Vol. 9. No. 1 (2026)
Page: 9582-9590

THE L2-EFFNET FOR HUMAN SKIN DISEASE MULTICLASS CLASSIFICATION

Kurnia Adi Cahyanto[1]
Universitas Diponegoro, Semarang, Indonesia
kelixo@gmail.com

Kusworo Adi[2]
Universitas Diponegoro, Semarang, Indonesia
kusworoadi@lecturer.undip.ac.id

Catur Edi Widodo[3]
Universitas Diponegoro, Semarang, Indonesia
catur.ediwidodo@gmail.com

Abstract

Skin diseases are common health problems that are often underestimated as most are mild and can be treated with over-the-counter medications, however, some types such as melanoma can be cancerous and deadly if not treated properly. Melanoma is caused by overexposure to ultraviolet light and has a 99% cure rate if diagnosed early, but decreases to 20% in advanced stages. In Indonesia, the uneven distribution of medical personnel and geographical challenges cause many cases of skin diseases not to be diagnosed well. This study develops a multi-class skin disease classification model using transfer learning by utilizing pre-trained models such as EfficientNetV2S to overcome data imbalance and improve accuracy. The dataset used is a combination of ISIC and Atlas Dermatology, which has been curated into 31 classes with a total of 3,399 samples, and data augmentation was performed to increase the sample size. The modified L2-EfficientNetV2S model showed a testing accuracy of 88.33%, higher than previous studies, demonstrating the potential of using deep learning in the early diagnosis of skin diseases to improve healthcare.

Keywords: Index Terms—Skin Disease, Diagnosis, EfficienNetV2S, Accuracy, Health

INTRODUCTION

Skin diseases are common in society, as they result from health-related effects. They are often overlooked, as, most are relatively harmless and can usually be treated with mild prescriptions from the pharmacy. However, there are some types that are sometimes cancerous and can be deadly if not treated properly (Khan et al., 2021).

These skin disease problems can be temporary or permanent and may be pleasant or unpleasant. Some are the result of environmental influences, while others may be inherited. Some skin conditions are mild, while others can be fatal and cancerous (Rafay & Hussain, 2023). One of the causes is excessive exposure to ultraviolet (UV) light on the skin which causes cell mutations in the skin and disruption of melanocytes. This type of melanoma has a good cure rate if found early, with a 5-year relative survival rate of 99%. However, as it spreads faster to other parts of the body than non-cancerous skin cancers, the 5-year relative survival rate decreases to 20% in the long-term stage.

In all regions of Indonesia, this is a disease that is often encountered and tends to be contagious so not a few polyclinics or health facilities are dominated by the emergence of this skin disease (Purnama et al., 2019), where patients are spread across various regions and islands in Indonesia. Community health centers have been established throughout Indonesia to provide health services to patients, including those with skin diseases. Unfortunately, medical personnel in these centers are not evenly distributed and tend to be concentrated in big cities or provinces. Also, as an archipelago, transportation in Indonesia tends to be difficult to reach remote areas. As such, more serious types often go untreated for long periods of time due to a lack of understanding of therapeutic procedures and available medical support (Setiawan et al., 2021).

Figure. 1.
Examples of types of skin diseases: Melanoma (left) and Basal Cell Carcinoma (right)

Given the above facts, early diagnosis of the disease is needed to improve the survival rate (Li et al., 2021). An alternative is required to aid in the early detection of skin diseases because medical capabilities can occasionally be limited. In recent years, the influence of machine learning using computers has become a necessity that spans all fields (Shanthi et al., 2020), including the medical field, especially those related to skin health or dermatology. Dermatology is a bioscience discipline concerned with the diagnosis and treatment of skin problems (Jeong et al., 2023). A wide range of dermatologic diseases change regionally and seasonally according to temperature, humidity, and other environmental conditions.

By using technology to assist in diagnosis, physicians can see more patients and make fewer mistakes in their diagnoses. Among these is deep learning, which uses convolutional neural networks (CNNs) to classify different forms of skin diseases based on visual images of the affected area (Karthik et al., 2022). The goal is to identify the patient's skin disease type early on (Al-masni et al., 2020). Pre-existing or pre-trained models are leveraged in deep learning, particularly in transfer learning, to handle picture data (Sadik et al., 2023). The output of these pre-existing models can be altered and tailored to certain image recognition needs.

In addition, often with multi-class skin diseases that are usually unbalanced (Calderón et al., 2021), sometimes the training accuracy results are much different from the accuracy when applied to test data. This is what causes overfitting (Soumare et al., 2021). This is what causes overfitting. Therefore, this research aims to handle the imbalance of data and improve the accuracy of previous research in classifying multiclass skin diseases.

REVIEW OF LITERATURE

There have been many studies on machines to classify skin diseases, but there is not much research on multiclass-related searches. Therefore, we only take related searches that have relatively many classes. In 2021, in their review paper, (Li et al., 2021), with a dataset from Dermnet consisting of 23 classes, using pre-trained VGG-16, VGG-19, and GoogleNet gave a top-1 accuracy of 73%. This review discusses future challenges to be able to better classify human skin diseases, including overfitting, imbalance of human skin disease data, and noisy data caused by different types of cameras with different resolutions so that heterogeneity, lack of diversity, lack of medical history and clinical meta-data of patients, clarity about methods, and selection of appropriate deep learning models.

Then according to (Rafay & Hussain, 2023), a combined dataset from ISIC (ISIC, 2019) and Atlas Dermatology (Silva, 2025) was used to eventually curate the dataset into 31 classes, of which 20 were used for testing and 80 for training. There are 994 test data and 3916 training data in this dataset. This training data is pre-segmented using Center Zoom, Rotation, Brightness, Shear, Vertical Flip, and Horizontal Flip during the preprocessing step. Using EfficientNet-B2 as the transfer learning and 10 epochs, the maximum accuracy for the test data is 87.15%. As of this article's conclusion, the author believes that accuracy can still be raised.

(Karthik et al., 2022), combined datasets from various sources with a total of 4930 images collected to represent four different skin diseases. To increase the diversity in the dataset, data augmentation was applied, and a total of 17.329 images were generated for model building and evaluation then observed the use of Efficient Channel Attention (ECA) on EfficientNetV2, but only for 4 classes with a testing accuracy of 84.70%. In another study, (Cahyanto et al., 2023), discussed the effect of L2-Regularization on improving the accuracy of test data for ISIC 2019 and Dermnet data.

RESEARCH METHOD

Data Collection

The dataset used was taken from research conducted by (Rafay & Hussain, 2023). This dataset is a combination of ISIC (ISIC, 2019) and Atlas Dermatology (Silva, 2025), which has been curated to ensure data richness and standardization.

There are 561 skin illnesses in the Atlas Dermatology dataset overall, yet some classes have insufficient sample sizes to adequately train the deep learning model. Consequently, a cutoff of 80 samples or more per class was established, yielding a dataset including 24 classes and 3.399 samples. The ISIC dataset included nine classes for skin diseases, but prior to the merger, the Atlas Dermatology dataset removed two of these classes. This led to the inclusion of seven additional classes in the dataset.

We applied data augmentation to the photos to address the issue of a small sample size. As a result, the curated dataset utilized in this work had 3.399 samples total, representing 31 different skin conditions. Additionally, we divided the dataset into 20% for the validation and test datasets, and 80% for the training dataset.

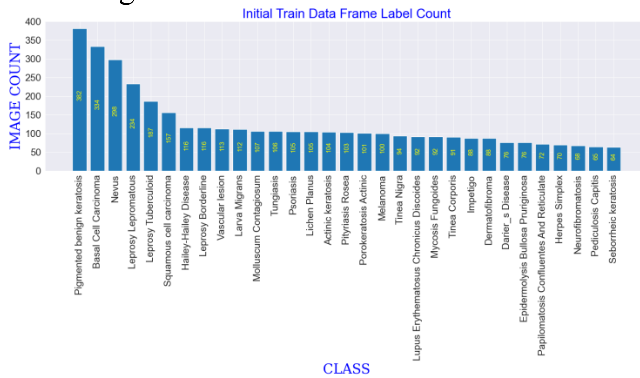

Figure. 2.
Training Dataset Distribution

Data Augmentation

Data augmentation scheme using 180^0 rotation, center crop 0.25, random brightness contrast 0.2, shear 0.5, and horizontal and vertical flip 0.5. The center crop is used because there are some images where the background takes up too large a part of the whole image, thus reducing the object information that is more necessary.

Figure. 3.
Data Augmentation Result

We perform augmentation on the training data, then equalize the amount of data in each class with the class that has the most data. In the end, the data from each class in the training data amounted to 382 data points.

The aim of this data augmentation is to enhance the diversity of the data's position and shape, while also augmenting its quantity, with the expectation that the model will gain more insight from specific patterns and contours

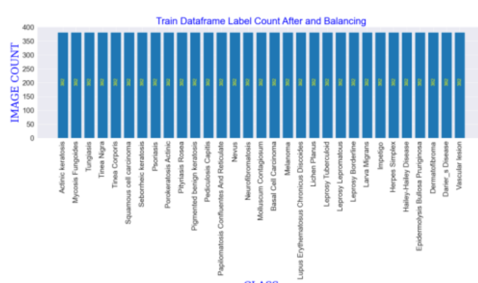

Figure 4.
Training Dataset Distribution after Trimming

To compare the extent of overfitting, our proposed framework uses two accuracy values, namely accuracy for training data and accuracy for test data.

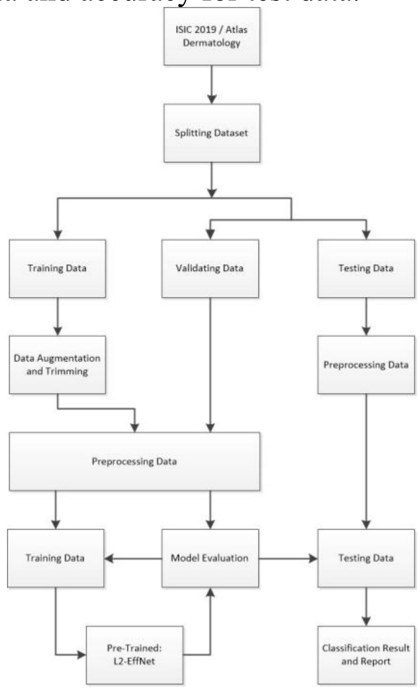

Figure 5.
Framework Method

L2-EffNet Model

This step completes the data modeling process to develop the classification model. We construct the model using the EfficientNetV2 architecture, a novel neural network scaling technique. With 11x faster training speed and a 6.8x smaller model size, the EfficientNetV2 model set of convolutional neural networks (CNN) offers better image categorization capabilities. EfficientNetV2 uses a smaller 3x3 kernel with multiple layers, fused-MBConv and fused-MBConv added at the first layer, as well as a mobile inverted bottleneck convolution (MBConv) block with a smaller expansion ratio (Tan & Le, 2021).

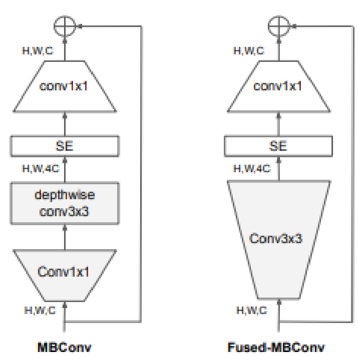

Figure 6.
Structure of MBConv and Fused-MBConv (Tan & Le, 2021)

The figure above depicts the architecture of MBConv and Fused-MBConv. In EfficientNetV2, there are slight differences in the layers used by the Fused-MBConv and Mobile

Inverted Bottleneck Convolution (MBConv) blocks. Fused-MBConv uses Conv3x3, while MBConv uses Conv3x3 and Conv1x1 in depth.

In our proposed EfficientNetV2S construction, before fully connected (FC), Batch Normalization is added with a momentum of 0.99, dense with L2 of 0.02, and DropOut of 0.25 to prevent overfitting.

We temporarily set the learning rate (LR), which will continue to decrease as the accuracy increases. When no increase in accuracy occurs after 7 epochs, the algorithm will stop.

Table 1
EfficientNetV2S + L2 Architecture

Stage	Operator	Stride	#Channels	Layers
0	Conv3x3	2	24	1
1	Fused-MBConv1, k3x3	1	24	2
2	Fused-MBConv4, k3x3	2	48	4
3	Fused-MBConv1, k3x3	2	64	4
4	MBConv4, k3x3, SE0.25	2	128	6
5	MBConv6, k3x3, SE0.25	1	160	9
6	MBConv4, k3x3, SE0.25	2	256	-
7	BatchNormalization 0.99	2	1280	-
8	Dense + L2 0.02	2	256	-
9	DropOut 0.25	2	256	-
10	Conv1x1 & Pooling & FC	-	256	1

This architecture was trained using hardware specifications: Lenovo-L13, 11th Gen Intel(R) Core(TM) i7-165G7, RAM 16 Gb, eGPU Razer Core X with VGA NVIDIA GeForce RTX3060. Untuk software menggunakan Windows 10 Home Single Language 64-Bit OS, Python 3.9.7, and Tensorflow 2.10.1.

As for the hyperparameter adjustment, it can be explained by the AI model below.

Table 2
Use of Hyperparameters in AI Model

Hyperparameter	Value:
Transfer Learning	EfficientNetV2S
Kernel Size	3 x 3
Image Size	384 x 384
Batch Size	16
Pooling	MaxPooling
Stride	1 x 1
Padding	Same
Optimizer	Adamax
Acitvation Function	ReLU and Softmax
Regularization	L1, L2, DropOut and Early Stopping

Hyperparameter	Value:
Learning Rate	Adaptive LR = 0.001

The learning rate setting starts from 0.001, but it is configured that if there is no improvement in validation accuracy for 2 epochs, the learning rate value will be multiplied by 10%.

RESULTS AND DISCUSSION

The experiments conducted resulted in a maximum epoch configuration of 50. Computation stopped at the 18th epoch with a duration of approximately 2 hours and 28 minutes. During the duration of the computation, the learning rate decreased until the 17th epoch reached 10^{-7}.

Table 3
Accuracy in Epoch Session

Epoch Session	LR	Train Accuracy	Test Accuracy
1 - 7	0.001	97.15%	85.92%
8 - 12	0.0001	99.13%	89.13%
13 - 14	0.00001	99.17%	89.54%
15 - 16	0.000001	99.27%	89.13%
17 - 18	0.0000001	99.21%	88.33%

The final result of training accuracy is 99.25% and for testing accuracy is 88.33% with an automatic decrease in learning rate by 4 times from the initial learning rate value. It can also be seen that the distance from the first epoch to the next epoch of each session is also getting smaller as the learning rate decreases.

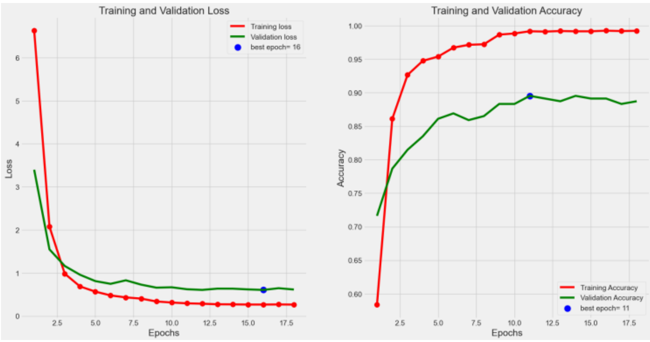

Figure.7.
The Result of Graph Training and Validation of L2-EffNet

From the graph above, it can also be seen that for Training Loss, the epoch with the best results is the 16th epoch, while for Training Accuracy, the epoch with the best results is the 11^{th} epoch.

Table 4
Comparative Analysis with Previous State-of-the-Art Skin Disease Classification Using Deep Learning

Ref.	Year	Authors	Datasets	Model	Accuracy
Our Proposed			Augmented Manually curated from Atlas Dermatology and ISIC Skin Disease Dataset (31 classes)	L2-EffNetV2S	88.33%
(Rafay & Hussain, 2023)	2023	Rafay et al.	Manually curated from Atlas Dermatology and ISIC Dataset (31 classess)	Fine-tuned EfficientNet-B2	Top-1 Accuracy: 87.15%
(Karthik et al., 2022)	2022	Karthik et al.	DermNet NZ, Derm7Pt, DermatoWeb, and Fitzpatrick17k databases	EfficientNetV2	84.70%

When compared to previous research (Rafay & Hussain, 2023), this proposed research has higher accuracy with the same number of classes.

CONCLUSION

The modified L2-EfficientNetV2S model, which used hyperparameters and fine-tuning, was able to produce a model with a higher test accuracy value of 88.33% with a total of 31 dataset variables. For further research, it can be done by using a different transfer learning model, increasing the number of datasets, or increasing computer hardware specifications so as not to wait too long to perform the computation process.

REFERENCES

Al-masni, M. A., Kim, D. H., & Kim, T. S. (2020). Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. *Computer Methods and Programs in Biomedicine*, *190*, 105351. https://doi.org/10.1016/j.cmpb.2020.105351

Cahyanto, K. A., Yani, M., Rendi, Kustiyo, A., & Haryanto, T. (2023). The Effects of L2 Regularization in EfficientNet for Human Skin Disease Multi-Class Classification. *Proceedings - 2023 6th International Conference on Computer and Informatics Engineering: AI Trust, Risk and Security Management (AI Trism), IC2IE 2023*, 262–265. https://doi.org/10.1109/IC2IE60547.2023.10331108

Calderón, C., Sanchez, K., Castillo, S., & Arguello, H. (2021). BILSK: A bilinear convolutional neural network approach for skin lesion classification. *Computer Methods and Programs in Biomedicine Update*, *1*, 100036. https://doi.org/10.1016/j.cmpbup.2021.100036

ISIC. (2019). *ISIC challenge dataset*. ISIC.

Isbahi, M. B., Zuana, M. M. M. ., & Mariana, E. R. . (2022). The Technology Strategy in Website Communication Media in Improving Business Activities. *Majapahit Journal of Islamic Finance and Management*, *1*(2), 126–138. https://doi.org/10.31538/mjifm.v1i2.17

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

Karthik, R., Vaichole, T. S., Kulkarni, S. K., Yadav, O., & Khan, F. (2022). Eff2Net: An efficient channel attention-based convolutional neural network for skin disease classification. *Biomedical Signal Processing and Control*, *73*(October 2021), 103406. https://doi.org/10.1016/j.bspc.2021.103406

Khan, M. A., Muhammad, K., Sharif, M., & ... (2021). Multi-class skin lesion detection and classification via teledermatology. *IEEE Journal of ...*.

Li, H., Pan, Y., Zhao, J., & Zhang, L. (2021). Skin disease diagnosis with deep learning: A review. *Neurocomputing*, *464*, 364–393. https://doi.org/10.1016/j.neucom.2021.08.096

Purnama, I. K. E., Hernanda, A. K., Agung, A., Ratna, P., Nurtanio, I., & Hidayati, A. N. (2019). *Disease Classification based on Dermoscopic Skin Images Using Convolutional Neural Network in Teledermatology System*. 1–5.

Rafay, A., & Hussain, W. (2023). EfficientSkinDis : An EfficientNet-based classification model for a large manually curated dataset of 31 skin diseases. *Biomedical Signal Processing and Control*, *85*(December 2022), 104869. https://doi.org/10.1016/j.bspc.2023.104869

Sadik, R., Majumder, A., Biswas, A. A., Ahammad, B., & Rahman, M. M. (2023). An in-depth analysis of Convolutional Neural Network architectures with transfer learning for skin disease diagnosis. *Healthcare Analytics*, *3*(January), 100143. https://doi.org/10.1016/j.health.2023.100143

Setiawan, N. B., Natalia, F., Ferdinand, F. V., Sudirman, S., & Ko, C. S. (2021). Classification Of Skin Diseases And Disorders Using Convolutional Neural Network On A Mobile Application. *ICIC Express Letters*, *12*(8), 715–721. https://doi.org/10.24507/icicelb.12.08.715

Shanthi, T., Sabeenian, R. S., & Anand, R. (2020). Automatic diagnosis of skin diseases using convolution neural network. *Microprocessors and Microsystems*, *76*, 103074. https://doi.org/10.1016/j.micpro.2020.103074

Silva, S. F. da. (2025). *Atlas Dermatology*.

Soumare, H., Benkahla, A., & Gmati, N. (2021). Deep learning regularization techniques to genomics data. *Array*, *11*(May), 100068. https://doi.org/10.1016/j.array.2021.100068

Tan, M., & Le, Q. V. (2021). *EfficientNetV2: Smaller Models and Faster Training*.

Zuana, M. M. M., Toha, M., & Isbahi, M. B. (2024). Exploration of Community Empowerment in a Village as the Entrance to a Lake in East Java. *Malacca: Journal of Management and Business Development* , *1*(1), 47–55. https://doi.org/10.69965/malacca.v1i1.52