International Journal of Advanced Research in Engineering Innovation
e-ISSN: 2682-8499 | Vol. 6, No. 1, 1-9, 2024
http://myjms.mohe.gov.my/index.php/ijarei

Classification 19 Type of Skin Condition by Using Convolution Neural Network

Muhammad Faris Hamdani[1*], Noranhairul Ariffin Abdullah[1], Gordon Kechendai Nyanggau[1]

[1] Jabatan Kejuruteraan Elektrik, Politeknik Mukah KM7.5 Jalan Oya, 96400 Mukah, Sarawak, Malaysia

*Corresponding Author: faris.poli1992@gmail.com

Received: 15 December 2023 | Accepted: 10 February 2024 | Published: 1 March 2024

DOI: https://doi.org/10.55057/ijarei.2024.6.1.1

Abstract: *Convolutional Neural Network (CNN) is a vision computerization analysis that assists in understanding the deep learning models. This study will carry out to purpose one method for machine learning files that can be used for classifying skin disease through deep learning based. This study will determine the precision(p), recall(r) and score(F1) of model to classify the type of skins into own categories. The focus is to generate the architecture hdf5 type file that can be automated for the diagnosis of 18 common skin diseases and 1 normal skin by using data from public clinical images dataset and patient information using deep learning pre-trained EfficientNetB7 model. The implement a simple image classification model using TensorFlow to rebuild one of parameter hospitality for nursery robot development in future. The image classification model will classify images of various skin disease problems into labelled classes.*

Keywords: Convolutional Neural Network (CNN), EffiecientNetB7 model

1. Introduction

Skin disease a significant healthcare challenge in medical industry which is requires accurate and timely diagnosis for effective classification and after treatment. Computer vision and machine learning techniques have shown promising results in classification object in past studies by using convolution neural network (CNN)(Yamashita et al., 2018). In this study, machine learning will classify the various types of skin disease into 19 categories including normal skin by using sing deep learning pre-trained EfficientNetB7 model to measure the robustness and accuracy for the model architecture. This design of model will generate a .h5 file and will be implemented into the hardware that can run the machine learning process in a future development. This model is concerned with robustness, accuracy, losses, and data collection. With the advancement of technology, the skin observing framework can be structured and executed for early detection of skin infections by through image classification (Kulshrestha et al., 2020). The purpose of this study is to evaluate how well EfficientNetB7 performance toward skin images to classify a modest-scale dataset of 2,604 images containing various types of skin diseases normal skin.

2. Literature

Humans use vision to adapt and understand their surroundings, but computer vision works on duplicating human vision but in an electronic form to see and interpret an image. It must be given the ability to detect, identify, and process images in the same way that human vision does. (Teoh et al., 2021) There a lot of method, techniques and algorithm was provided in low scale dataset to comply the skin disease classification with a popular pre-trained deep learning model, such as ResNet50, EfficientNetB7, InceptionV3, and VGG16(Mishra et al., 2023). But these studies are more focusing on EfficientNetB7 which are the highest accuracy on previous studies. (Selim et al., 2022; Shoukat et al., 2021; Wang et al., 2022). The method EffecientNetB7 have a lot of variants starting with EfficientNetB0 develop around 2019 by google researcher and introduce the development model variants until to EfficientNetB7 where each variant has depending on result by own accuracy percentage.

EfficientNetB7 is a specific variant of the EfficientNet family of convolutional neural networks (CNNs). The EffecientNet model was proposed by researchers with paper titled, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." The main idea behind EfficientNet is to provide better accuracy and efficiency by scaling the model's depth, width, and resolution. Instead of simply making models larger, the researcher introduced a compound scaling method that balances the three dimensions to find an optimal configuration (Tan & Le, 2019). EfficientNetB7 is the largest variant in the EfficientNet family with higher accuracy compared to other versions. The character "B" in EfficientNetB7 indicates that this variant of the EfficientNet model is based on the B-version scaling coefficients proposed in the research paper. The number "7" specifically refers to the compound scaling coefficient used for this variant (Tan & Le, 2019). There are a lot of researchers using a lower data scale applied to the EffecientNetB7 with high accuracy result. This study is focusing on large scale datasets with different values of epoch training cycle for combination of 19 type of skin.

Previous studies have demonstrated the exceptional performance of the EfficientNet model, showcasing its superiority over other architectures with the highest testing accuracy. As a result, the researchers were successful in designing a model that achieved a testing accuracy of 71% using a 70% training split, which was considered low. To address this, data augmentation was introduced to the 70% training split, which consisted of small quantity of image samples. However, it was noted that the 70/30 data split did not yield significant gains in accuracy. Consequently, the researchers decided to modify the data split, opting for an 80/20 train-test split. This adjustment resulted in a more substantial improvement, raising the accuracy score to 74% (Rafay & Hussain, 2023). The result shows if adjustment of dataset train-test split is affected to the accuracy score.

There is another evidence that indicates previous studies of benchmarking EfficientNetB7, InceptionResNetV2, InceptionV3, and Xception Artificial Neural Networks Applications for Aortic Pathologies Analysis. The result shown EfficientNetB7 model achieved the highest accuracy of 97.01% after training and validating in comparing their accuracy (Miserlis et al., 2023). The result gives high impact to all researchers to design the best method, algorithms, and techniques by proposing EfficientNetB7 as a mechanism or system for medical researchers. This model can apply to any hardware that can be running GPU or TensorFlow for future AI robotic development. To improve the accuracy can be increase by using image segmentation method, Pre-processing skin photos improved their quality for categorization by removing minute noise. The skin image's edges were then delineated using ground truth, and the backdrop was then removed. Then, the dataset image after segmentation applied in the transfer learning models. (Alwakid et al., 2023) This method purposes a lot of techniques depend a class input that be needed to classify but these studies a primary to find out the result on large scale dataset without image segmentation.

International Journal of Advanced Research in Engineering Innovation
e-ISSN: 2682-8499 | Vol. 6, No. 1, 1-9, 2024
http://myjms.mohe.gov.my/index.php/ijarei

Besides that, there are a few researchers using a convolution neural method to design a model layer of skin cancer detection. The author state that by using deep learning and image processing can remove a crucial part such hair-like noise from lesion. If the unnecessary image is not removed or done properly, the rate of success in classifying the lesion will decrease. (Alwakid et al., 2023) The image can recognize by convolution neural network on the certain area was affected. By using this method, the image can be improved to classify which class a more accurate and suitable to be compared in training dataset image.

Overall, EfficientNetB7's success in achieving high accuracy has made it a powerful tool for medical research and skin disease analysis. With ongoing advancements, deep learning techniques promise to contribute significantly to skin disease diagnosis and understanding. At the end of this study, the model will be generated by using pre-trained TensorFlow model and the model will be tested to validate the accuracy of model using method image augmentation. If the result is above 75% percent proven from 2,604 image dataset, the file that has been generate will be implanted into hardware that can running AI model such as Raspberry Pi 3B+ or Nvidia Jetson.

3. Methodology

Data preparation is the action of gathering the data needed to make a more convenient and reasonable before feeding machine learning. Massaging it into a format that's computer-readable and understandable and asking questions of it to check it for completeness and bias. Before loading the image dataset, all image files in each folder that have already been categorized need to be renamed into ''image_X " which is X for indexing files name by rearrange into increasing numbering. The reason for rearranging all folders to the fixed name is to make a dataset image readable. The all-dataset images will be manipulated and analysis to provide data structure like DataFrames and Series, which make more easier to work with structures data. The data structure will be organized into 19 types of skin categories as shown in Figure 1.

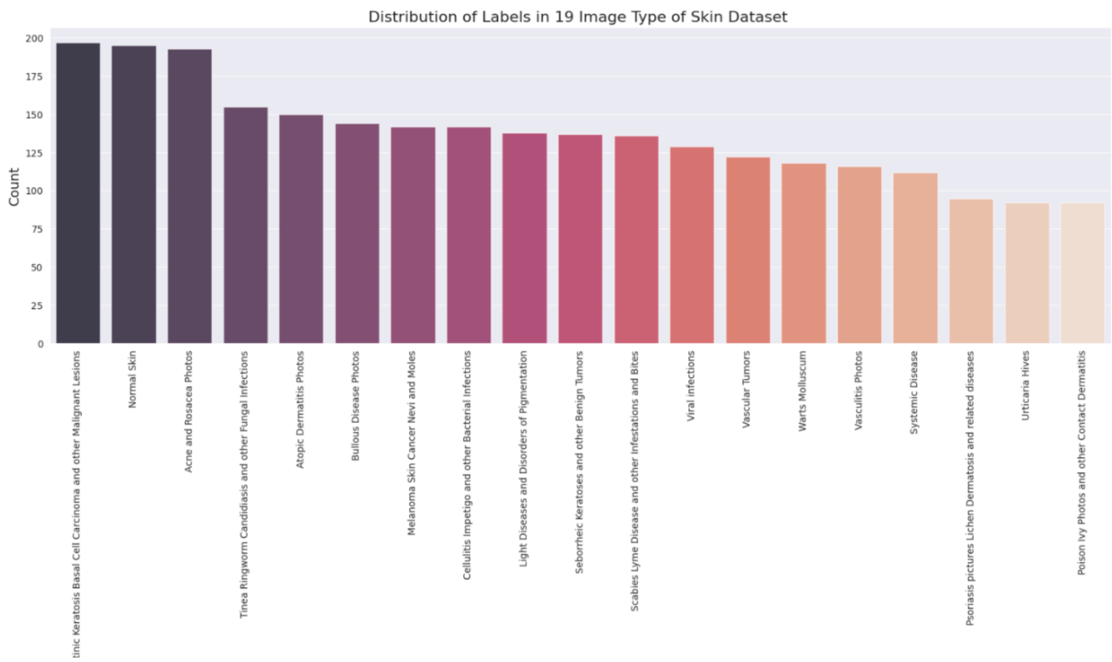

Figure 1: Distribution of 19 type skins categories.

International Journal of Advanced Research in Engineering Innovation
e-ISSN: 2682-8499 | Vol. 6, No. 1, 1-9, 2024
http://myjms.mohe.gov.my/index.php/ijarei

Based on Figure 1, the graph shows the numbers of images that have been categorized into 19 types of skins categories. The numbers files can be added on future development to make a dataset more robust and reliable. The dataset image file will assign and categorize following the folder name to make data easier to pre-process when running machine learning. The visualization can be tested in random image from data set to shows that all file's name is already follow the folder name. The random image dataset was plot as shown in Figure 2.

Figure 2: Random image that already organized with labels name.

Figure 2 shows the random sample of image are already rearrange with labels name depending on categories image files located in folder name. The data will be split into three categories such as Training, Validation and Testing. The training data will be used to train the deep learning CNN model and its parameter will be fine-tuned with the validation data. The objective of this purpose method is to check the performance of the data that will be evaluated using the data that has not complied in model or the data has not previously seen before. These studies are divided into 3 classes which is 1,668 validated image filenames belonging to train dataset,416 validated image filenames belonging to validation and lastly 521 validated image filenames belonging to testing purpose.

By using "***ImageDataGenerator***" class in TensorFlow or Keras is provided to generate batches of augmented image from directory source files before feed to next process which is a module for data augmentation pipeline. This parameter is used to split the data into training and validation sets. It indicates the fraction of images that should be used for validation during training. In this case, 20% of the data will be used for validation, and the remaining 80% will be used for training.The next process is applied a pipeline proposed by TensorFlow's "***keras.Sequential***" and "***layer.experimantal.preprocessing***" module to provide a pipeline applies a series data augmentation techniques to augment image before fed into deep learning. The sequential module provides resizing, rescaling, random flip, random rotation, random zoom, and random contrast before feeding into training process.

Data image augmentation is a technique commonly used in machine learning and computer vision to artificially increase the size of a training dataset by applying various type transformations to the existing images such as rotation, flip, zoom, brightness, contrast, color jitter etc. The goal is to improve the generalization and robustness of a machine learning model

International Journal of Advanced Research in Engineering Innovation
e-ISSN: 2682-8499 | Vol. 6, No. 1, 1-9, 2024
http://myjms.mohe.gov.my/index.php/ijarei

by exposing by using variations of input data. Data Augmentation is suitable for limitation dataset that available to used and it can prevent overfitting and enhances the model's ability.

Data image augmentation is a technique commonly used in machine learning and computer vision to artificially increase the size of a training dataset by applying various type transformations to the existing images such as rotation, flip, zoom, brightness, contrast, color jitter etc. The goal is to improve the generalization and robustness of a machine learning model by exposing by using variations of input data. Data Augmentation is suitable for limitation dataset that available to used and it can prevent overfitting and enhances the model's ability.

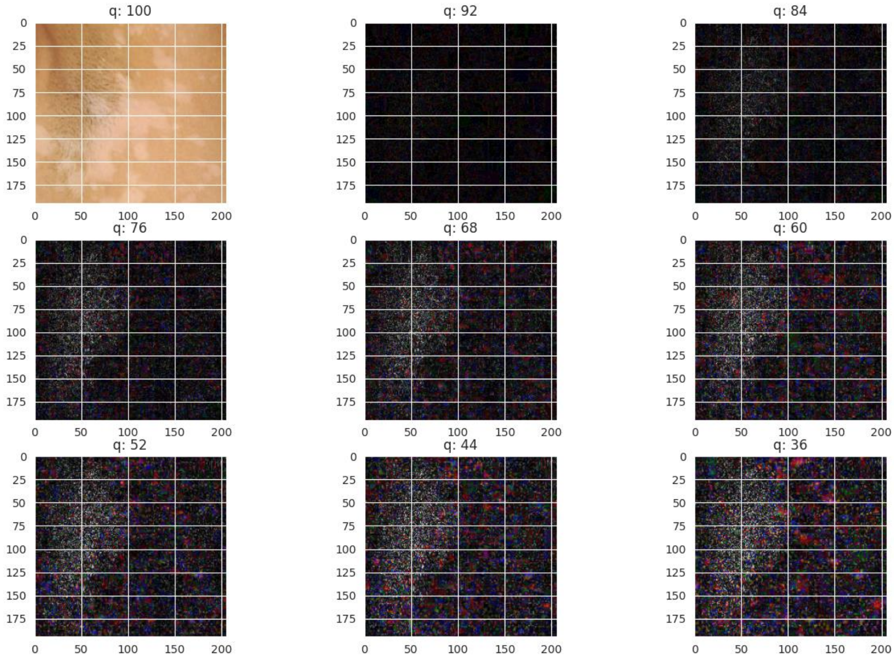

Figure 3: The example image of Light Diseases and Disorders of Pigmentation

Figure 3 shows the data image can be extracted into different types of variants with different brightness, saturation, and color jitter to feed into an input dense for model architecture before training. This technique can improve the accuracy and robustness of training model. But the model depends also on the value of epoch iteration and type of GPU used in the model design. The pre-trained model is used as a feature extractor and its output will be used as an input to custom layers of the model.

The layers of the model called "Dense" which is its connected layers with dropout are added to the model. Dropout is a technique to avoid overfitting on model architecture. The "Adam Optimizer" was applied to this model because of a lot of advantages such as adaptive learning rate, efficiency, low memory, sparse gradients well, invariant to rescaling, intuitive hyperparameter, bias correction and works well with noise.(Kingma & Ba, 2014) The epoch for this model is 100 which is the model has learned and adjusted its parameters based on the training example for hundred time. The pre-trained model will generate .h5 file as model of layer that can be applied to any web-based or hardware application such as hugging face, Raspberry Pi 3B+ or Nvidia Jetson.

The model evaluation can be recalculated based on the metrics. The model will be tested to check the accuracy which measures the fraction of predictions the model.(Jordan et al., n.d.)Accuracy measures the fraction of correct predictions made by the model out of the total number of predictions. It is a common metric for binary and multi-class classification problems. The formula for accuracy is:

$$Accuracy = \frac{True\ Positives + True\ Negatives}{True\ Positives\ +\ True\ Negatives\ +\ False\ Positives\ +\ False\ Negatives}$$

Precision is the fraction of true positives (correctly predicted positive samples) out of the total number of predicted positive samples (true positives + false positives). In multi-class classification problems, precision is calculated for each class separately and then averaged among the classes. The formula for precision is:

$$Precision\ (P) = \frac{True\ Positives}{True\ Positives\ +\ False\ Positives}$$

Recall, also known as sensitivity or true positive rate, is the fraction of true positives out of the total number of actual positive samples (true positives + false negatives). For multi-class classification problems, recall is calculated for each class individually and then averaged among the classes. The formula for recall is:

$$Recall\ (R) = \frac{True\ Positives}{True\ Positives\ +\ False\ Negatives}$$

The F1 score is the harmonic mean of precision and recall. It is a single metric that balances both precision and recall and is useful when there is an imbalance between classes. Like precision and recall, the F1 score is also calculated for each class separately in multi-class problems and then averaged among the classes. The formula for the F1 score is:

$$F1\ Score\ (R) = 2\ x\ \frac{Precision\ (p)\ x\ Recall}{Precision +\ Recall}$$

These metrics are valuable tools for assessing the overall performance of a classification model and understanding how well it performs in correctly identifying positive and negative samples. Depending on the problem and the importance of precision and recall, different metrics may be emphasized to evaluate the model's effectiveness.

4. Results

The focus of this research is to define the relation between accuracy and total loss by using epochs in 1000 iteration in training on GPU RTX5000 applied a EffecientNetB7 training model.These validations show the accuracy and loss of model after the training was completed.

International Journal of Advanced Research in Engineering Innovation
e-ISSN: 2682-8499 | Vol. 6, No. 1, 1-9, 2024
http://myjms.mohe.gov.my/index.php/ijarei

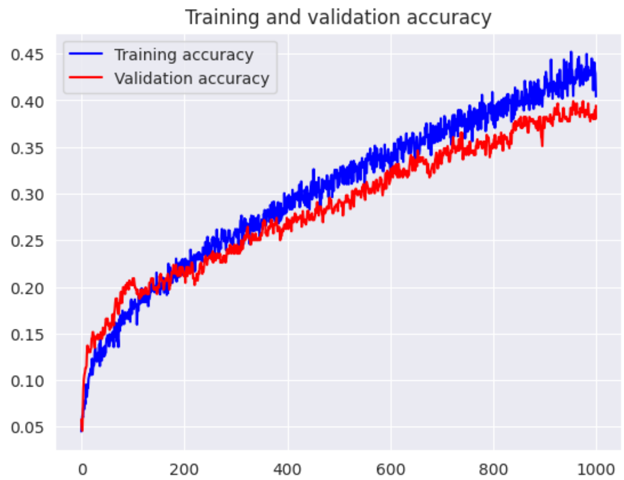

Figure 4: Validation Accuracy.

Regarding from table Figure 4 shows the result after 1000 epochs iteration cycle after training was proceeding. Based on Figure 4, the value of accuracy training is much slightly higher than validation accuracy data. The data proves that if the iteration of epoch increasing will produce more precision and accuracy on the model classification. The graph shows fluctuation on every step of epoch in training session caused by the low-quality and unstable data image on each type of skin disease. This causes certain data images to not focus on subject matter toward their own classification. The validation can be misjudgment on exact type of skin during training. The training accuracy is achieving 45% maximum performance within iteration epochs 800 to 1000 epochs, the best weighted average value is 0.393016 in term of precision. This research can be improved by increasing the value of epochs training for future development.

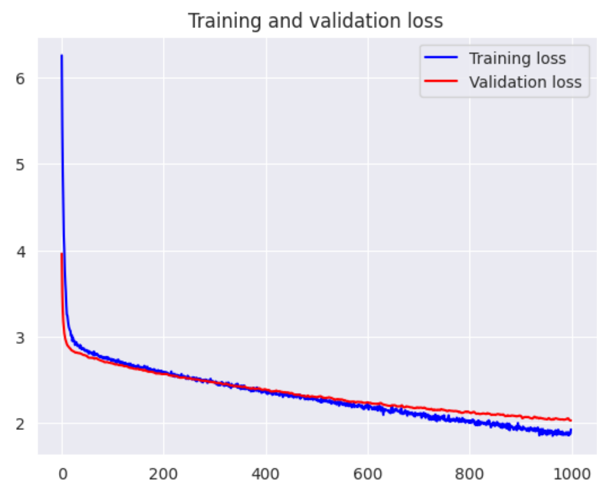

Figure 5: Validation Loss.

Figure 5 shows good feedback on validation loss graph decreasing per epoch iteration, these data proven the validation loss graph can be decreasing if the epoch training are increasing. The validation loss graph shows its frequent drop toward 1000 iteration epochs cycle. If the epochs training was increasing into 2000, the expected research findings will produce more low validation loss twice compared to latest findings.

A confusion matrix is a table that lists how many predictions a classification model on a set of test data made correctly and incorrectly. Rows and columns of a square matrix are typically used to represent the predicted and actual class labels, respectively. The entries in the matrix show how many test samples there were overall, and how many of those were successfully or wrongly identified by the model. An in-depth analysis of the model's performance, including metrics like accuracy, precision, recall, and F1-score for each class, may be found in a confusion matrix. It can be used to diagnose issues with the model's predictions and to pinpoint specific instances where the model is off. Figure 6 shows the result at the end of training in 1000 epochs iteration by using Graphical Processing Unit (GPU) model RTX5000.

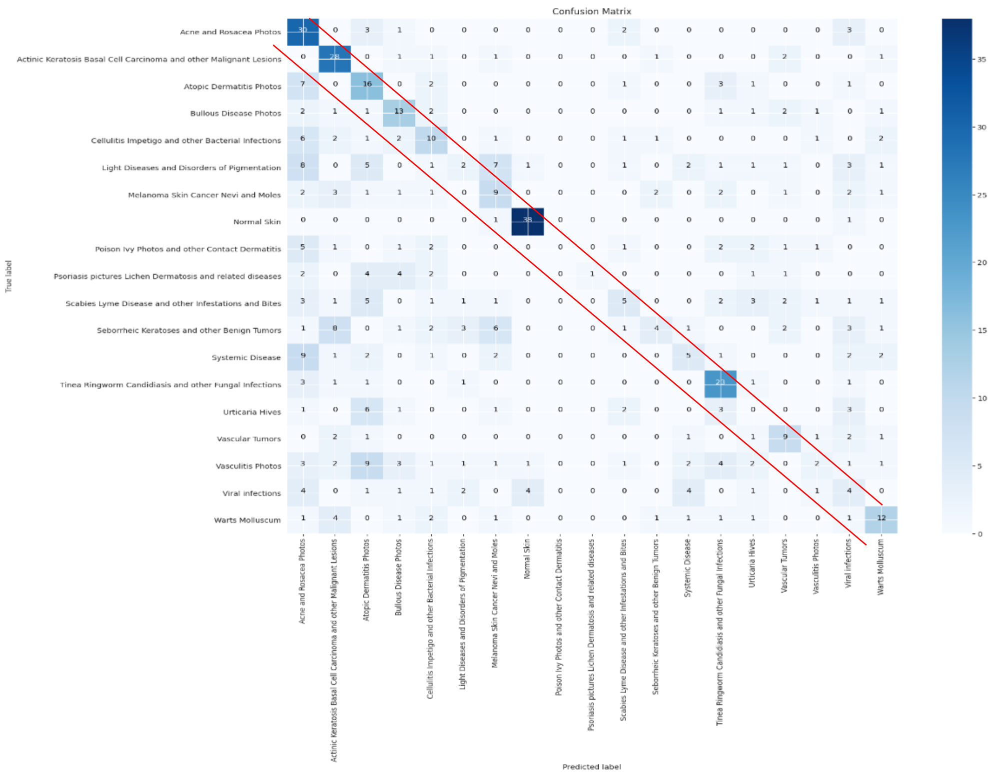

Figure 6: Confusion Matrix Skin Data

Based on Figure 6 shows that the confusion matrix table indicates that the quantity of predicted image skin type disease versus true label classification. The box with dark blue color means that a lot of data predict on same label type of skins. These results can be determined for accuracy on each data image if the predicted image and true label have drop-in same location in straight line (red color) means that the image validation are corrected, meanwhile if the image are not categorize falling into straight line means that the validation are false.

Based on the confusion matrix table, there a have some of type skin with high score true validation which is Acne and Rosacea Photos, Actinic Keratosis Basal Cell Carcinoma and other Malignant Lesions, Atopic Dermatitis Photos, Bullous Disease Photos and Normal Skin because of quality data image a more precise to the subject compare with other data image.

5. Conclusion

In conclusion, the exploration of EfficientNetB7 for skin disease classification has revealed a validation accuracy of 40%, indicating both challenges and promising opportunities for enhancing the model's performance. While the current accuracy may seem modest, it serves as a crucial starting point for further refinement. This initial assessment underscores the importance of addressing key considerations, including Data Quality and Quantity, Data Augmentation and Preprocessing, Hyperparameter Tuning, Continuous Evaluation and

Iteration, and Model Interpretability. The model can improve by tackling on others aspects, there is substantial potential to elevate the model's accuracy and applicability for skin disease classification. This research not only provides a solid foundation for improvement in dermatological applications but also lays the groundwork for extending the methodology to various types of physical diseases in future development.

Acknowledgement
The author would like to thank Politeknik Mukah for providing the support for these studies. The authors would like to give appreciation and gratitude to Atlasdermatologico.com, Kaggle and DermNet’s for providing public data access.

References

Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. Insights into Imaging, 9(4), 611–629.

Kulshrestha, C., Juyal, P., Sharma, S., & Juyal, S. (2020). Machine Learning Algorithms based Skin Disease Detection. International Journal of Innovative Technology and Exploring Engineering, 9, 4044–4049.

Teoh, K. H., Ismail, R. C., Naziri, S. Z. M., Hussin, R., Isa, M. N. M., & Basir, M. S. S. M. (2021). Face Recognition and Identification using Deep Learning Approach. Journal of Physics: Conference Series, 1755(1).

Mishra, N., Jahan, I., Nadeem, Md. R., & Sharma, V. (2023). A Comparative Study of ResNet50, EfficientNetB7, InceptionV3, VGG16 models in Crop and Weed classification. 2023 4th International Conference on Intelligent Engineering and Management (ICIEM), 1–5.

Selim, T., Elkabani, I., & Abdou, M. A. (2022). Students Engagement Level Detection in Online e-Learning Using Hybrid EfficientNetB7 Together With TCN, LSTM, and Bi-LSTM. IEEE Access, 10, 99573–99583.

Shoukat, A., Akbar, S., Hassan, S. A. E., Rehman, A., & Ayesha, N. (2021). An Automated Deep Learning Approach to Diagnose Glaucoma using Retinal Fundus Images. 2021 International Conference on Frontiers of Information Technology (FIT), 120–125.

Wang, T., Sun, J., Li, Y., & Toe, T. T. (2022). Detecting and Classifying Misplaced Catheters III Chest X-Rays Based on Efficient net B7. 2022 IEEE 10th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), 10, 1133–1138.

Tan, M., & Le, Q. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In K. Chaudhuri & R. Salakhutdinov (Eds.), Proceedings of the 36th International Conference on Machine Learning (Vol. 97, pp. 6105–6114). PMLR.

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, 104869.

Miserlis, D., Munian, Y., Ferrer Cardona, L. M., Teixeira, P. G. R., DuBose, J. J., Davies, M. G., Bohannon, W., Koutakis, P., & Alamaniotis, M. (2023). Benchmarking EfficientNetB7, InceptionResNetV2, InceptionV3, and Xception Artificial Neural Networks Applications for Aortic Pathologies Analysis. Journal of Vascular Surgery, 77(6), e345.

Alwakid, G., Gouda, W., Humayun, M., & Jhanjhi, N. Z. (2023). Diagnosing Melanomas in Dermoscopy Images Using Deep Learning. Diagnostics, 13(10).

Kingma, D. P., & Ba, J. (2014). *Adam: A Method for Stochastic Optimization*.