Journal of Computer Science

Research Article

Empirical Study on Modified Pre-Trained CNN Architectures for Fitzpatrick17k Skin Diseases Prediction Modelling

[1]Steven Matthew, [1]Nur Afny Catur Andryani, [1]Rojali, [2,3]Srie P. Gondokaryono, [3,4]Dedianto Hidajat and [3]July I. Rahardja

[1]Department of Computer Science, Bina Nusantara University, Jakarta, Indonesia
[2]Department of Dermatology and Venereology, Faculty of Medicine, Padjadjaran University, Indonesia
[3]Indonesian Society of Dermatology & Venerology, Indonesia
[4]Department of Dermatology and Venereology, Faculty of Medicine and Health Science, University of Mataram, Indonesia

Article history
Received: 10-11-2024
Revised: 13-01-2025
Accepted: 09-05-2025

Corresponding Author:
Steven Matthew
Department of Computer Science, Bina Nusantara University, Jakarta, Indonesia
Email: steven.matthew@binus.ac.id

Abstract: Preventing missed diagnosis of skin diseases is critical for enhancing patient outcomes. Technology support empowered by AI based algorithms have been recently developed to minimize the burden. Convolutional Neural Networks (CNNs) are increasingly utilized for image classification tasks including skin's diseases prediction modelling. Its advancement architectures such as pre-trained models, VGG16 and ResNet50 provide a strong foundation to up-skill the prediction capability. However, it takes higher resources. This research proposed empirical study of pre-trained modified CNN's architecture in handling skin's diseases identification which include the challenging open skin's diseases dataset, Fitzpatrick 17k as the task. The proposed modified CNN is pre-trained by using large skin diseases dataset ISIC 2019. The study involves the performance evaluation of pre-trained model and modified CNN architecture for classifying skin diseases using the Fitzpatrick17k dataset, which includes a diverse representation of skin tones and conditions. Simulation's findings demonstrate that the proposed pre-trained modified CNN architecture improved performance by up to 6% in accuracy compared to the baseline VGG16 model. While freezing specific layers enhanced model accuracy, this approach introduced trade-offs, such as a decrease in precision, which should be addressed in future research to optimize overall model performance. In addition, this research elaborates discussion on overfitting issue in handling Fitzpatrick17k-C dataset.

Keywords: Modified CNN, Skin's Diseases, Fitzpatrick17k, Fitzpatrick17k-C, Transfer Learning, Pre-Trained Model

Introduction

Skin diseases are a significant public health issue that can lead to death (Fauziyyah *et al.*, 2023) and early detection through automated image analysis could improve patient outcomes by facilitating timely intervention. Convolutional Neural Networks (CNNs) have shown impressive results in various image classification tasks and are increasingly used in dermatology for diagnosing skin diseases (Cano *et al.*, 2021); (Arun Prakash *et al.*, 2023); (Islam *et al.*, 2023). However, while pre-trained models like VGG16 and ResNet50 offer a strong starting point, recent literature suggests that modified CNN architectures (Rahman *et al.*, 2022); (Nigat *et al.*, 2023), specifically tailored and fine-tuned for a specific dataset, can yield better performance than generic pre-trained models when applied to unique and specialized datasets like Fitzpatrick17k.

While many studies using the Fitzpatrick17k dataset have focused on skin tone analysis (Chiu *et al.*, 2023) and benign-malignant classification (Almuzaini *et al.*, 2023); (Daneshjou *et al.*, 2022), few have explored the direct prediction of specific skin diseases.

In this study, we evaluate the effectiveness of modifying a pre-trained CNN model to improve performance on the Fitzpatrick17k dataset (Groh *et al.*, 2021). The aim of proposing evaluation on pre-trained modified CNN is improving prediction ability while at the same time retaining less complexity. Hence, it is feasible to deploy in such low-resource devices. In addition, a discussion elaboration on the overfitting issue of handling the Fitzpatrick 17k-C dataset is also

presented. It contributes to potential future works on improving the generality of the prediction model.

Materials and Methods

In this section, the research materials and methodology are elaborated. It consists of the overall research methodology, the proposed modified CNN architecture, the simulation set, and an explanation of the datasets.

Research Methodology

The proposed works are started with a clear state of the art to address the challenges of building skin disease prediction modeling on the given Fitzpatrick17k Dataset. Hence, the research started with a literature review to clear up the research formulation and its appropriate addressed research works. The flow is illustrated in Figure (1).

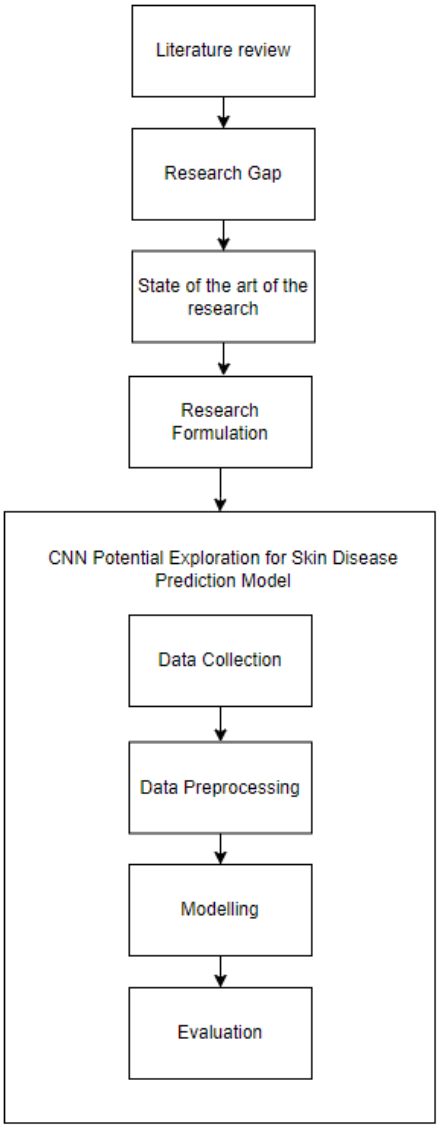

Fig. 1: Research architecture

Numerous studies have explored skin disease classification, ranging from traditional machine learning approaches (Andryani *et al.*, 2023); (Ravi, 2022) to deep learning methods (Cano *et al.*, 2021); (Nigat *et al.*, 2023); (Ravi, 2022); (Cai *et al.*, 2023). Deep learning research commonly utilizes models pretrained on the ImageNet dataset. However, recent studies (Rahman *et al.*, 2022); (Nigat *et al.*, 2023) have shown that tailoring CNN models for specific datasets can yield better performance and often result in smaller model sizes. In addition, in terms of testing images, most of the previous researches are working on ISIC (Cano *et al.*, 2021); (Andryani *et al.*, 2023); (Ravi, 2022); (Cai *et al.*, 2023); (Wang *et al.*, 2023); (Jeong *et al.*, 2023). Recently, there has been a mobile phone-based skin diseases dataset released, Fitzpatrick17k, which includes a more challenging image to infer. This proposed research is focusing on handling this particular dataset.

Pretrained models are widely used in medical imaging due to their ability to quickly adapt to new datasets, thanks to prior knowledge and because they reduce training costs (Iman *et al.*, 2023); (Salehi *et al.*, 2023).

Pretraining on ImageNet requires substantial computational resources due to its size of over 1 million images. Therefore, in this study, we chose to pretrain our model on the ISIC 2019 dataset, which is more relevant to our final dataset and significantly less computationally demanding. It is also intended to result in better weight initialization in modeling the Fitzpatrick dataset.

Evaluation of pre-trained CNN architecture dealing with skin disease prediction modeling on the Fitzpatrick 17k dataset was executed by using classification performance metrics and learning curves.

Evaluated Modified CNN Architecture

The modified CNN that is proposed is a modified version of the model proposed by Vulpe-Grigoraşi. This model is also used as a benchmark to evaluate the performance of the proposed model. It is pre-trained CNN architecture by using the ISIC 2019 dataset. Its additional enhancement is also on the feature extraction process which was executed by adding an additional convolutional layer, max pooling, batch_normalization, and dropout layer and replacing the flattened layer with a GlobalAveragePooling2D layer to improve performance.. A random search algorithm was implemented to find the optimal filter sizes for each convolutional layer and the ideal number of units for the dense layers in the modified CNN. In the current simulation, the detail of the proposed modified CNN architecture is illustrated in Figure (2).

GlobalAveragePooling2D (GAP) is used to replace the flattened layer because of its superior performance, GAP resulted in a smaller network because of the averaging process instead of just reshaping the feature

map. Because of the smaller feature map, GAP can also give the benefit of the regularization effect (Lin, 2013).

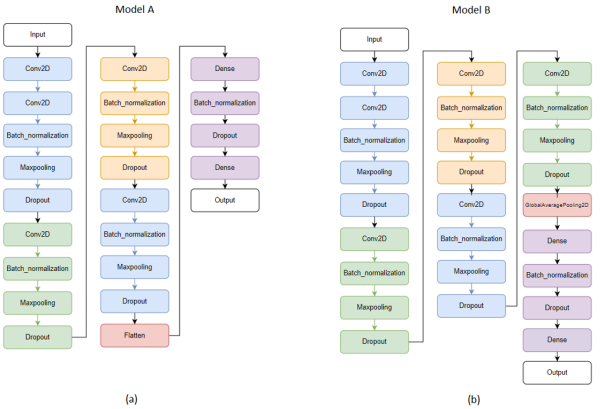

Fig. 2: The modified CNN Architecture (a) Vulpe-Grigorasi, (b) Proposed Pre-trained Modified CNN Architecture

In the proposed modified CNN architecture, each added hidden layer comprises a convolutional layer, max-pooling, batch normalization, and dropout, forming a modular structure that enhances the network's feature extraction and regularization capabilities. Adding an additional convolutional layer allows the model to capture more complex and abstract features from the data, progressing from detecting low-level patterns like edges to high-level abstractions such as shapes or textures as depth increases (Simonyan, 2014). Following each convolutional layer, we employ max-pooling to downsample feature maps, effectively capturing the most prominent features and enhancing translation invariance. This process reduces spatial dimensions, optimizing the model's computational efficiency and contributing to the network's robustness (Sermanet, 2013).

To improve stability and convergence speed, each hidden layer also includes batch normalization. This normalizes the input to each layer, reducing internal covariate shifts, which allows for faster convergence during training. Batch normalization also acts as a regularizer, supporting the network's generalization to new data (Ioffe and Szegedy, 2015). Finally, we integrate dropout after batch normalization, where a fraction of neurons are randomly deactivated during training. This method prevents overfitting by ensuring that neurons do not become overly specialized to specific features, (Srivastava *et al.*, 2014) thereby fostering a more generalized model.

Together, these elements—convolutional filtering, pooling, normalization, and dropout—create a balanced architecture that maintains complexity in feature extraction while effectively managing overfitting risks.

For pretraining, we used the ISIC 2019 dataset on the same model architecture, with the only difference being the number of classes—8 for ISIC 2019 as opposed to the 16 classes in Fitzpatrick17k-C. The model was trained for 50 epochs using the Adam optimizer with a learning rate of 0.001. Consistent data augmentation techniques were applied across all datasets, including horizontal flipping, a 10-degree rotation, and 10% zoom augmentation.

About Dataset

The Fitzpatrick17k-C dataset, developed by Abhishek, K. in 2024, is a comprehensive skin lesion dataset that enhances the original Fitzpatrick17k dataset which was initially developed by Groh *et al.* through critical modifications to improve usability and reliability for machine learning applications. Notably, Fitzpatrick17k-C addresses two significant issues: data duplication and standardized data splitting. The duplication process removes redundant images found in the original dataset, resulting in a cleaner and more efficient set for model training and evaluation. Additionally, the dataset incorporates a carefully designed split of training and testing sets, preventing data leakage and enabling consistent performance comparisons across research studies. By ensuring that images from the same patient or similar sources do not overlap between training and testing, Fitzpatrick17k-C supports more reliable evaluations, reducing the variability in performance metrics typically introduced when datasets are split differently by individual researchers. This improvement enhances the integrity of the dataset and the reliability of model comparisons across studies.

Fitzpatrick17k-C comprises 11,394 images across 114 skin condition classes. For the purposes of our research, we narrowed down this set to include 16 skin conditions specified in the Standard Kompetensi Dokter Indonesia (SKDI). This focused selection aligns with medical standards in Indonesia and yields a total of 2,535 images. The selected classes include acne vulgaris, allergic contact dermatitis, drug eruption, folliculitis, ichthyosis vulgaris, lichen simplex, perioral dermatitis, pityriasis rosea, psoriasis, scabies, seborrheic dermatitis, Stevens-Johnson syndrome, sun-damaged skin, tick bite, urticaria, and vitiligo.

Unlike many standard datasets, Fitzpatrick17k-C does not enforce a specific resolution for its images; however, many of the images exceed the size of 224x224 pixels. This offers flexibility and potential for capturing finer details, though it necessitates preprocessing for tasks that require standardized input sizes. In summary, the deduplication, data leakage prevention, and class diversity of Fitzpatrick17k-C make it a valuable resource for dermatological research, while our specific selection of 16 classes allows for focused application aligned with Indonesian medical competencies.

ISIC 2019 was also used to pretrain the model. The dataset contains 8 skin lesions with a total image of 25.331. The high resolution of these dermoscopic images —sized at 224x224 pixels and in full color—enhances

the dataset's value for accurate image-based diagnosis. Dermoscopy is a critical imaging modality in dermatology as it captures detailed, close-up views of skin lesions, making it an ideal data source for training deep learning models. The fine-grained structure captured in dermoscopic images, including pigmentation patterns and lesion borders, supports nuanced classification.

Simulation Set Up

To demonstrate the performance of the proposed architecture, a simulation scenario is set up. It is described in Table (1).

Table 1: Simulation setup for skin lesion classification

Controlled Variables	Configuration
Dataset	Modelling Dataset: Fitzpatrick17k-C, Pre-trained Model Dataset: ISIC2019
Image resolution	Resized to 224×224
Data augmentation	Horizontal flipping, 10° rotation, 10% zoom
ISIC 2019 split ratio	Training: 80%, Validation: 10%, Testing 10%
Training Epochs	Fitzpatrick17k-C: 150 epochs, ISIC 2019 (Pretraining): 50epochs
Optimizer	Adam optimizer with a learning rate of 0.001
Fine-tuning strategy	Incrementally frozen convolutional layer, starting from none until all convolutional layers are frozen.
Performance Metrics	Accuracy, Precision, Recall, F1-Score

The preprocessing for the Fitzpatrick17k-C dataset involved resizing all images to 224×224 pixels to match the ISIC 2019 dataset resolution, ensuring consistency and simplifying model training then the ISIC 2019 dataset was split into training, validation, and testing sets with an 80:10:10 ratio. For both datasets, identical data augmentation techniques were applied to the training set to improve performance, including horizontal flipping, 10 degrees of rotation, and 10% zoom.

For training parameters on the Fitzpatrick17k-C dataset, all models were trained for 150 epochs to ensure the model already reached the optimal performance. For pretraining on ISIC 2019, the model was trained for 50 epochs.

Both pretraining and fine-tuning were done using the Adam optimizer with a learning rate of 0.001.

In the fine-tuning stage, we explored all possible configurations of freezing convolutional layers. Starting with no frozen layers, we sequentially froze layers—from only the first convolutional layer up to all convolutional layers—to observe how layer freezing impacts model performance. The overall training procedure is summarized in the pseudocode shown in Figure (3), providing a high-level view of the implementation logic.

```
Input: Dataset Fitzpatrick17k-C
Output: Fine tuning model
1.Start:
2.  X_train, Y_train <- LoadDataset(Fitzpatrick17-C_train)
3.  X_val, Y_val <- LoadDataset(Fitzpatrick17-C_validation)
4.  X_test, Y_test <- LoadDataset(Fitzpatrick17-C_test)
5.  X_train_pre <- PreprocessData(X_train)
6.  X_val_pre <- PreprocessData(X_val)
7.  X_test_pre <- PreprocessData(X_test)
8.  X_train_aug <- AugmentationImage(X_train_pre)
9.  base_model <- LoadModel(pretrained_model)
10. base_model_freeze <- SetFrozenLayer()
11. parameter <- LoadParameters()
12. model <- CompileModel(base_model_freeze, parameter)
13. callback <- [Checkpoint(parameter,checkpoint_path)]
14. TrainModel[model, X_train_aug,Y_train,X_val,Y_val,callback,parameter]
15. best_model <- LoadWeight(checkpoint_path)
16. EvaluateModel(best_model,X_test,Y_test)
17. End
```

Fig. 3: Pseudocode for model training

Performance was measured using Accuracy, Precision, Recall and F1-Score. Accuracy represents the percentage of all predictions that were correct, measuring the overall correctness of the model. Precision represents the proportion of correctly predicted positive cases out of all cases predicted as positive, indicating how often positive predictions are accurate, while Recall represents the proportion of actual positive cases that were correctly identified by the model, showing the model's ability to capture all true positives. The F1-Score provides a balanced metric by combining Precision and Recall.

All experiments were conducted on a Windows 10 system using WSL2, running Python 3.11.7 and TensorFlow 2.15.0 on a Nvidia 3080Ti GPU.

Results and Discussion

All the necessary classification performance metrics and learning curves are used to quantitatively evaluate the model. It is presented in Tables (2-4) and Figs. (4-8).

Table 2: Performance metrics for each model on the Fitzpatrick17-C dataset (in percentage)

Model	Acc	Precision	Recall	F1
Vulpe Model	22	17	17.31	14.5
Resnet50V2	42.35	45.87	28.76	30.37
VGG16	39.22	38.89	27.95	29.71
CNN without Pretrained	40.59	34.55	29.42	29.12
CNN Pretrained without Freezing	42.55	43.04	36.72	37.06
Freeze conv2d	40.78	34.73	31.53	30.53
Freeze conv2d_1	44.71	42.96	35.84	36.09
Freeze conv2d_2	44.71	43.59	37.65	38.15
Freeze conv2d_3	45.1	42.7	38.2	38.53
Freeze conv2d_4	44.12	44.96	36.97	37.84
Freeze conv2d_5	36.67	37.32	30.67	30.57

Table 3: Performance metrics for each model on the Fitzpatrick17k dataset (in percentage)

Model	Acc	Precision	Recall	F1
Resnet50	46	49	47	47
InceptionV3	49	51	49	47
Our Model	51	51	51	50

Table 4: Model size and parameters for each model

Model	Size	Total parameter
Vulpe	9.47 MB	2.482.472
VGG16	57.66 MB	15.116.112
Resnet50V2	91.12 MB	23.887.664
Modified CNN	2.26 MB	593.600

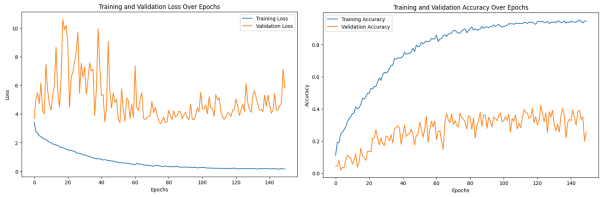

Fig. 4: Performance Learning Curve for Vulpe Model

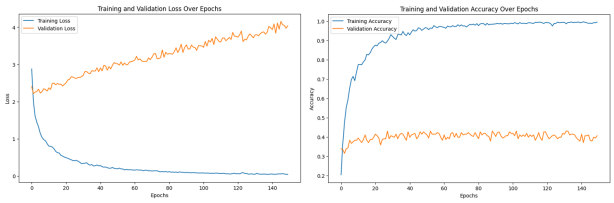

Fig. 5: Performance Learning Curve for VGG16

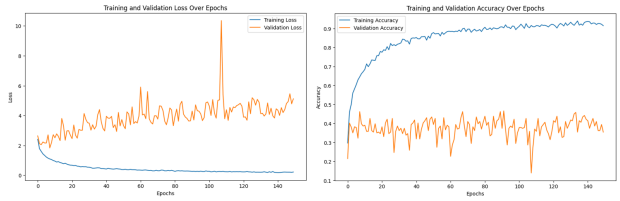

Fig. 6: Performance Learning Curve for Resnet50V2

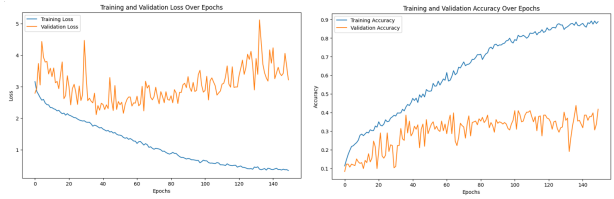

Fig. 7: Performance Learning Curve for Modified CNN without pre-trained

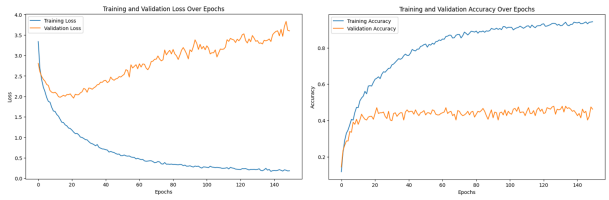

Fig. 8: Performance Learning Curve for Modified CNN with pre-training

The best performance was achieved by freezing the first convolutional layer up to conv2d_3. This improvement is attributed to the fact that the pretrained weights in the early layers tend to lose their precision when not frozen, especially due to the noisy data from the Fitzpatrick17k-c dataset. However, when all convolutional layers were frozen, performance significantly declined, as the model was unable to adapt to the new dataset. Pretraining led to significant improvements in the recall of most classes, with notable gains such as a 16% increase in Folliculitis recall, 20% in Lichen simplex, 36% in Ichthyosis Vulgaris, 40% in Pityriasis rosea, and 45% in Urticaria. However, while freezing conv2d_2 resulted in slightly lower overall performance, it still managed to classify all classes successfully.

Although the average accuracy is slightly lower than that of conv2d_3, the recall performance showed improvement for 10 classes, achieving better or comparable recall rates. Furthermore, pretraining resulted in lower loss values and enhanced training stability, with significantly reduced fluctuations during each training epoch. When using randomly initialized weights, the validation accuracy plateaued after 40 epochs; in contrast, with transfer learning, the performance reached a plateau after just 10 epochs.

Compared to the original Fitzpatrick17k dataset benchmarks, our model achieved superior performance than another model, reaching an accuracy and recall of 51%. This demonstrates our model's effective adaptation to the dataset, achieving higher reliability in both overall and class-specific predictions. In this research, our model not only outperformed other pretrained networks like VGG16 and ResNet50V2 in terms of classification performance but also proved to be the most computationally efficient. With a model size of only 2.26 MB and a total of 593,600 parameters, our Modified CNN achieves top results while maintaining a low computational cost compared to the larger, more resource-intensive VGG16 and ResNet50V2 networks.

Overfitting Issues

Fitzpatrick17k will be released in early 2021 by Groh *et al.* It consists of 114 classes of diseases with various skin-based conditions. The images are taken using mobile phones which have various backgrounds and less lesion detail compared to dermoscopy images. Hence, this dataset is very challenging to model. This study reveals that overfitting exists in all employed and proposed architectures. It is described in Figs. (4-8). There is a significant gap between the training and validation performance which indicates the overfitting. It is earlier predicted because of the dataset complexity.

Discussion

Transferred Learning Performance

The results of our study reveal important insights into the challenges and potential of transfer learning for skin disease classification using the Fitzpatrick17k-c dataset. Initially, the proposed pre-trained modified CNN model yielded a modest accuracy of 40.59% on the

Fitzpatrick17k-c dataset, with particularly low recall rates (under 10%) for classes like Ichthyosis Vulgaris, Pityriasis rosea, Sun-damaged skin and Urticaria. This indicates that despite the modified architecture, the model struggled to generalize well due to the diverse and sometimes low-resolution nature of the dataset.

Pretraining the model on the ISIC 2019 dataset proved beneficial, leading to a notable improvement in overall accuracy and recall across most classes. By experimenting with different layer-freezing configurations, we found that freezing the early layers, up to conv2d_3, yielded the best performance. This finding aligns with research suggesting that freezing the initial layers helps preserve essential features learned from the source dataset (ISIC 2019) while allowing the model to adapt higher-level layers to the target data (Fitzpatrick17k-C). Interestingly, while freezing up to conv2d_2 resulted in slightly lower accuracy than conv2d_3, it enhanced recall for 10 classes, suggesting that freezing fewer layers may help retain information relevant to fine-grained classification.

However, when all convolutional layers were frozen, performance dropped considerably. This observation underscores the importance of adapting deeper layers to the specific characteristics of Fitzpatrick17k-C; without this adaptation, the model could not adequately capture the variability in this dataset, which includes skin lesions of varied texture and context.

A class-level analysis of recall rates post-pretraining revealed marked improvements for several classes. For example, recall increased by 16% for Folliculitis, 20% for Lichen simplex, 36% for Ichthyosis Vulgaris, 40% for Pityriasis rosea, and 45% for Urticaria, indicating the transfer learning approach’s value for previously underperforming classes. Additionally, pretraining led to more stable training with lower validation loss and convergence after 10 epochs, whereas randomly initialized models plateaued around 40 epochs, further highlighting the stability benefits of transfer learning.

To contextualize our model's performance, we compared it with results from pre-trained models on the original Fitzpatrick17k dataset. Our modified model demonstrated superior performance across all metrics, as shown in Table (3). This performance difference reinforces the advantage of data-specific models and highlights the benefits of the enhanced dataset structure in Fitzpatrick17k-c, which includes deduplication and standard data splitting. The primary limitation of this study was the dataset's low lesion resolution, which affected the model's ability to capture detailed features of the skin lesions. Future research could explore approaches like super-resolution techniques or lesion-focused cropping to enhance lesion visibility. Additionally, further exploration of layer-freezing strategies could reveal additional combinations that yield optimal results. Lastly, we recommend assessing the model's generalizability on external datasets to verify the stability of performance improvements across varied conditions.

Addressing Overfitting Issues

It can be inferred from the above learning curves that all the learning algorithms working on this dataset are facing the same overfitting issues. Beyond the challenge of the datasets itself, pre-processing data might help to explain the case. Instead of simple geometric transformation-based augmentation, such as flipping, rotation, and resizing, it might not enrich the required information to learn. On the opposite, it potentially supplies redundant information that causes weak generalization. Hence, more advanced image augmentation or other pre-processing approaches that address some aspects below are recommended for future development:

- Ensuring data distribution on the data augmentation is significant to avoid data drift during the training testing procedure. The significant bias of the model which leads to overfitting might occur due to data augmentation which ignores the data distribution aspect. Hence, domain adaptation-based image augmentation might be beneficial to address this issue.
- The intention of image augmentation is to enforce the information enrichment on improving the generalization ability of the model hence overfitting can be reduced. More advanced image augmentation methods such as SalfMix (Choi *et al.*, 2021) and KeepAugment (Gong *et al.*, 2021), which take more localized considerations on producing augmented data are promising to address this issue.
- One of the main issues in developing prediction modeling on images or vision is information localization. The intended localized area which significantly contributes to the recognition model might be mostly disturbed by the other additional information coming from the background. Hence, if possible appropriate annotation should be enforced. However, annotation might be time-consuming and costly. The combination of the few-shot learning framework which only requires less annotated data, might be beneficial.
- The images taken in Fitzpatrick17k are mobile phone images that lack lighting sensitivity. It might ruin the data distribution in the intra-class images. It could be corrected by the D65 lighting standard. Before the image augmentation is executed, lighting correction into the D65 lighting standard is executed. It promotes a uniform standard of lighting and hence will retain the core feature information for each skin disease more consistently.

Conclusion

This study highlights the challenges of adopting CNNs for skin disease classification on challenging Fitzpatrick 17k-C datasets and demonstrates the benefit of transfer learning and tailored layer-freezing techniques. The proposed pre-trained modified CNN model aims to explore the potential of less complex architecture that is able to learn the skin's diseases better. The simulation shows, that the proposed model succeeds in improving the accuracy in less complex architecture compared to pre-trained models such as ResNet50V2, VGG16, and Vulpe model. However, as presented in the predecessor proposed model by other researchers, the model is still facing overfitting in handling the particular Fitzpatrick17k-C dataset. Regarding the overfitting issue, the study explores discussion on several aspects that aim to reduce overfitting in future works. Ensuring data distribution and information enrichment are some consideration keywords for addressing potential image augmentation method approach to providing more data on the modeling process. These two considerations are addressed to achieve the generalization ability of prediction modeling. In addition, to improve the modeling performance, two aspects of data pre-processing on handling localized regions of interest and lighting correction are also presented. The insight into the pre-processing data direction is beneficial to improve the prediction modeling performance in future works.

Acknowledgment

This research is a part of the research project" AI-based pediatric Tele dermatology development in Indonesia" collaboration of Computer Science BINUS University, Pusat Riset Komputasi BRIN (National Research and Innovation Agency) and Kelompok Studi Dermatologi Anak PERDOSKI (Indonesian Society of Dermatology and Venereology).

Funding Information

The research is funded by BINUS Research funding and RIIM BRIN 2024-2025.

Author's Contributions

Steven Matthew: Conducting literature Review, conducting simulation, and preparing the manuscript.

Nur Afny Catur Andryani: Designing research, Designing the state of the art, Designing the problem-solving, Analysis, and Revising manuscript.

Rojali: Review manuscript.

Srie P. Gondokaryono, July I. Rahardja and Dedianto Hidajat: Domain Expertise on validating the skin's diseases case and dataset.

Ethics

This manuscript is an original work. The corresponding author declares that no ethical concerns are associated with this submission. The datasets are open datasets which legally published.

Conflicts of Interest

The authors have no competing interests to declare relevant to this article's content.

References

Almuzaini, A. A., Dendukuri, S. K., & Singh, V. K. (2023). Toward Fairness Across Skin Tones in Dermatological Image Processing. *2023 IEEE 6th International Conference on Multimedia Information Processing and Retrieval (MIPR)*, 1-7. https://doi.org/10.1109/mipr59079.2023.00030

Andryani, N. A. C., Juwono, F. H., & Sianturi, P. J. (2023). Artificial intelligence approach for automatic multiclass skin diseases identification. *Journal of Engineering Science and Technology*, *18*(4), 99-111.

Arun Prakash, J., Asswin, C., Ravi, V., Sowmya, V., & Soman, K. (2023). Pediatric pneumonia diagnosis using stacked ensemble learning on multi-model deep CNN architectures. *Multimedia Tools and Applications*, *82*(14), 21311-21351. https://doi.org/10.1007/s11042-022-13844-6

Cai, G., Zhu, Y., Wu, Y., Jiang, X., Ye, J., & Yang, D. (2023). A multimodal transformer to fuse images and metadata for skin disease classification. *The Visual Computer*, *39*(7), 2781-2793. https://doi.org/10.1007/s00371-022-02492-4

Cano, E., Mendoza-Avilés, J., Areiza, M., Guerra, N., Mendoza-Valdés, J. L., & Rovetto, C. A. (2021). Multi skin lesions classification using fine-tuning and data-augmentation applying NASNet. *PeerJ Computer Science*, *7*, e371. https://doi.org/10.7717/peerj-cs.371

Chiu, C.-H., Chung, H.-W., Chen, Y.-J., Shi, Y., & Ho, T.-Y. (2023). Toward Fairness Through Fair Multi-Exit Framework for Dermatological Disease Diagnosis. *Medical Image Computing and Computer Assisted Intervention - MICCAI 2023. MICCAI 2023. Lecture Notes in Computer Science*, 97-107. https://doi.org/10.1007/978-3-031-43898-1_10

Choi, J., Lee, C., Lee, D., & Jung, H. (2021). SalfMix: A Novel Single Image-Based Data Augmentation Technique Using a Saliency Map. *Sensors*, *21*(24), 8444. https://doi.org/10.3390/s21248444

Daneshjou, R., Yuksekgonul, M., Cai, Z. R., Novoa, R., & Zou, J. Y. (2022). Skincon: A skin disease dataset densely annotated by domain experts for fine-grained debugging and analysis. *Advances in Neural Information Processing Systems*, 18157-18167.

Fauziyyah, R. N. P., Komariah, M., & Herliani, Y. K. (2023). Sunlight Exposure and Protection Behavior as Prevention of Skin Cancer in Nursing Students. *Indonesian Journal of Cancer*, *17*(1), 1. https://doi.org/10.33371/ijoc.v17i1.921

Gong, C., Wang, D., Li, M., Chandra, V., & Liu, Q. (2021). KeepAugment: A Simple Information-Preserving Data Augmentation Approach. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 1055-1046. https://doi.org/10.1109/cvpr46437.2021.00111

Groh, M., Harris, C., Soenksen, L., Lau, F., Han, R., Kim, A., Koochek, A., & Badri, O. (2021). Evaluating Deep Neural Networks Trained on Clinical Images in Dermatology with the Fitzpatrick 17k Dataset. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1820-1828. https://doi.org/10.1109/cvprw53098.2021.00201

Iman, M., Arabnia, H. R., & Rasheed, K. (2023). A Review of Deep Transfer Learning and Recent Advancements. *Technologies*, *11*(2), 40. https://doi.org/10.3390/technologies11020040

Ioffe, S., & Szegedy, C. (2015). Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *Proceedings of the 32nd International Conference on Machine Learning*, 448-456.

Islam, M., Reza, Md. T., Kaosar, M., & Parvez, M. Z. (2023). Effectiveness of Federated Learning and CNN Ensemble Architectures for Identifying Brain Tumors Using MRI Images. *Neural Processing Letters*, *55*(4), 3779-3809. https://doi.org/10.1007/s11063-022-11014-1

Jeong, H. K., Park, C., Henao, R., & Kheterpal, M. (2023). Deep Learning in Dermatology: A Systematic Review of Current Approaches, Outcomes, and Limitations. *JID Innovations*, *3*(1), 100150. https://doi.org/10.1016/j.xjidi.2022.100150

Lin, M., Chen, Q., & Yan, S. (2013). Network in network. *ArXiv:1312.4400v3*. https://doi.org/10.48550/arXiv.1312.4400

Nigat, T. D., Sitote, T. M., & Gedefaw, B. M. (2023). Fungal Skin Disease Classification Using the Convolutional Neural Network. *Journal of Healthcare Engineering*, *2023*(1), 370416,. https://doi.org/10.1155/2023/6370416

Rahman, K., Amin, T. B., Rahman, M. S., & Sakib, G. (2022). *An efficient approach to detect melanoma skin cancer using a custom CNN model*.

Ravi, V. (2022). Attention Cost-Sensitive Deep Learning-Based Approach for Skin Cancer Detection and Classification. *Cancers*, *14*(23), 5872. https://doi.org/10.3390/cancers14235872

Salehi, A. W., Khan, S., Gupta, G., Alabduallah, B. I., Almjally, A., Alsolai, H., Siddiqui, T., & Mellit, A. (2023). A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope. *Sustainability*, *15*(7), 5930. https://doi.org/10.3390/su15075930

Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., & LeCun, Y. (2013). Overfeat: Integrated recognition, localization and detection using convolutional networks. *ArXiv:1312.6229v4*. https://doi.org/10.48550/arXiv.1312.6229

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *ArXiv*, 1409.1556. https://doi.org/10.48550/arXiv.1409.1556

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, *15*(1), 1929-1958.

Wang, Y., Su, J., Xu, Q., & Zhong, Y. (2023). A Collaborative Learning Model for Skin Lesion Segmentation and Classification. *Diagnostics*, *13*(5), 912. https://doi.org/10.3390/diagnostics13050912