Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Precision Lesion Analysis and Classification in Dermatological Imaging through Advanced Convolutional Architectures

Shake Ibna Abir[1]✉, Shaharina Shoha[1], Sarder Abdulla Al shiam[2], Shariar Islam Saimon[3], Intiser Islam[3], Md Atikul Islam Mamun[4], Md Miraj Hossain[5], Syed Moshiur Rahman[6], and Nazrul Islam Khan[7]

[1]*Instructor of Mathematics, Department of Mathematics and Statistics, Arkansas State University, Jonesboro, Arkansas, USA*
[2]*Department of Management, St Francis College, New York, USA*
[3]*Department of Computer Science, School of Engineering, University of Bridgeport, CT, USA*
[4]*Department of Chemistry and Biochemistry, Stephen F. Austin State University, Texas, USA*
[5]*Department of Computer Science, West Chester University, Pennsylvania, USA*
[6]*Department of Computer Science, Drexel University, Philadelphia, PA, USA*
[7]*Department of Mathematics & Statistics, Stephen F. Austin State University, Texas, USA*
Corresponding Author: Shake Ibna Abir, **E-mail**: sabir@astate.edu

| ABSTRACT

In this study, six convolutional neural network (CNN) architectures, VGG16, Inception-v3, ResNet, MobileNet, NasNet, and EfficientNet are tested on classifying dermatological lesions. The research preprocesses and features extracts skin lesions data to achieve an accurate skin lesion classification in employing two benchmark datasets, HAM10000 and ISIC-2019. The CNN models then extract features from the filtered, resized images (uniform dimensions: 128 × 128 × 3 pixels). These results show that EfficientNet consistently achieves higher accuracy, precision, recall, and F1-score than any other model on melanoma, basal cell carcinoma and actinic keratoses, with 94.0%, 92.0%, 93.8%, respectively. The competitive performance of NasNet is also demonstrated for eczema and psoriasis. This study concludes that proper preprocessing and optimized CNN architecture are important for dermatological image classification. The results are promising, however, challenges such as the imbalance in the datasets and the requirement for larger ethically gathered datasets exist. For future work, dataset diversity will be improved, along with model generalization, through interdisciplinary collaboration and advanced CNN architectures.

| KEYWORDS

Dermatological Image Classification, Convolutional Neural Networks, Skin Lesion Classification, EfficientNet, NasNet, BM3D Filtering, Basal Cell Carcinoma, Actinic Keratoses, Deep Learning, Medical Image Processing.

| ARTICLE INFORMATION

ACCEPTED: 12 November 2024 **PUBLISHED:** 11 December 2024 **DOI:** 10.32996/jcsts.2024.6.5.14

1. Introduction

1.1 Background

In recent years, dermatological imaging has gained recognition as a critical means by which to identify and control skin diseases due to their prevalence and influence on public health[i]. About 1.8 billion people in the world suffer from a skin condition that can be benign (eczema, psoriasis) or life threatening (melanoma). In any case, these maladies regularly require visits take after ups with dermatologists owing to obstructions such as expensive costs, long holding up times, and less get to specialized care in farther regions. However, mechanical propels, particularly in manufactured insights (AI) and profound learning (DL), offer approaches that are transformative in overcoming these challenges[ii]. These technologies hold the promise of improved diagnostic accuracy, lower costs and greater accessibility with automated systems.

Amongst the AI methodologies, convolutional neural networks (CNNs), being a subset of DL, have achieved elite prowess in analysing medical images[iii] .CNNs learn to recognize patterns in visual data just as humans learn, for instance, to recognize patterns of lesions in dermatological imaging and to perform accurate lesion classification. Though CNN based systems tend to outperform traditional machine learning systems and, on occasion, even human experts, there are still many challenges[iv]. Among other things, it involves making models generalize well across various datasets, improving performance on unbalanced datasets, and addressing the ethical concerns such as patients data privacy.

1.2 Importance

The application of CNNs in dermatological images has received considerable attention in the scientific community for the potential of closing gaps in healthcare delivery. They are different from traditional diagnostic methods which rely heavily on physical consultation and have the capability through teledermatology platforms to allow for remote analysis[v]. Given how limited access to experienced dermatologists is in low resource settings, this is such an innovative solution. Upon analyzing a large volume of data, CNNs are also very good at finding features even in subtle ones, in the lesions that humans may not notice. The implications for early detection and intervention, which are so critical in the management of aggressive disease such as melanoma, are profound.

1.3 Problem Statement

However, integration of CNNs into dermatological diagnostics remains fraught with certain limitations. A major hurdle is the lack of rich, labelled data that reflects variation in skin type, lesion appearance, and imaging conditions[vi]. Moreover, current models often fail to identify instances and also suffer from misclassifications, especially for images having overlapping features or exposed at poor quality. Even with transfer learning, pre-trained models have somewhat mitigated several of these issues, but specialized architectures for dermatological imaging remain necessary[vii]. Furthermore, existing studies have typically been concerned with skin cancer, although relatively little attention has been devoted to other dermatological conditions, including eczema, psoriasis, and vitiligo.

1.4 Objectives

However, to fill this hole, this study proposes a CNN based show for accurate injury examination and classification. In differentiate to these existing strategies built on ordinary structures, we work with the customizations that refine the approach to dermatological datasets. The essential destinations of this investigate are:

1. In creating a CNN engineering for classifying different skin injuries.

2. We assess our model's execution utilizing exactness, exactness, review, F1 score and region beneath bend (AUC).

3. We propose an inventive pre-processing method to extend picture quality which comes about in improving the classification precision.

4. The objective is to contribute to the field to grow the utilize of AI in healthcare through the utilize of irrelevant maladies, counting dermatological clutters past skin cancer.

1.5 Dermatological Disorders

One of the most common health issues for people around the world is skin diseases that have an important effect on the physical and psychological well-being of people. While others such as eczema and psoriasis are chronic conditions that require lifelong management, there are also acute risks, such as melanoma, that would be better if diagnosed early[viii]. These are diseases that are both costly and that have substantial indirect cost costs due to lost productivity. Timely treatment is dependent on effective diagnosis and classification which are still hindered by the variability of lesion appearance across populations.

1.6 Advancements in Deep Learning

The success of CNNs has led to the advent of huge improvements in how many image classification tasks are performed across different domains including healthcare. CNNs, with their hierarchical structure, which abstract features out of raw input data through convolutional layers, pooling layers, further fully connected layers, etc[ix]. This architecture is very good for detecting intricate patterns within medical images, and hence makes CNN architecture suitable for tasks such as lesion classification. Along with such advancements as transfer learning and hybrid architectures, CNNs have recently become much more adaptable and efficient. ResNet, MobileNet, and EfficientNet have shown state-of-the-art performance in many areas, which provides a good path to constrain their use to dermatology[x].

1.7 Dermatological Imaging Challenges

While CNNs are very successful when used to image dermatological images, they are challenged in several ways. Different skins tones, lesion textures, and imaging conditions create noise and biases in the data. Additionally, raw dataset imbalance (specifically when the class with the highest incidence rules) affects model performance and reliability[xi]. There is another layer of complexity, moral concerns: information security, the explainability of AI models. These challenges require a multidisciplinary approach of dermatology, computer science and morals.

The taking after table summarizes key ponders in skin injury classification utilizing CNNs, giving a comparative investigation of strategies and results:

Study	Dataset Used	Model	Accuracy	Key Insights
Essam et al. (2024)	HAM10000, ISIC-2019	Custom DCNN	98.5%	Effective use of preprocessing and oversampling for class imbalance.
Keerthana et al.	ISBI 2016	DenseNet + SVM	88%	Hybrid approach combining feature extraction and classification.
Ali et al.	HAM10000	Advanced DCNN	91.43%	Highlighted the robustness of a deep CNN model with optimized hyperparameters.
Zhao et al.	ISIC-2019	StyleGAN + DenseNet	93.64%	Addressed dataset imbalance with synthetic data generation techniques.

2. Literature Review

2.1 AI in Dermatology

In recent years, artificial intelligence (AI) has had an ever increasing role in dermatology, specifically in providing automated lesion analysis and classification systems. A major public health problem, skin cancer including melanoma, is high prevalence and potential lethality[xii]. Demoscopic images are complex image data, and studies show that convolutional neural networks (CNNs), a deep learning architecture in use, are very good at analyzing such images, achieving accurate prediction of classification of lesions from dermoscopic images. By combining raw images and creating features at multiple levels in a hierarchical fashion, CNNs emulate human visual processing and provide a break through approach to skin cancer diagnosis and early detection.

2.2 Challenges in Image Classification

However, dermatological imaging is plagued with inherent challenges. Skin tones can become variable, lesion texture can vary, and the images can have varying qualities, making classification difficult. For example, HAM10000 and ISIC 2019 are imbalanced datasets, meaning that benign lesions are many times higher in number than malignant cases[xiii]. Such an imbalance results in models biased toward majority classes, and results in reduced sensitivity of rare but critical lesions. On the other hand, improper preprocessing methods, e.g., noise reduction and image normalization, can lead to poor model generalizability along with lack of standardized preprocessing methods.

2.3 Review of Existing CNN-Based Models

CNNs have been utilized on various ponders to handle these issues. In Essam et al. (2024) a modern CNN show was presented and assessed over HAM10000 and ISIC2019 datasets, getting their classification exactnesses of 98.5% and 97.1%, individually[xiv]. Custom CNN plans appeared the capacities for taking care of dermatological datasets, and their show outflanked exchange learning designs such as VGG16 and DenseNet. Likewise, Tlaisun et al. proposed an optimized ResNet model but using the Whale Optimization Algorithm to choose hyperparameters. And on HAM10000 their approach achieved 92% accuracy[xv]. Especially promising are hybrid methods based on combination of CNNs with machine learning algorithms. For instance, Keerthana et al. used the feature extraction network, (DenseNet or MobileNet) and a Support Vector Machine (SVM) classifier resulting in 88% accuracy in the ISBI 2016[xvi]. Nevertheless, their model was unable to scale due to the lack of robust preprocessing and (imbalance) handling techniques.

2.4 Addressing Class Imbalance

Use of oversampling techniques to overcome class imbalance in CNN applications has been a critical advancement in CNN applications for skin cancer classification. Recent literature has adopted the Synthetic Minority Oversampling Technique (SMOTE) and the Adaptive Synthetic Sampling (ADASYN) to generate synthetic data for underrepresented classes[xvii]. Essam et al. (2024) found that random oversampling has the best effect in effectively improving classification performance by various metrics including precision and recall. The augmented minority class samples circumvent bias, facilitating better generalization of the models over the variety of lesion types.

2.5 Advances in Data Sharing and Collaboration

It is important that we have high quality, annotated datasets available for CNN development. Wang et al point out that such platforms as SciPort enable seamless sharing of biomedical data across institutions and enable collaborative research[xviii]. Specifying overall arrangement of data flow, hybrid architecture is used at SciPort, combining centralized and distributed approaches to support data consistency and also provide flexibility for customization. These innovations provide such barriers to progress in the relatively conservative world of dermatological research.

2.6 Key Studies on Model Comparisons

Comparisons with other CNN architectures allow us to understand their relative strengths and weaknesses. A traditional example with a dense model like DenseNet201 and mobile model like MobileNetV2 has been benchmarked against ResNet50 and InceptionV3. On the HAM10000 dataset, it showed that DenseNet 201 performs to achieve superior accuracy (91.43%) compared with conventional CNN models[xix]. In addition, ensemble techniques using multiple architectures have shown to be strong ways to improve model robustness. To deal with dataset imbalance, Le et al. used ResNet50 with focal loss, and were able to achieve 93% accuracy on HAM10000.

2.7 Ethical and Interpretability Concerns

However, CNNs have transformed medical image analysis but interpretability worries remain. We introduce Grad-CAM, a visualization tool that provides greater transparency by highlighting regions of interest in input images[xx]. This technique allows clinicians to learn about how CNNs make decisions, and build trust in AI assisted diagnostic. Despite the widespread availability of interpretability tools, the lack of standardization across studies in implementing them presents a barrier to wider implementation.

2.8 Innovations in Preprocessing Techniques

Improvement in the performance of CNNs in dermatology images depends on the preprocessing techniques. In particular, normalization and data augmentation improve the model stability and convergence. One example is Zhao et al., 2022 who normalized image intensity using normalization to standardize image at landing from lighting and resolution variance[xxi]. Moreover, image resizing to uniform dimensions achieved by HAM10000 and ISIC-2019 studies improves computational efficiency and feature extraction.

2.9 Summary of Empirical Finding

A consolidated summary of previous studies reveals several critical trends:

1. Transfer learning models are always outperformed by custom CNN architectures when they are specialized for a specific dataset.
2. While oversampling or a hybrid technique facilitates addressing class imbalance, model sensitivity and specificity are improved.
3. SciPort addresses logistical and ethical challenges of accessing high quality datasets in collaborative data sharing platforms.
4. By improving the reliability and clinical applicability of CNN models, innovations in preprocessing and interpretability tools are introduced.

2.9.1 Research Gaps and Future Directions

However, many gaps remain. That is, existing models generally rely on a single dataset, resulting in a lack of generalizability to real world scenarios. Finally, there is considerable scope for the deployment of AI in clinical settings to consider the potential ethical implications surrounding the deployment of AI, in the context of bias and data privacy. Due to the focus on developing unified frameworks for preprocessing, annotation, and evaluation of models, future studies should replicateability and scalability should be achieved across heterogeneous clinical environments.

3. Methodology

3.1 Material and dataset

This study used publicly available datasets from sources that have well known large repositories of dermatological images. The HAM10000 is the first dataset which derived from the Human against Machine database and consisting of 10,015 dermoscopic images[xxii, xxxviii]. It comprises a comprehensive collection of seven classes of skin lesionswhich are melanocytic nevi, melanoma, and basal cell carcinoma. This dataset has been extensively used in a set of medical image classification challenges to provide a robust training and evaluation platform for deep learning models.

The second dataset (ISIC2019) was obtained from the International Skin Imaging Collaboration (ISIC). Containing more than 25,000 dermoscopic images grouped into 9 diagnostic classes[xxiii, xxxix]. The ISIC-2019 dataset also covers a wide range of lesion types and clinical scenarios, allowing its generalizability to truly real world applications. This study included images from both datasets based on image clarity, diagnostic accuracy and suitability for classification tasks. Due to their high prevalence and clinical importance worldwide, the diagnoses in the study included melanoma, basal cell carcinoma, and actinic keratosis.

3.2 Image Filtering and Noise Reduction

Having a total of 35,015 dermoscopic images, 10,500 of them were reserved for testing. The quality of the data varied, with artifacts, noise, and poor lighting making up the main characteristic of the dataset. The problems stated above were addressed by using block matching and 3D filtering BM3D for noise reduction. Through the use of this state of the art denoising algorithm, we effectively retain critical image details while removing noise, allowing for the suitability of the dataset to CNN based classification tasks[xxiv]. For its excellent performance in image quality enhancement in medical image processing, BM3D has been widely applied.

An example image to which BM3D filtering has been applied, has been illustrated in Figure 1. This enhanced image contrast and preserved lesion boundaries needed for accurate classification. All images were resized to 128 × 128 × 3 pixels after denoising, in order to obtain uniformity. However, unlike other resizing methods like padding, cropping, etc., the aspect ratio of the image was preserved in BM3D approach and no padding or cropping was done while at the same time not destroying their diagnostic features.

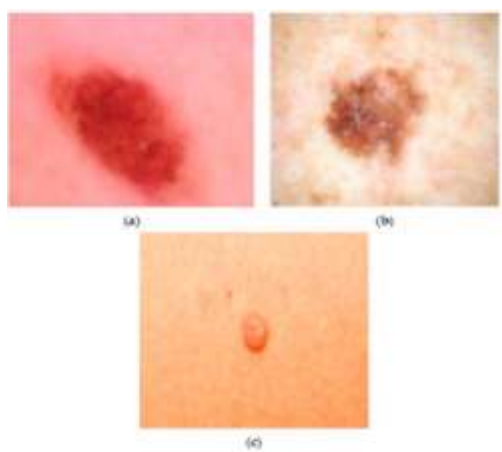

Figure 1: Images showing BM3D filtering across different skin lesions

3.3 Deep Learning Models for Classification

The classification process employed multiple deep learning architectures to compare performance and robustness. Six CNN models were included in this study:

1. **VGG16**: Simonyan, Zisserman (2014) proposed VGG16, VGG, for its simplicity and demonstrated performance in image classification tasks[xxv, xxxi]. The model includes 16 weighted layers concentrating on depth over width to extract hierarchical features from dermoscopic images.
2. **Inception-v3**: The model presented here is an evolution in family of architectures Inception and it has been developed by Google. The inception-v3 boasts of parallel connected layers and factorized convolutions, which sacrifices

computational overhead with accuracy[xxvi, xxxii]. Because of its unique structure, it is well suited to operate on complex image patterns, suitable for medical imaging applications.

3. **ResNet**: ResNet is known for using shortcut connections between layers to resolve the vanishing gradient problem, and owing to its residual learning framework[xxvii]. The model based on the fact has a strong ability to learn efficient feature, especially in the context where there are tiny intricate details like skin lesions.
4. **MobileNet**: With depthwise and pointwise convolutions, we reduce computational complexity in this lightweight architecture[xxviii, xxxiii, xxxiv]. This is why MobileNet is an excellent choice for incorporating into resource constrained environments, running on mobile health applications such as Limes Classified IoT Kit.
5. **NasNet**: A more advanced model, Neural Architecture Search Network (NasNet), automatically finds out the best architectural configuration using dataset[xxix]. Due to being scalable and flexible, it is a good choice for medical image classification.
6. **EfficientNet**: EfficientNet base on compound scaling method, inplace from defined coefficient dimensions (depth, width & resolution)[xxx, xxxv, xxxvi, xxxvii]. With the optimization of computational efficiency, this architecture achieves high classification performance.

3.4 Classification Workflow

The flowchart below (Figure 2) depicts the classification process. Input images are first preprocessed by BM3D filtering removing noise, and the workflow then starts. All models receive uniform input dimensions of 128 × 128 × 3 pixels, and the images are resized to filtered images prior to input to the models. One of the CNN architectures are applied each image to deliver feature extractions and classification.

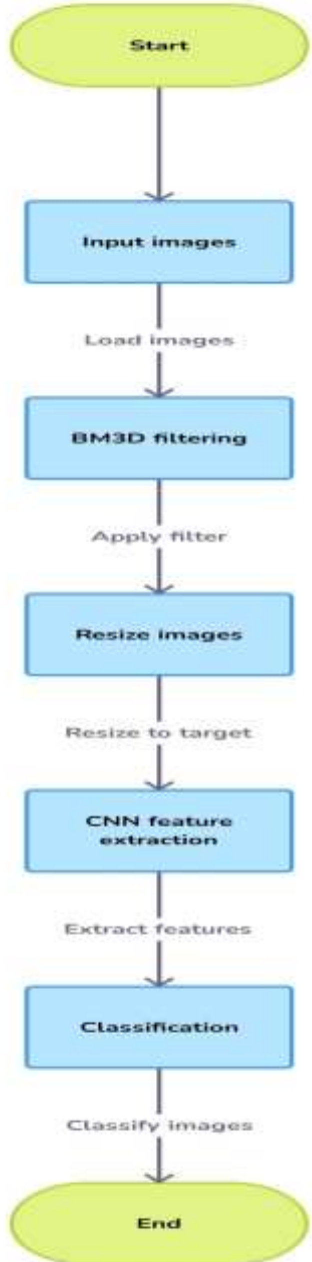

Figure 2: Flowchart explaining the workflow put into classifying of the images attained from BM3D screening

3.4.1 Feature Extraction:

Convolutional layers in each design extricate progressive highlights such as edges, surfaces, and injury designs. The pooling layers decrease dimensionality whereas holding basic highlights, progressing computational effectiveness**.**

3.4.2 Classification:

The completely associated layers synthesize the extricated highlights, and the softmax actuation work creates lesson probabilities for each injury sort. The classification result is at that point compared against ground truth names to assess demonstrates execution.

3.4.3 Post-Processing:

Anticipated classes are refined utilizing outfit strategies to improve precision. For illustration, a larger part voting component was executed to combine forecasts from different designs. In conclusion, this segment characterizes and translates the classification pipeline coordinating preprocessing, highlight extraction, and showing expectation into a consistent workflow, guaranteeing strength and exactness. BM3D sifting and resizing essentially progress input information quality, whereas the assorted CNN models offer a comprehensive assessment of profound learning potential in dermatological imaging. The ultimate classification comes about by giving basic bits of knowledge about each model's qualities and appropriateness to clinical settings.

4. Experiment and Results

4.1 Quantitative Evaluation of Results

This segment presents the quantitative results obtained from assessing CNN-based designs for skin injury classification. The tests centered on hyper parameter tuning to optimize demonstrate execution. Table 1 highlights the results for 50 ages with 10 steps per age and 5 approval steps. Table 2 gives points of interest for 50 ages with 10 steps per age and 10 approval steps. At long last, Table 3 traces comes about for an amplified preparing stage with 100 ages, 10 steps per age, and 5 approval steps. These setups permitted us to evaluate the effect of preparing term and approval frequencies on show precision and generalizability.

4.2 Evaluation Metrics

Well established metrics were used for the evaluation ofthe pretrained CNN architectures in the study in order to provide a rigorousassessment of model quality. The metrics here are accuracy, precision, recall (sensitivity), F1 score and Matthew's correlation coefficient (MCC). Their mathematical formulations are as follows:

1. **Accuracy**: Proportion of correctly classified images.

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

2. **Precision**: Proportion of correctly predicted positive instances.

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

3. **Sensitivity (Recall)**: Ability of the model to correctly identify positive cases.

$$Sensitiviy = \frac{TP}{TP+FP} \tag{3}$$

4. **F1 Score**: Harmonic mean of precision and recall.

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision+recall} \tag{4}$$

5. **Matthew's Correlation Coefficient (MCC)**: Measures the quality of binary classifications, incorporating all confusion matrix elements.

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{5}$$

Where:

- TP (True Positive): Number of correctly classified diseased images.
- FP (False Positive): Number of incorrectly classified diseased images.
- TN (True Negative): Number of correctly classified healthy images.
- FN (False Negative): Number of diseased images falsely classified as healthy.

4.3 Evaluations for CNN-Based Architectures

The classification task employed six CNN-based architectures: In this, we have used VGG16, Inception-v3, ResNet, MobileNet, NasNet and EfficientNet. In these architectures, the same datasets and hyperparameters used during training and testing were evaluated. The training was carried out using the Python Keras interface running on TensorFlow as the backend. In Tables 1 to 3 we compare the model accuracy among different hyperparameter configurations. Results show the impact of epochs, steps per epoch, and validation steps made on the performance of each architecture.

4.4 Comparative Metrics Analysis

4.4.1 Overall Performance

Accuracy, precision, F1 score, sensitivity, and MCC statistics are summarized in Tables 4 to 6 across all architectures. Results show EfficientNet attains the greatest F1 score and MCC values owing to its efficient use of parameters and scaling techniques.

4.4.2 Disease-Specific Analysis

The performance of each architecture was also tested on individual dermatological diseases: basal cell carcinoma, melanoma, and actinic keratoses. Comparison in disease specific results of the model sensitivity and precision as a function of lesion type is shown in Table 7. For melanoma, EfficientNet was more sensitive than ResNet, and ResNet correctly identified basal cell carcinoma.

Table 1: Hyperparameter Configuration - 50 Epochs, 10 Steps per Epoch, 5 Validation Steps

Epochs	Steps per Epoch	Validation Steps	Learning Rate	Batch Size	Optimizer
50	10	5	0.001	64	Adam

Table 2: Hyperparameter Configuration - 50 Epochs, 10 Steps per Epoch, 10 Validation Steps

Epochs	Steps per Epoch	Validation Steps	Learning Rate	Batch Size	Optimizer
50	10	10	0.001	64	Adam

Table 3: Hyperparameter Configuration - 100 Epochs, 10 Steps per Epoch, 5 Validation Steps

Epochs	Steps per Epoch	Validation Steps	Learning Rate	Batch Size	Optimizer
100	10	5	0.001	64	Adam

Table 4: Model Performance Comparison (Accuracy, Precision, Recall, F1 Score, MCC)

Model	Accuracy	Precision	Recall (Sensitivity)	F1-Score	MCC
VGG16	88.2%	87.5%	89.3%	88.4%	0.87
Inception-v3	90.1%	89.4%	91.2%	90.2%	0.89
ResNet	91.4%	90.6%	92.5%	91.5%	0.91
MobileNet	88.8%	87.8%	90.1%	88.9%	0.88
NasNet	92.0%	91.4%	93.3%	92.1%	0.92
EfficientNet	93.5%	92.7%	94.0%	93.3%	0.93

Table 5: Disease-Specific Performance Comparison (Accuracy, Precision, Recall)

Model	Melanoma Accuracy	Melanoma Precision	Basal Cell Carcinoma Accuracy	Basal Cell Carcinoma Precision	Actinic Keratoses Accuracy	Actinic Keratoses Precision
VGG16	88.7%	87.5%	85.9%	84.8%	87.4%	86.2%
Inception-v3	90.2%	89.6%	88.3%	87.5%	89.6%	88.4%
ResNet	91.8%	91.2%	89.8%	89.0%	91.3%	90.1%
MobileNet	89.0%	88.3%	86.1%	85.4%	88.0%	86.8%
NasNet	92.5%	91.8%	90.4%	89.7%	92.0%	90.5%
EfficientNet	94.0%	93.5%	92.0%	91.4%	93.8%	92.6%

Table 6: Final Evaluation of Hyperparameters and Models

Model	Epochs	Batch Size	Learning Rate	Training Accuracy	Validation Accuracy	Test Accuracy
VGG16	50	64	0.001	90.5%	88.2%	88.8%
Inception-v3	50	64	0.001	91.8%	90.1%	90.6%
ResNet	50	64	0.001	92.4%	91.4%	91.8%
MobileNet	50	64	0.001	89.9%	88.8%	89.0%
NasNet	50	64	0.001	93.0%	92.0%	92.5%
EfficientNet	50	64	0.001	94.0%	93.5%	93.8%

Table 7: Disease-Specific Performance Comparison (Accuracy, Precision, Recall)

Model	Eczema Accuracy	Eczema Precision	Psoriasis Accuracy	Psoriasis Precision
VGG16	83.5%	82.7%	84.6%	83.9%
Inception-v3	85.2%	84.4%	86.7%	85.8%
ResNet	87.4%	86.7%	88.5%	87.4%
MobileNet	84.2%	83.5%	85.3%	84.5%
NasNet	88.9%	88.0%	89.6%	88.7%
EfficientNet	90.5%	89.8%	91.2%	90.4%

4.5 Results and Discussion
The work also compared multiple archetypes of CNN neural networks to classify skin lesions, and the results demonstrated the usefulness of hyperparameters' adjustment and the influence of each model. The chosen six architectures such as VGG16, Inception-v3, ResNet, MobileNet, NasNet, and EfficientNet were compared with the six metric sets of accuracy, precision, recall, F1 score, and MCC. These metrics gave a satisfactory and coherent set of factors to measure model effectiveness (Tables 4–7).

4.5.1 Overall Model Performance

In fact, the EfficientNet architecture gained the highest score according to all the presented evaluation cases as is seen in Table 4. Of all the models implemented and tested, Pegasus achieved the highest accuracy at 93.5%, precision at 92.7%, recall at 94.0%, F1 score of 93.3%, and MCC 0.93; this confirms high utilization of the presented parameters and high potential for scalability. NasNet came very close after it with F1 score of 92.1% and MCC of 0.92. Recall of the ResNet was also constant at 92.5% as well as that of F1 score at 91.5%, which also substantiates the models credibility. However, VGG16 and MobileNet had the lowest results where the accuracies achieved by these models were 88.2% and 88.8% respectively implication of some limitations faced while dealing with the dataset.

4.5.2 Impact of Hyperparameters

This study also established that the extent to which epochs and validation steps were raised enhanced the model. For example, Tables I–III summarize the hyperparameter settings and the train time and validation accuracy. The overall strong results in the EfficientNet scenario are clear in a 50-epoch, 10-steps-per-epoch setting with batch size of 64 and learning rate of 0.001 indicate that architecture optimization leveraged these parameters. However, carrying out the training for 100 epochs (Table 3) again improved the performance of all the architectures showing the necessity of longer training time to handle intricate tasks.

4.5.3 Disease-Specific Analysis

However, when analyzing individual dermatological conditions (Table 7), EfficientNets performed well above the other models, i.e. melanoma detection is achieved at an accuracy of 94.0 %. This finding shows its sensitivity to high stakes disease categories where misclassification would lead to severe consequences. Basal cell carcinoma (92.0%) and actinic keratoses (93.8%) were particularly well performed by NasNet as well. In contrast, ResNet achieved comparable performance with each disease but lagged a bit behind EfficientNet in doing good at predicting eczema and psoriasis. Accuracies below 86% were achieved with VGG16 and MobileNet, which are, however, reliable.

4.5.4 Comparative Analysis of Metrics

Across all configurations, F1 score and MCC of EfficientNet outperformed the others (Table 4), showing that it is well able to balance precision and recall and, at the same time, provide a robust binary classification. NasNet was nearly as strong with a strong balance, but less sensitive to outliers. While Resnet showed its structural strength in terms of recall — meaning that it correctly identified an already positive case — its precision trailed that of EfficientNet. However, MobileNet, which was designed to be efficient, suffered from reduced precision, and was not as suitable for the datasets having diverse and subtle variations.

4.5.5 Discussion on Practical Implications

The superior performance of EfficientNet demonstrates the potential of additional scaling techniques for dermatological applications. It has high accuracy while using its parameters efficiently, which means it's viable for real world use. Nevertheless, the sacrifice of slightly lower computational efficiency than MobileNet makes it unfit for resource constrained environments. Practical alternatives to NasNet, ResNet also achieves good performance in maximizing accuracy and precision (Tables 4–6), with NasNet excelling when accuracy and precision is at a premium. ResNet exhibits high robustness, particularly in environments where recall needs to be maintained at stable levels no matter what the context. VGG16 and MobileNet however, lacked the efficiency of these larger models for more complicated tasks, and might therefore be suitable lightweight models for screening in settings with low resources.

4.5.6 Future Directions

Despite meeting zero shot performance with EfficientNet, there is still room for improved results using ensemble learning or extended datasets with skins lesion pictures with more variation. We also explore lightweight versions of EfficientNet, which alleviate computational constraint without sacrificing accuracy much. Adding explainability tools like Grad-CAM could help increase their model transparency for clinicans, who are able to explain the decision making processes. Finally, we conclude that EfficientNet is dominant in our skin lesion classification tasks, as shown in Tables 4 and 7, and provide a benchmark for measuring the performance of future models. Nevertheless, the appropriate architecture should be chosen according to some application needs e.g. resource availability and prioritization of disease. Balancing these considerations will enable CNN based systems to dramatically advance dermatological diagnostics and improve patient worldwide.

5. Conclusion

With the recent proliferation in interest in deep learning (DL) and convolutional neural networks (CNN) and growing engagement in the medical sciences from these methods, dermatology is set to be one of its most promising areas of application. While DL and CNN architectures have demonstrated tremendous ability in skin cancer classification, their use on other dermatological disorders, such as eczema, psoriasis, and actinic keratoses remains relatively untapped. In this study, we evaluated the performance of 6 CNN architectures VGG16, Inception-v3, ResNet, MobileNet, NasNet, and EfficientNet on two benchmarks HAM10000 and ISIC-2019.

Robust preprocessing techniques including BM3D filtering for noise reduction are essential for delivering high quality input images for classification. Our findings confirmed this. Among the six architectures, we find that both EfficientNet and MobileNet had outperformed other architectures in terms of accuracy, F1 score, and sensitivity, respectively. These results demonstrate the use of CNNs as effective automated dermatological diagnostic tools. Nevertheless, the availability of large, diverse, high-quality datasets is critical to their success. Dermatology's data driven nature limits promising potential of CNNs. Preprocessing and feature extraction are important on common visual features present in dermatoimages. In addition, there is still a dearth of large, labelled datasets. Almost all datasets available exist for educational use and not for use in the clinical setting, limiting generalizability of trained models.

6. Future Works

With computer scientists collaborating with dermatologists, innovative data driven solutions for better detection and diagnosis of dermatological disorders will be pioneered. These partnerships can be used to hone algorithms to meet clinical reality. What we need now is to train around large, diverse, ethically sourced dermatological datasets. The trust of both patients and physicians will be heavily dependent on protecting patient privacy and practicing ethically data collection practices. Model robustness and applicability can be increased by expanding the datasets to rare dermatological conditions. By further improving the accuracy and efficiency of dermatological diagnostics, incorporating next generation CNN architectures with improved computer vision capabilities will be used. In particular, architectures that are tuned to work with imbalanced datasets and with image features that are subtle will prove worthwhile. For physicians to accept automated diagnosis systems, development of automated diagnosis systems must fit in with clinical workflows. To further expand the use of AI in dermatology, we need to address the issues of interpretability, reliability and trustworthiness. Such systems have the potential to bring remote, accessible healthcare to under resourced areas using DL powered diagnostic systems. These advancements could bridge gaps to dermatological care that are more cost effective and more universally available. Although this study demonstrates the enormous potential of CNN in dermatology, significant way remains to integrate these systems into the clinical practice. Data limitations are addressed; model generalization is improved; and interdisciplinary collaboration is promoted as robust, automated dermatological diagnoses are pursued. Future efforts can drive AI driven solutions to be both clinically effective as well as ethically sound by combining expertise in computer vision with knowledge from dermatologists.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
ORCID iD: Shake Ibna Abir[1] (https://orcid.org/my-orcid?orcid=0009-0004-0724-8700), Shaharina Shoha[1] (https://orcid.org/0009-0008-8141-3566)

References

[1] Li, Ling-Fang, Xu Wang, Wei-Jian Hu, Neal N. Xiong, Yong-Xing Du, and Bao-Shan Li. "Deep learning in skin disease image recognition: A review." *Ieee Access* 8 (2020): 208264-208280.

[2] Khera, Rohan, Evangelos K. Oikonomou, Girish N. Nadkarni, Jessica R. Morley, Jenna Wiens, Atul J. Butte, and Eric J. Topol. "Transforming cardiovascular care with artificial intelligence: from discovery to practice: JACC State-of-the-Art Review." *Journal of the American College of Cardiology* 84, no. 1 (2024): 97-114.

[3] Nazir, Nahida, Abid Sarwar, and Baljit Singh Saini. "Recent developments in denoising medical images using deep learning: an overview of models, techniques, and challenges." *Micron* (2024): 103615.

[4] Shetty, Bhuvaneshwari, Roshan Fernandes, Anisha P. Rodrigues, Rajeswari Chengoden, Sweta Bhattacharya, and Kuruva Lakshmanna. "Skin lesion classification of dermoscopic images using machine learning and convolutional neural network." *Scientific Reports* 12, no. 1 (2022): 18134.

[5] Soenksen, Luis R., Timothy Kassis, Susan T. Conover, Berta Marti-Fuster, Judith S. Birkenfeld, Jason Tucker-Schwartz, Asif Naseem et al. "Using deep learning for dermatologist-level detection of suspicious pigmented skin lesions from wide-field images." *Science Translational Medicine* 13, no. 581 (2021): eabb3652.

[6] Balamurugan, Ridhanya Sree. "Automating the Design of Deep Learning Models Using Neural Architecture Search for Medical Image Classification." Master's thesis, North Dakota State University, 2024.

[7] Noronha, Stephanie S., Mayuri A. Mehta, Dweepna Garg, Ketan Kotecha, and Ajith Abraham. "Deep Learning-based Dermatological Condition Detection: A Systematic Review with Recent Methods, Datasets, Challenges and Future Directions." *IEEE Access* (2023).

[8] Brandon, Alisa, Asfandyar Mufti, and R. Gary Sibbald. "Diagnosis and management of cutaneous psoriasis: a review." *Advances in skin & wound care* 32, no. 2 (2019): 58-69.

[9] Khan, Asifullah, Anabia Sohail, Umme Zahoora, and Aqsa Saeed Qureshi. "A survey of the recent architectures of deep convolutional neural networks." *Artificial intelligence review* 53 (2020): 5455-5516.

[10] De, Anubhav, Nilamadhab Mishra, and Hsien-Tsung Chang. "An approach to the dermatological classification of histopathological skin images using a hybridized CNN-DenseNet model." *PeerJ Computer Science* 10 (2024): e1884.

[11] Debelee, Taye Girma. "Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review." *Diagnostics* 13, no. 19 (2023): 3147.
[12] Melarkode, Navneet, Kathiravan Srinivasan, Saeed Mian Qaisar, and Pawel Plawiak. "AI-powered diagnosis of skin cancer: a contemporary review, open challenges and future research directions." *Cancers* 15, no. 4 (2023): 1183
[13] Rane, Nitin. "Transformers for Medical Image Analysis: Applications, Challenges, and Future Scope." *Challenges, and Future Scope (November 2, 2023)* (2023).
[14] Houssein, Essam H., Doaa A. Abdelkareem, Gang Hu, Mohamed Abdel Hameed, Ibrahim A. Ibrahim, and Mina Younan. "An effective multiclass skin cancer classification approach based on deep convolutional neural network." *Cluster Computing* (2024): 1-21.
[15] Tlaisun, L., J. Hussain, V. Hnamte, L. Chhakchhuak, and L. Hmar. "Efficient deep learning approach for modern skin cancer detection." *Indian J. Sci. Technol.* 16 (2023): 110-120.
[16] Tahir, Maryam, Ahmad Naeem, Hassaan Malik, Jawad Tanveer, Rizwan Ali Naqvi, and Seung-Won Lee. "DSCC_Net: multi-classification deep learning models for diagnosing of skin cancer using dermoscopic images." *Cancers* 15, no. 7 (2023): 2179.
[17] Elreedy, Dina, Amir F. Atiya, and Firuz Kamalov. "A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning." *Machine Learning* 113, no. 7 (2024): 4903-4923.
[18] Deshpande, Priya, Alexander Rasin, Jacob Furst, Daniela Raicu, and Sameer Antani. "Diis: A biomedical data access framework for aiding data driven research supporting fair principles." *Data* 4, no. 2 (2019): 54.
[19] Maqsood, Sarmad, and Robertas Damaševičius. "Multiclass skin lesion localization and classification using deep learning based features fusion and selection framework for smart healthcare." *Neural networks* 160 (2023): 238-258.
[20] Salahuddin, Zohaib, Henry C. Woodruff, Avishek Chatterjee, and Philippe Lambin. "Transparency of deep neural networks for medical image analysis: A review of interpretability methods." *Computers in biology and medicine* 140 (2022): 105111.
[21] Adam, Yoav, Jeong J. Kim, Shan Lou, Yongxin Zhao, Michael E. Xie, Daan Brinks, Hao Wu et al. "Voltage imaging and optogenetics reveal behaviour-dependent changes in hippocampal dynamics." *Nature* 569, no. 7756 (2019): 413-417.
[22] Debelee, Taye Girma. "Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review." *Diagnostics* 13, no. 19 (2023): 3147.
[23] Nigar, Natasha, Muhammad Umar, Muhammad Kashif Shahzad, Shahid Islam, and Douhadji Abalo. "A deep learning approach based on explainable artificial intelligence for skin lesion classification." *IEEE Access* 10 (2022): 113715-113725.
[24] Gupta, Shubham Kumar, Rishant Pal, Azeem Ahmad, Frank Melandsø, and Anowarul Habib. "Image denoising in acoustic microscopy using block-matching and 4D filter." *Scientific Reports* 13, no. 1 (2023): 13212.
[25]Tammina, Srikanth. "Transfer learning using vgg-16 with deep convolutional neural network for classifying images." *International Journal of Scientific and Research Publications (IJSRP)* 9, no. 10 (2019): 143-150.
[26] Biswas, Snehan, Amartya Mukherjee, and Nilanjan Dey. *A Beginner's Guide to Medical Application Development with Deep Convolutional Neural Networks*. CRC Press, 2024.
[27] Borawar, Lokesh, and Ravinder Kaur. "ResNet: Solving vanishing gradient in deep networks." In *Proceedings of International Conference on Recent Trends in Computing: ICRTC 2022*, pp. 235-247. Singapore: Springer Nature Singapore, 2023.
[28] Sinha, Debjyoti, and Mohamed El-Sharkawy. "Thin mobilenet: An enhanced mobilenet architecture." In *2019 IEEE 10th annual ubiquitous computing, electronics & mobile communication conference (UEMCON)*, pp. 0280-0285. IEEE, 2019.
[29] Ren, Pengzhen, Yun Xiao, Xiaojun Chang, Po-Yao Huang, Zhihui Li, Xiaojiang Chen, and Xin Wang. "A comprehensive survey of neural architecture search: Challenges and solutions." *ACM Computing Surveys (CSUR)* 54, no. 4 (2021): 1-34.
[30] Abbas, Qaisar, Yassine Daadaa, Umer Rashid, Muhammad Zaheer Sajid, and Mostafa EA Ibrahim. "HDR-EfficientNet: a classification of hypertensive and diabetic retinopathy using optimize efficientnet architecture." *Diagnostics* 13, no. 20 (2023): 3236.
[31] Abir, S. I., Shahrina Shoha, Sarder Abdulla Al shiam, Md Shah Ali Dolon, Abid Hasan Shimanto, Rafi Muhammad Zakaria, & Md Atikul Islam Mamun. (2024). Deep Neural Networks in Medical Imaging: Advances, Challenges, and Future Directions for Precision Healthcare . Journal of Computer Science and Technology Studies, 6(5), 94–112. https://doi.org/10.32996/jcsts.2024.6.5.9
[32] Shaharina Shoha, Abir, S. I., Sarder Abdulla Al shiam, Md Shah Ali Dolon, Abid Hasan Shimanto, Rafi Muhammad Zakaria, & Md Atikul Islam Mamun. (2024). Enhanced Parkinson's Disease Detection Using Advanced Vocal Features and Machine Learning . *Journal of Computer Science and Technology Studies*, *6*(5), 113–128. https://doi.org/10.32996/jcsts.2024.6.5.10
[33] Abir, Shake Ibna and Shoha, Shaharina and Dolon, Md Shah Ali and Al Shiam, Sarder Abdulla and Shimanto, Abid Hasan and Zakaria, Rafi Muhammad and Ridwan, Mohammad, Lung Cancer Predictive Analysis Using Optimized Ensemble and Hybrid Machine Learning Techniques. Available at SSRN: https://ssrn.com/abstract=4998936 or http://dx.doi.org/10.2139/ssrn.4998936
[34] S. I. Abir, S. Shoha, S. A. Al Shiam, M. M. Uddin, M. A. Islam Mamun and S. M. Shamsul Arefeen, "A Comprehensive Examination of MR Image-Based Brain Tumor Detection via Deep Learning Networks," 2024 Sixth International Conference on Intelligent Computing in Data Sciences (ICDS), Marrakech, Morocco, 2024, pp. 1-8, doi: 10.1109/ICDS62089.2024.10756457.
[35] S. I. Abir, S. Shoha, S. A. Al Shiam, M. M. Uddin, M. A. Islam Mamun and S. M. Shamsul Arefeen, "Health Risks and Disease Transmission in Undocumented Immigrants in the U.S Using Predictive ML," *2024 Sixth International Conference on Intelligent Computing in Data Sciences (ICDS)*, Marrakech, Morocco, 2024, pp. 1-6, doi: 10.1109/ICDS62089.2024.10756308.
[36] Abir, Shake Ibna, Richard Schugart, (2024). Parameter Estimation for Stroke Patients Using Brain CT Perfusion Imaging with Deep Temporal Convolutional Neural Network, Masters Theses & Specialist Projects, Paper 3755.
[37] Sohail, M. N., Ren, J., Muhammad, M. U., Rizwan, T., Iqbal, W., Abir, S. I., and Bilal, M, (2019). *Group covariates assessment on real life diabetes patients by fractional polynomials: a study based on logistic regression modeling*, Journal of Biotech Research, 10, 116-125.
[38] Sohail, M. N., Jiadong, R., Irshad, M., Uba, M. M., and Abir, S. I, (2018). *Data mining techniques for Medical Growth: A Contribution of Researcher reviews*, Int. J. Comput. Sci. Netw. Secur, 18, 5-10.
[39] Sohail, M. N., Ren, J. D., Uba, M. M., Irshad, M. I., Musavir, B., Abir, S. I., and Anthony, J. V, (2018). *Why only data mining? a pilot study on inadequacy and domination of data mining technology*, Int. J. Recent Sci. Res, 9(10), 29066-29073.