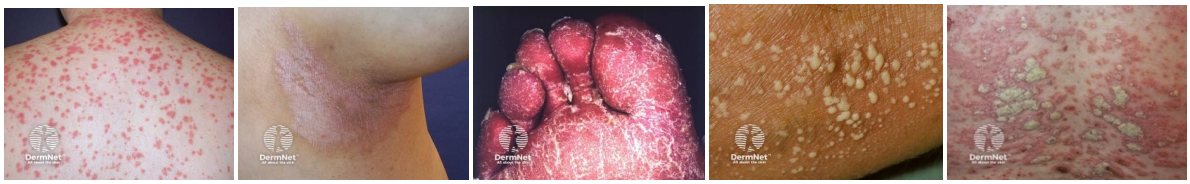

Figure 2. Representative lesion images. A: Guttate, B: Inverse, C: Erythrodermic, D: Pustular, E: Plaque. Image sourced from DermNet (www.dermnetnz.org)

Table 1. Data distribution across each psoriasis subtype in the dataset

Subtype	Training (N)	Validation (n)	Test (n)
Erythrodermic	200	50	50
Guttate	180	40	40
Inverse	220	55	55
Plaque	250	60	60
Pustular	180	45	45
Total	1030	250	250

Values represent the number of images in each dataset split

feature maps into a 1D vector, facilitating the transition to fully connected layers.

To mitigate the risk of overfitting, a dropout layer with a 50% dropout rate is incorporated, randomly deactivating half of the neurons during training to promote generalization. Subsequently, a dense layer with 256 neurons and ReLU activation function is added to introduce non-linearity and capture complex patterns in the data.

Finally, the output layer consists of 5 neurons with softmax activation, corresponding to the five dermatological disease classes, including psoriasis, to produce probabilistic predictions for multi-class classification. This architecture balances the benefits of transfer learning with custom modifications to optimize performance for the specific task of dermatological disease recognition.

Performance Metrics

To prevent overfitting, multiple strategies were employed, including dropout, frozen VGG16 weights, data augmentation, proper dataset splitting, and early stopping. The final model achieved an accuracy of approximately 96% on the training set and 90% on the test set, demonstrating strong generalization capability in classifying psoriasis from images. The training progress across epochs is illustrated in Figure 3.

Confusion Matrix Analysis

The confusion matrix provides a detailed breakdown of the model's classification performance, indicating how many samples from each class were correctly predicted versus misclassified. The results demonstrate strong discriminative capability, with the majority of samples in each class being accurately predicted as their true class. Specifically, most diagonal entries (representing correct predictions) exhibit significantly higher values compared to off-diagonal elements (misclassifications), suggesting robust class-specific feature learning. This alignment between predicted and actual labels further validates the model's effectiveness in distinguishing psoriasis from other conditions, as evidenced by the high accuracy metrics (96% training, 90% testing). Minor off-diagonal misclassifications may reflect inherent visual similarities between certain lesion types, warranting further investigation. The confusion matrix is depicted in Figure 4.

Model Deployment

This model has been deployed as a web-based tool to assist in the diagnosis of dermatological scars, with the

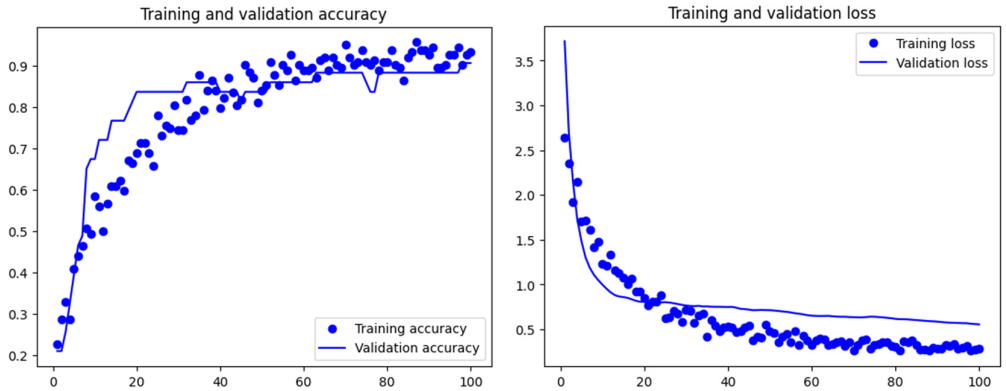

Figure 3. Training progress across epochs according to the developed model

[DOI: 10.47176/mjiri.40.16]

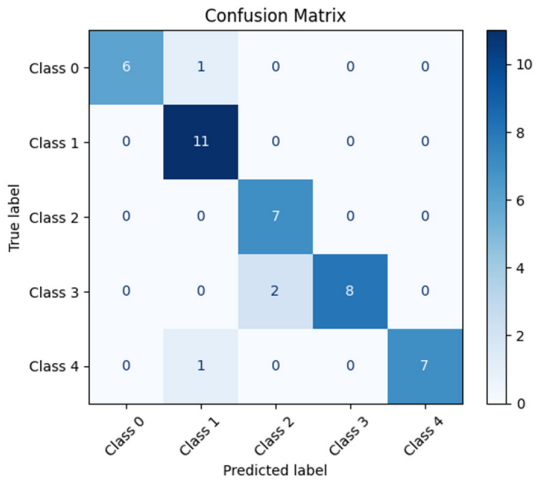

Figure 4. Confusion matrix showcasing classification accuracy across psoriasis subtypes.

oversight and approval of a specialist dermatologist. Patients can upload an image of the affected skin scar to the website, and the system provides a probabilistic classification of the scar type. The results, along with the uploaded image, are sent and shared with the dermatologist through a dedicated panel for review. The dermatologist can then evaluate the system's output and provide a final diagnosis, ensuring clinical accuracy and reliability. Additionally, the dermatologist can independently use the model within the panel to analyze patient images and make informed decisions, enhancing diagnostic precision and supporting more effective treatment planning for dermatological scars.

Any real-world deployment involving patient data must rigorously adhere to privacy and security protocols, even though the dataset used in this study is anonymized and publicly available. Depending on the jurisdiction, this includes encrypted storage, secure image transmission (e.g., HTTPS), authenticated access to user data, and adherence to ethical and legal frameworks like the Health Insurance Portability and Accountability Act (HIPAA) and the General Data Protection Regulation (GDPR) (30, 37).

Discussion

The use of deep learning techniques in conjunction with image processing to diagnose psoriasis represents a significant advancement in the application of artificial intelligence in dermatology. Using convolutional neural networks (CNNs) to capture and interpret the complex visual patterns associated with psoriatic lesions, we developed a dependable deep learning model in this study that can identify the condition from clinical skin images.To maximize the accessibility and practical applicability of our model, we implemented it as a web-based tool that allows healthcare professionals and patients to easily upload and analyze skin images in real time. This approach not only enhances the diagnostic process, but also bridges the gap between advanced computational techniques and clinical practice, providing a scalable solution for early diagnosis and management of psoriasis. The use of such a tool demonstrates the potential of AI in transforming skin care from traditional methods to innovative ones, especially in areas with limited access to dermatologists.

Our results are consistent with earlier research that used deep learning to diagnose psoriasis (38-40). However, our approaches differ in terms of our methodology from existing models, likely due to employing effective preprocessing techniques such as data augmentation and transfer learning, or innovative architecture design. Unlike prior studies that focused solely on model development (41-43), Additionally, our work offers a web-based, easily navigable decision support tool for practical uses.

The strengths of this study lie in its integration of image processing and deep learning to enhance automated psoriasis diagnosis, bridging the gap between advanced AI techniques and clinical practice. The model leverages CNNs and transfer learning to extract complex visual patterns associated with psoriasis, ensuring efficient and accurate feature representation. A key strength is the customized deep learning architecture, which balances pretrained feature extraction with additional layers to optimize performance while mitigating overfitting (41, 44).

ReLU-activated dense layers and dropout layers are used to further enhance pattern recognition and generalization (42, 45), which is also used in our study.

The deployment of our model as a web-based tool represents a significant step toward bridging the gap between advanced AI-based techniques and clinical practice. This tool enables clinicians in resource-limited settings to access more accurate diagnostic capabilities without the need for specialized equipment or expertise. In reality, developing AI models without integrating them into practical applications offers limited value today (37, 46, 47).

Additionally, developing such tool empowers patients to monitor their condition as soon as possible, potentially leading to earlier intervention and improved outcomes. Additionally, the web-based implementation of the model significantly improves accessibility and usability, allowing both healthcare professionals and patients to analyze skin images in real time. This approach provides a scalable, cost-effective solution for early detection, particularly benefiting regions with limited access to dermatologists. Despite the performance gap between training and test accuracy, the study highlights the importance of dataset diversity and lays the groundwork for further optimization through data augmentation and additional regularization techniques.

Among the factors affecting the results, we can mention the choice of model, image enhancement techniques, and preprocessing strategies. Unlike Zhang et al. and Chen et al. (24), our study utilized a smaller dataset. Nevertheless, we were able to achieve high accuracy metrics (96% training, 90% testing), which we attribute to both the choice of model and the development techniques we employed, including careful model tuning and image enhancement. The VGG16 model was selected for its simple yet power-

[DOI: 10.47176/mjiri.40.16]

ful architecture, consisting of uniform 3×3 convolutional layers stacked with increasing depth (48), making it highly suitable for fine-tuning and practical implementation. With 16 weight layers, VGG16 effectively captures rich hierarchical features while maintaining architectural simplicity, which is ideal for transfer learning applications (49). Its widespread use in domains where interpretability and robustness are critical further supports our decision (50). We chose VGG16 because of its interpretability, ease of design, and demonstrated performance in small datasets, even though more recent architectures like ResNet and EfficientNet have demonstrated encouraging results in medical imaging tasks. VGG16 is a viable and dependable option for this application because prior research (51, 52) has shown that it performs competitively, particularly when optimized on sparse medical image data. For instance, Tajbakhsh et al. employed VGG16 as a feature extractor in medical image classification tasks such as lung nodule detection, demonstrating that fine-tuning the model led to significant performance gains, particularly in scenarios with limited labeled data (52).

Similarly, Shin et al. compared multiple CNN architectures, including VGG16, in various medical imaging contexts and found that VGG16 consistently outperformed other models such as AlexNet when fine-tuned, particularly for thoraco-abdominal lymph node detection and interstitial lung disease classification. Their findings reinforced the model's strength in generalizing features learned from natural images to medical imaging tasks, due to its deep and transferable representations (51). In addition to model architecture, other studies have shown that advanced image preprocessing methods, such as contrast-limited adaptive Histogram Equalization (CLAHE), can improve visual contrast and highlight finer details, thereby boosting classification accuracy. Incorporating such techniques could potentially enhance our model's performance further. Lastly, the use of more sophisticated data augmentation methods—such as random erasing or color jittering—could introduce greater variability and help simulate real-world scenarios more effectively, contributing to improved generalization and robustness in future implementations.

Although the results are encouraging, our study has limitations. Despite its reliability, the training and validation dataset might not adequately represent the range of psoriasis presentations across skin types and demographics. The high diversity of skin diseases and the comparatively small dataset also pose difficulties. Factors like internet access and image quality may also have an impact on the web-based tool's performance, which may limit its usefulness in particular contexts.

Although, unlike the studies of Zhang et al. and Chen et al. (53), our study used a smaller dataset, the application of advanced data augmentation and transfer learning methods was able to overcome the limitations of the data size. Curating a larger, multi-center dataset will be part of future work to increase the model's generalizability and robustness. As is well known in dermatological datasets, the dataset's high levels of intra-class variability and inter-class visual similarities may introduce bias during training, which could impact the model's capacity to generalize (32). It should be noted that this study only used a held-out dataset from the same source for internal validation. For stronger generalizability, future studies will include external validation on independent datasets collected from diverse devices, environments, and patient populations.

Conclusion

A convolutional neural network-based model for psoriasis subtype classification using clinical images was created and validated in this study. The model successfully distinguished between the five main psoriasis subtypes, attaining 96% accuracy on the training set and 90% accuracy on the test set. These findings support our objective of developing a scalable, easily accessible diagnostic tool to help physicians identify psoriasis variations early and accurately.

The model's deployment as a web-based tool marks a step forward in integrating AI technologies into clinical dermatology, particularly in settings with limited access to specialists. However, while our findings are promising, they are grounded in the specific dataset and methodological choices of this study. Broader validation across diverse skin types, clinical settings, and image sources is necessary to confirm generalizability.

This study supports the feasibility of AI-powered diagnostic tools in dermatology and highlights the need for further research into their clinical and ethical integration. Based on the results, we propose that deep learning–based image classification holds potential to enhance diagnostic accuracy for psoriasis, provided that robust, transparent, and clinically validated systems are developed.

Acknowledgment

The authors would like to thank the Health Information Management Department, Tehran University of Medical Sciences, for their general support and the colleagues who provided technical assistance during this project.

Conflict of Interests

The authors declare that they have no competing interests.

Authors' Contributions

Hoorie Masoorian designed the study, developed the methodology, prepared and processed the dataset, implemented the deep learning model, and drafted the manuscript.

Marsa Gholamzadeh contributed to data analysis and assisted in revising the manuscript.

Alireza Firooz provided clinical consultation and evaluated the medical validity of the work.

Reza Safdari supervised the overall research process, contributed to the study design, and reviewed the manuscript critically.

All authors reviewed and approved the final version of the manuscript.

Ethical Considerations

This study was approved by the Tehran University of

[DOI: 10.47176/mjiri.40.16]

Medical Sciences Ethics Committee

Ethics approval code.: IR.TUMS.SPH.REC.1402.003. All methods were performed according to the relevant guidelines and regulations.

Funding Support

The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data, or in writing the manuscript.

Data availability

The dataset used in this study was obtained from a publicly accessible source on the Kaggle platform. The processed data and trained model generated during this study can be made available by the corresponding author upon reasonable request.

AI Use Statement

Artificial intelligence methods, particularly deep learning techniques, were employed as the main approach for image classification in this study.

AI-based writing assistance tools were used solely to improve the language and readability of the manuscript. These tools had no role in the analysis, modeling, or generation of the research results.

References

1. Lee HJ, Kim M. Challenges and future trends in the treatment of psoriasis. Int J Mol Sci. 2023;24(17):13313.
2. Griffiths CE, Armstrong AW, Gudjonsson JE, Barker JN. Psoriasis. The Lancet. 2021;397(10281):1301-15.
3. Khawaja AR, Bokhari SMA, Rasheed T, Shahzad A, Hanif M, Qadeer F, et al. Disease severity, quality of life, and psychiatric morbidity in patients with psoriasis with reference to sociodemographic, lifestyle, and clinical variables: a prospective, cross-sectional study from Lahore, Pakistan. The primary care companion for CNS disorders. 2015;17(3):26260.
4. Mpofana N, Makgobole M, Nxumalo CT, Pillay P. Psoriasis: Clinical Features and Its Impact on Quality of Life. 2024.
5. Ni C, Chiu MW. Psoriasis and comorbidities: links and risks. Clinical, cosmetic and investigational dermatology. 2014:119-32.
6. Rungapiromnan W. Examining Multimorbidity in Patients with Psoriasis and the Impact of Biologic Therapies on the Risk of Major Cardiovascular Events: The University of Manchester (United Kingdom); 2018.
7. Kuchekar AB, Pujari RR, Kuchekar SB, Dhole SN, Mule PM, Vidyapeeth B, et al. Psoriasis: A comprehensive review. Int J Pharm Life Sci. 2011;2(6):857-77.
8. Rendon A, Schäkel K. Psoriasis pathogenesis and treatment. Int J Mol Sci. 2019;20(6):1475.
9. Borgbjerg J. Development and Application of a Web-Based Platform for Assessment of Observer Performance in Medical Imaging. 2022.
10. Kassem MA, Hosny KM, Damaševičius R, Eltoukhy MM. Machine learning and deep learning methods for skin lesion classification and diagnosis: a systematic review. Diagnostics. 2021;11(8):1390.
11. Karalis VD. The integration of artificial intelligence into clinical practice. Applied Biosciences. 2024;3(1):14-44.
12. Liu Z, Wang X, Ma Y, Lin Y, Wang G. Artificial intelligence in psoriasis: Where we are and where we are going. Experimental Dermatology. 2023;32(11):1884-99.
13. Yu K, Syed MN, Bernardis E, Gelfand JM. Machine learning applications in the evaluation and management of psoriasis: a systematic review. J Psoriasis Psoriatic Arthritis. 2020;5(4):147-59.
14. Luo A, Gou S, Tong N, Liu B, Jiao L, Xu H, et al. Visual interpretable MRI fine grading of meniscus injury for intelligent assisted diagnosis and treatment. NPJ Digital Medicine. 2024;7(1):97.
15. Panagoulias DP, Tsoureli-Nikita E, Virvou M, Tsihrintzis GA. Dermacen Analytica: A Novel Methodology Integrating Multi-Modal Large Language Models with Machine Learning in tele-dermatology. arXiv preprint arXiv:240314243. 2024.
16. Tariq S, Raza R, Sargano AB, Habib Z. Magnification independent breast cancer analysis using vision transformer. Multimedia Tools and Applications. 2024:1-29.
17. Akshay B, Pulari SR, Murugesh T, Vasudevan SK. Machine Learning: A Comprehensive Beginner's Guide: CRC Press; 2024.
18. Goceri E. Medical image data augmentation: techniques, comparisons and interpretations. Artificial Intelligence Review. 2023;56(11):12561-605.
19. Mumuni A, Mumuni F. Data augmentation: A comprehensive survey of modern approaches. Array. 2022;16:100258.
20. Wang G, Wang Y, Bao X, Huang D, editors. Rotation Has Two Sides: Evaluating Data Augmentation for Deep One-class Classification. The Twelfth International Conference on Learning Representations; 2023.
21. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data. 2019;6(1):1-48.
22. Nalepa J, Marcinkiewicz M, Kawulok M. Data augmentation for brain-tumor segmentation: a review. Frontiers in computational neuroscience. 2019;13:83.
23. Jiang Z-P, Liu Y-Y, Shao Z-E, Huang K-W. An improved VGG16 model for pneumonia image classification. Applied Sciences. 2021;11(23):11185.
24. Liu Z, Xu Z, Jin J, Shen Z, Darrell T, editors. Dropout reduces underfitting. International Conference on Machine Learning; 2023: PMLR.
25. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res. 2014;15(1):1929-58.
26. Zhang Z, editor Improved adam optimizer for deep neural networks. 2018 IEEE/ACM 26th international symposium on quality of service (IWQoS); 2018: Ieee.
27. Mao A, Mohri M, Zhong Y, editors. Cross-entropy loss functions: Theoretical analysis and applications. International conference on Machine learning; 2023: PMLR.
28. Rabinovich SG, Rabinovich M. Evaluating measurement accuracy: Springer; 2010.
29. Rojas E, Pérez D, Calhoun JC, Gomez LB, Jones T, Meneses E, editors. Understanding soft error sensitivity of deep learning models and frameworks through checkpoint alteration. 2021 IEEE International Conference on Cluster Computing (CLUSTER); 2021: IEEE.
30. Thakur A, Gupta M, Sinha DK, Mishra KK, Venkatesan VK, Guluwadi S. Transformative breast Cancer diagnosis using CNNs with optimized ReduceLROnPlateau and Early stopping Enhancements. Int J Comput Intell Syst. 2024;17(1):14.
31. Amr T. Hands-On Machine Learning with scikit-learn and Scientific Python Toolkits: A practical guide to implementing supervised and unsupervised machine learning algorithms in Python: Packt Publishing Ltd; 2020.
32. Hicks SA, Strümke I, Thambawita V, Hammou M, Riegler MA, Halvorsen P, et al. On evaluation metrics for medical applications of artificial intelligence. Scientific reports. 2022;12(1):5979.
33. Rainio O, Teuho J, Klén R. Evaluation metrics and statistical tests for machine learning. Scientific Reports. 2024;14(1):6086.
34. Muhammad NA, Rehman A, Shoaib U. Accuracy based feature ranking metric for multi-label text classification. Int J Adv Comput Sci Appl. 2017;8(10).
35. Loughrey J, Cunningham P. Using early-stopping to avoid overfitting in wrapper-based feature selection employing stochastic search. Trinity College Dublin, Department of Computer Science; 2005.
36. Diukarev V, Starukhin Y. Proposed methods for preventing overfitting in machine learning and deep learning. Asian J Res Comput Sci. 2024;17(10):85-94.
37. Nayyar A, Gadhavi L, Zaman N. Machine learning in healthcare: review, opportunities and challenges. Machine

[DOI: 10.47176/mjiri.40.16]

learning and the internet of medical things in healthcare. 2021:23-45.
38. Hammad M, Pławiak P, ElAffendi M, El-Latif AAA, Latif AAA. Enhanced deep learning approach for accurate eczema and psoriasis skin detection. Sensors. 2023;23(16):7295.
39. Yaseliani M, Maghsoodi AI, Hassannayebi E, Aickelin U. Diagnostic clinical decision support based on deep learning and knowledge-based systems for psoriasis: From diagnosis to treatment options. Computers & Industrial Engineering. 2024;187:109754.
40. Yu Z, Kaizhi S, Jianwen H, Guanyu Y, Yonggang W. A deep learning-based approach toward differentiating scalp psoriasis and seborrheic dermatitis from dermoscopic images. Frontiers in Medicine. 2022;9:965423.
41. He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
42. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems. 2012;25.
43. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014.
44. Zhou Z, Shin J, Zhang L, Gurudu S, Gotway M, Liang J, editors. Fine-tuning convolutional neural networks for biomedical image analysis: actively and incrementally. Proceedings of the IEEE conference on computer vision and pattern recognition; 2017.
45. He K, Zhang X, Ren S, Sun J, editors. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. Proceedings of the IEEE international conference on computer vision; 2015.
46. Nozari H, Sadeghi ME. Artificial intelligence and Machine Learning for Real-world problems (A survey). Int J Innov Eng. 2021;1(3):38-47.
47. Sinha S, Lee YM. Challenges with developing and deploying AI models and applications in industrial systems. Discover Artificial Intelligence. 2024;4(1):55.
48. Lindberg A, Larsson L. Evaluating the Performance of Extended Convolutional Networks: A Comparative Study between VGG-16 and VGG-23 for Image Classification. 2024.
49. Tammina S. Transfer learning using vgg-16 with deep convolutional neural network for classifying images. Int J Sci Res Publ. (IJSRP). 2019;9(10):143-50.
50. Ahmed F, Abbas S, Athar A, Shahzad T, Khan WA, Alharbi M, et al. Identification of kidney stones in KUB X-ray images using VGG16 empowered with explainable artificial intelligence. Scientific Reports. 2024;14(1):6173.
51. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE transactions on medical imaging. 2016;35(5):1285-98.
52. Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, et al. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE transactions on medical imaging. 2016;35(5):1299-312.
53. Li W, Chen C, Zhang M, Li H, Du Q. Data augmentation for hyperspectral image classification with deep CNN. IEEE Geoscience and Remote Sensing Letters. 2018;16(4):593-7.

[DOI: 10.47176/mjiri.40.16]