JURNAL ILMU KOMPUTER
Volume 3 Nomor 1, Juli, 2025
ISSN: 3031-125X (ONLINE)
Penerbit: Program Studi Teknik Informatika S-2 Universitas Pamulang

Optimizing Learning Rate, Epoch, and Batch Size in Deep Learning Models for Skin Disease Classification

Taufiqur Rahman [1], Sajarwo Anggai [2], Arya Adhyaksa Waskita [3]
[3)] Program Studi Teknik Informatika S-2, Universitas Pamulang
Email: [1]tr.mr09.lab@gmail.com, [2]sajarwo@mail.com, [3]aawaskita@gmail.com

ABSTRAK

Penelitian ini bertujuan untuk menemukan kombinasi terbaik dari learning rate, jumlah epoch, dan ukuran batch untuk melatih model deep learning dalam klasifikasi penyakit kulit. Percobaan dilakukan dengan menganalisis grafik perubahan nilai loss terhadap learning rate dalam skala logaritmik. Hasilnya menunjukkan bahwa learning rate sekitar 10^{-2} adalah yang paling optimal, sementara penggunaan 5×10^{-3} memberikan kestabilan tambahan selama pelatihan. Dalam hal jumlah epoch dan ukuran batch, berbagai kombinasi diuji, mulai dari 20 hingga 100 epoch, dengan batch size 32 hingga 128. Percobaan menunjukkan bahwa batch size 32 memberikan hasil terbaik, dengan akurasi validasi mencapai 97,35% dan nilai loss validasi terendah sebesar 0,1074. Meskipun batch size 128 lebih efisien dari segi waktu, akurasi yang dihasilkan sedikit lebih rendah. Dengan menggunakan 25 epoch dan batch size 32, model berhasil mencapai performa optimal tanpa tanda-tanda overfitting. Proses persiapan data juga menjadi fokus penting, seperti menyesuaikan ukuran gambar, normalisasi piksel, dan augmentasi data, agar sesuai dengan kebutuhan model seperti VGG-19, Inception-V4, dan ResNet-152. Visualisasi distribusi dataset membantu memastikan kualitas data dan keseimbangan kelas, sehingga model dapat memahami pola dengan lebih baik. Hasil penelitian ini memberikan panduan praktis untuk melatih model deep learning secara efektif dan efisien, khususnya dalam tugas klasifikasi penyakit kulit.

Kata kunci: Deep Learning, Klasifikasi Penyakit Kulit, Learning Rate, Epoch, Batch Size, Augmentasi Data, VGG-19, Inception-V4, ResNet-152, Visualisasi Dataset, Overfitting.

ABSTRACT

This study explores the best combination of learning rate, number of epochs, and batch size for training deep learning models to classify skin diseases. The experiments involved analyzing how loss changes with learning rates on a logarithmic scale. The findings reveal that a learning rate of approximately 10^{-2} is most effective, with 5×10^{-3} offering additional stability during training. Various combinations of epochs and batch sizes were tested, ranging from 20 to 100 epochs and batch sizes between 32 and 128. The results show that using a batch size of 32 yielded the best outcomes, achieving a validation accuracy of 97.35% and the lowest validation loss of 0.1074. While a batch size of 128 was more efficient in terms of time, it resulted in slightly lower accuracy. The model performed optimally with 25 epochs and a batch size of 32, avoiding any signs of overfitting. Data preparation also played a crucial role, involving steps like image resizing, pixel normalization, and data augmentation to align with the requirements of models such as VGG-19, Inception-V4, and ResNet-152. Visualizing the dataset distribution ensured data quality and class balance, allowing the model to better recognize patterns. This study offers practical insights for effectively and efficiently training deep learning models, particularly for tasks related to skin disease classification.

Keywords: Deep Learning, Skin Disease Classification, Learning Rate, Epoch, Batch Size, Data Augmentation, VGG-19, Inception-V4, ResNet-152, Dataset Visualization, Overfitting.

1. INTRODUCTION

Skin Diseases: Advancing Diagnosis with Deep Learning

Skin diseases are among the most common health conditions, ranging from mild issues to life-threatening illnesses like skin cancer. Early diagnosis plays a critical role in preventing severe complications and improving treatment outcomes. In recent years, deep learning approaches, particularly Convolutional Neural Networks (CNNs), have emerged as innovative solutions to assist healthcare professionals in detecting and classifying skin diseases automatically and accurately (Hajiarbabi, et.all (2023)[1].

Models such as Inception, VGG, and ResNet have demonstrated their effectiveness in recognizing complex patterns in medical images. A notable advantage of these models is their ability to leverage transfer learning, utilizing pre-trained weights from large datasets like ImageNet. This significantly accelerates the training process on smaller datasets (Developers, Ehsan et.all)[2], [3]. However, model performance heavily depends on the optimization of training parameters, such as learning rate, number of epochs, and batch size. For instance, a high learning rate might destabilize the model, while a low value could result in slow training (Eesteva, et.all., Hale, et,all.)[4], [5]

Data preparation is another crucial aspect in ensuring optimal model performance. Techniques like data augmentation, pixel normalization, and dataset visualization not only enhance the model's understanding of the data but also help prevent overfitting, especially when working with small datasets (Naeem, et.all., Qin, et.all.) [6], [7].

This research focuses on exploring how combinations of training parameters, such as learning rate, epochs, and batch size, influence the accuracy and efficiency of CNN models in skin disease classification. Previous studies have highlighted the importance of tuning training parameters to achieve optimal performance. For example, Agustina and Anisa (2024) [8], [9] utilized ResNet-50 for facial skin classification, while successfully employed the VGG-16 architecture to detect skin cancer. In this context, the research aims to identify the most reliable and efficient parameter combinations for similar tasks.

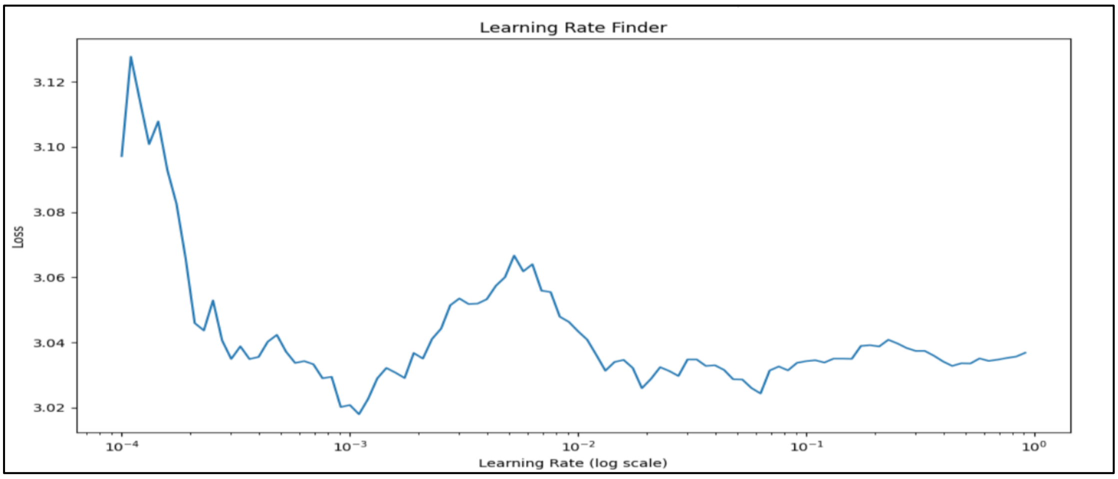
Beyond training parameters, this study emphasizes the significance of data preprocessing. As Fu'adah et al. (2020) [10]pointed out, normalization helps maintain stability during model training, while data augmentation, as applied by Ehsan Bazgir et al. (2024) [3], enhances dataset variability, reducing the risk of overfitting, particularly with limited datasets. Resizing images to align with pretrained architectures, such as VGG-19, Inception-V4, and ResNet-152, is also an essential step. For instance, Khan et al. (2023) [11] tailored their data to be compatible with Inception-V4, ensuring seamless integration with the optimized model.

This research not only prioritizes accuracy but also computational efficiency. By optimizing training parameters and employing appropriate data preprocessing, it aims to minimize training time and memory requirements without compromising model performance. This approach aligns with Hajiarbabi's (2023) [1]findings, which emphasize balancing efficiency and accuracy in deep learning. The results are expected to provide practical guidance for developing more effective deep learning models for diagnosing skin diseases. Additionally, the research aims to offer tangible benefits to healthcare, including real-world applications such as implementation in community clinics or pesantren, as proposed by Agustin and Putra (2023) [12].

2. METHOD

This research aims to identify the optimal combination of learning rate, number of epochs, and batch size in training deep learning models for skin disease classification. The first step is to test various learning rate values to determine the most effective one. In this experiment, the learning rate values are tested on a logarithmic scale, ranging from10^{-4} to 10^{-2}.

The results show that a learning rate of approximately 10^{-2} provides the best performance with consistent loss reduction, while a value of $5x10^{-3}$ offers additional stability during training.



Gambar 1. Percobaan Learning Rate

The next step in this research was to evaluate how variations in the number of epochs and batch size impacted the model's performance. Tests were conducted with epochs set to 20, 25, 50, 75, and 100, while the batch sizes tested were 32, 64, and 128. The findings showed that a batch size of 32 delivered the best results, achieving the lowest validation loss (0.1074) and the highest validation accuracy (97.35%). However, a batch size of 128 proved to be more efficient in terms of training time, as it required fewer iterations, though it produced slightly lower accuracy (96.77%). Ultimately, the most balanced combination was found to be 25 epochs with a batch size of 32, striking an optimal balance between accuracy and efficiency.

Preparing the data for training was a critical step to ensure the dataset was properly optimized for the models. Images were resized based on the model architecture: 224x224 pixels for VGG-19 and ResNet-152, and 299x299 pixels for Inception-V4. The pixel values were normalized to a scale of 0-1, which helped speed up training and stabilize the gradient updates. To further enhance the dataset, data augmentation techniques such as rotation, flipping, zooming, and brightness adjustments were applied. These methods increased dataset variability and reduced the risk of overfitting. Additionally, the dataset was carefully visualized to verify balanced class distributions and identify any potential issues.

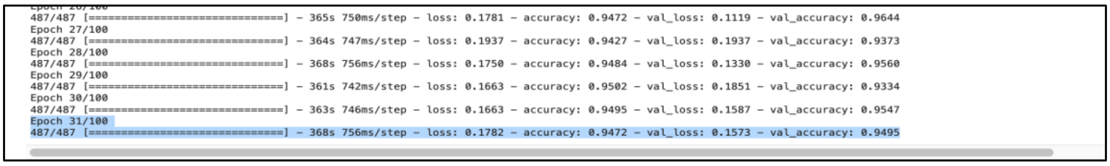
Three deep learning models—VGG-19, Inception-V4, and ResNet-152—were used in the training process, with all models leveraging pretrained weights from ImageNet. Fine-tuning was applied to the final layers of each model to align them with the number of classes in the dataset, enabling the models to be specifically tailored for the task of skin disease classification. This approach ensured the models were both efficient and highly accurate in their predictions.

The model's performance was evaluated using key metrics such as accuracy and loss to measure its ability to make predictions on validation and test datasets. To gain deeper insights into the results, a confusion matrix was used to analyze how predictions were distributed across different classes. Additionally, precision, recall, and F1-score were calculated, which are especially useful for assessing the model's performance on imbalanced classes.

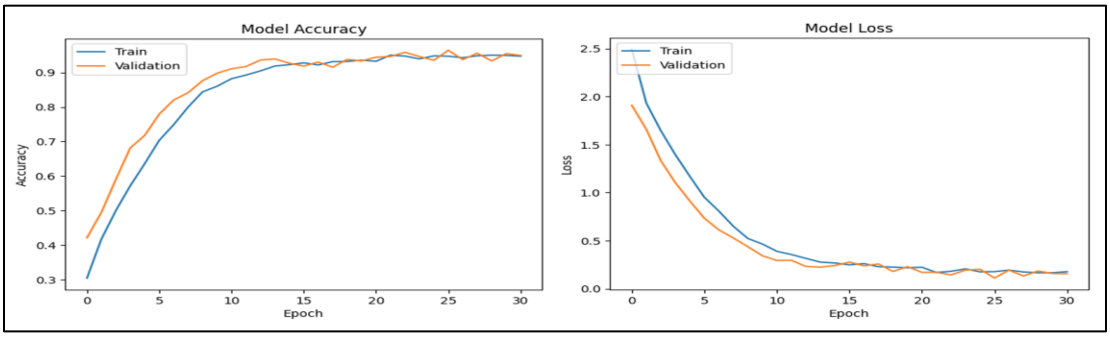
To understand how the training parameters affected the model, graphs of accuracy and loss were compared. This allowed for a clear visualization of trends and the impact of parameter variations. Efficiency was also taken into account by monitoring training time and memory usage, ensuring the model's development was not only effective but also resource-efficient.

This approach successfully identified the best combination of training parameters while maintaining a focus on computational efficiency. The findings are expected to serve as practical guidelines for further research, contributing to the development of accurate and efficient deep learning models for skin disease classification.

The research also demonstrated a sample scenario where, by optimizing learning rates, epochs, and batch sizes, the training process reached convergence efficiently, ensuring reliable and stable model performance.



Gambar 2. Epoch 31 **stopped training process



Gambar 3. Grafik Tingkat Konvergensi (Model Accuracy vs Model Loss)

3. RESULTS AND DISCUSSION

This section presents and analyzes the results of the experiments, highlighting their significance and implications. The discussion focuses on evaluating the model’s performance using key metrics, examining how the optimization of training parameters influenced the outcomes, and assessing the overall efficiency of the process. The goal is to provide clear insights and findings that demonstrate how the research objectives were achieved and to offer a comprehensive understanding of the results.

3.1. Learning Rate Experiment

The learning rate is a crucial parameter that significantly impacts the efficiency and stability of the deep learning model training process. This study identified several important patterns regarding the effect of the learning rate on the loss values:

a. **Small Learning Rate (10^{-4})**

With a small learning rate, the loss decreases steadily but very slowly. This is because the weight updates during each iteration are minimal, leading to a much longer training time to achieve optimal performance. While the graph shows a consistent downward trend, the lack of significant drops highlights the inefficiency in utilizing the model's full potential. Although this approach is safe and stable, it is not resource-efficient.

b. **Optimal Learning Rate (10^{-2})**

At this learning rate, the loss decreases significantly and reaches its lowest point. The graph illustrates that the model learns effectively without excessive fluctuations. The weight updates are large enough to speed up training yet remain within a stable range. This learning rate represents the optimal zone where the model learns efficiently and achieves the best balance between speed and stability.

c. **Large Learning Rate ($>10^{-2}$)**

When the learning rate is too large, the graph indicates a rise in the loss values, signaling model instability. This occurs because the weight updates overshoot the optimal values, leading to overshooting and making it difficult for the model to adapt to the data. This instability often results in poor convergence or, in extreme cases, training failure.

These findings emphasize the importance of carefully tuning the learning rate to achieve a balance between stability and training speed.

3.2. Experimenting with Epochs and Batch Sizes

This experiment aimed to find the best combination of the number of epochs and batch size to optimize the model's performance. The focus was on striking the right balance between accuracy, stability, and training efficiency by testing different configurations systematically. This approach ensures that the model performs well while using computational resources effectively.

a. Batch Size 128

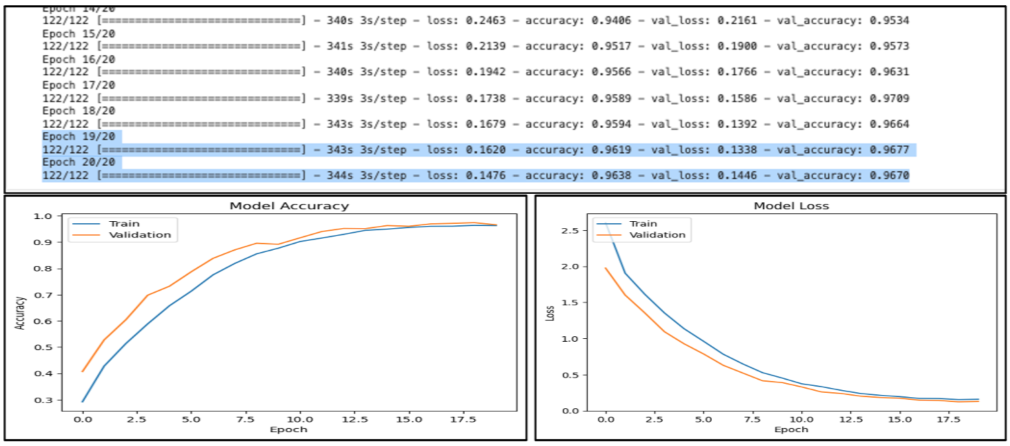
A larger batch size, such as 128, offers better training efficiency due to fewer iterations per epoch. However, the results showed that this batch size resulted in a higher validation loss (0.1109) and lower validation accuracy (96.77%). Larger batch sizes tend to smooth the gradient updates, which reduces the model's sensitivity to minor variations in the data. As a result, the model's ability to generalize is somewhat limited.

b. Batch size 32

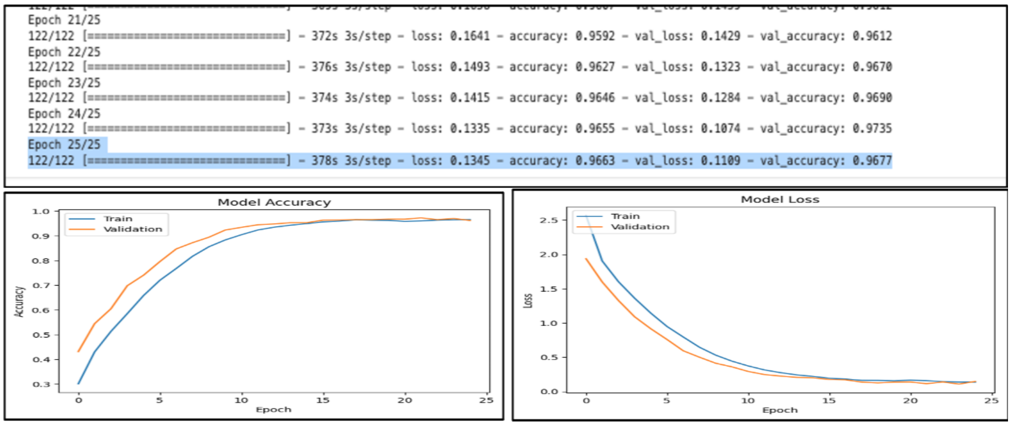
This batch size delivered the best performance, achieving the lowest validation loss (0.1074) and the highest validation accuracy (97.35%). Although training took longer due to the increased number of iterations, the smaller batch size allowed the model to capture more data variations. This enhanced the model's generalization capabilities, making it the optimal choice for achieving

c. Number of Epochs

1. The model performed optimally between 20 and 25 epochs. At this point, the loss graph stabilized, and validation accuracy reached its peak. The model successfully converged without signs of overfitting, as illustrated in Figures 4 and 5. This balance ensured that the model learned effectively while avoiding performance degradation.

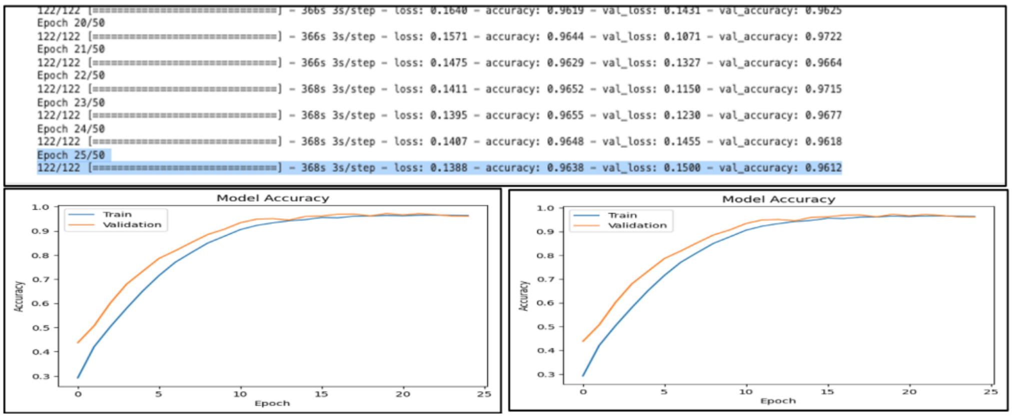


Gambar 4. *Epoch 20 Batchsize 128*



Gambar 5. Epoch 25 Batchsize 128

2. When training was extended beyond 50 epochs, no significant improvement in accuracy or reduction in loss was observed. On the contrary, the risk of overfitting increased, as the model began to overly adjust to the training data, which negatively impacted its ability to generalize effectively on validation data.



Gambar 6. Epoch 50 Batchsize 128

3.3. Impact of Data Preprocessing

Data preprocessing plays a crucial role in preparing the dataset for effective training of deep learning models. This section discusses how preprocessing steps, such as image resizing, normalization, and augmentation, influenced the performance of the model. Proper preprocessing not only enhances the model's ability to learn from the data but also minimizes the risks of overfitting and improves generalization on unseen data.

a. Normalization

The dataset was normalized to scale pixel values between 0 and 1. This ensured stable gradient updates during training, which helped accelerate model convergence. By avoiding excessively large or small gradient values, normalization reduced the risk of slowed training or model instability.

b. Data Augmentation

To enhance the dataset's variability, augmentation techniques such as rotation, flipping, zooming, and brightness adjustments were applied. These techniques helped the model recognize a broader range of patterns in the data, reducing the likelihood of overfitting, particularly with smaller datasets.

c. Image Resizing

All images were resized to align with the input size requirements of the chosen model architectures. For VGG-19 and ResNet-152, images were resized to 224x224 pixels, and for Inception-V4, they were resized to 299x299 pixels. This step ensured compatibility with pretrained model layers, optimizing their performance.

d. Dataset Visualization

The dataset was visualized to verify data quality, ensure balanced class distribution, and identify potential issues, such as low-quality images or imbalances among classes. This step ensured the data was suitable for training and minimized bias toward any specific class.

These preprocessing steps made the dataset more robust and optimized for training, allowing the model to better generalize and achieve accurate predictions.

4. Conclusion and Recommendations

This study successfully identified the optimal training parameters for deep learning models in skin disease classification. The best learning rate was 10^{-2}, with 5×10^{-3} providing extra stability during training. Smaller learning rates (10−410^{-4}) resulted in slower progress, while larger ones ($>10^{-2}$) led to instability and overshooting. The ideal setup for training was found to be 25 epochs with a batch size of 32, which achieved the lowest validation loss (0.1074) and the highest validation accuracy (97.35%). While a batch size of 128 offered faster training, it came with a slightly lower accuracy (96.77%). Extending training beyond 50 epochs did not improve performance and instead increased the risk of overfitting. Additionally, preprocessing steps such as normalization, data augmentation, and image resizing were critical for ensuring data quality, enhancing the model's generalization ability, and avoiding bias toward specific classes.

Based on these findings, the following suggestions are made for future research and development:

1. Explore Additional Parameters: Further investigation into parameters like optimizers and activation functions could lead to even better model performance.
2. Expand the Dataset: Using larger and more diverse datasets can help the model recognize more complex patterns and reduce overfitting risks.
3. Adopt Advanced Techniques: Leveraging technologies like federated learning or customized transfer learning for specific skin conditions can enhance accuracy and adaptability.
4. Optimize for Efficiency: Applying techniques such as model distillation or quantization could lower computational requirements without sacrificing accuracy.
5. Validate in Real-World Settings: Testing the model on clinical data is crucial to ensure it can be reliably applied in real medical environments.

Developing deep learning models for skin disease classification has the potential to revolutionize healthcare by providing more effective, efficient, and practical solutions. This research emphasizes the importance of addressing current challenges while driving innovation in skin disease diagnosis. By integrating these models into mobile apps or cloud platforms, access to diagnostic tools can be expanded, particularly in underserved areas. Collaboration with healthcare professionals will be key to ensuring these tools are clinically relevant and practical. Ultimately, this work sets the stage for deep learning to improve patient care, reduce diagnostic errors, and advance AI's role in healthcare.

5. DAFTAR PUSTAKA

[1] M. Hajiarbabi, "Skin cancer detection using multi-scale deep learning and transfer learning," *J Med Artif Intell*, vol. 6, no. 6, pp. 1–9, 2023, doi: 10.21037/jmai-23-67.

[2] T. Developers, "TensorFlow," Oct. 2024, *Zenodo*. doi: 10.5281/zenodo.13989084.

[3] Ehsan Bazgir, Ehteshamul Haque, Md. Maniruzzaman, and Rahmanul Hoque, "Skin cancer classification using Inception Network," *World Journal of Advanced Research and Reviews*, vol. 21, no. 2, pp. 839–849, 2024, doi: 10.30574/wjarr.2024.21.2.0500.

[4] A. Esteva *et al.*, "Deep learning-enabled medical computer vision," *NPJ Digit Med*, vol. 4, no. 1, pp. 1–9, 2021, doi: 10.1038/s41746-020-00376-2.

[5] J. Hale, "Deep Learning Framework Power Scores," https://towardsdatascience.com. [Online]. Available: https://towardsdatascience.com/deep-learning-framework-power-scores-2018-23607ddf297a

[6] A. Naeem, Ahmad;Shoaib, Muhammad; Farooq;Khelifi, Adel;Abid, "Malignant Melanoma Classification Using Deep Learning.pdf," 2020, [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9113301

[7] C. Qin *et al.*, "Dynamically aggregating MLPs and CNNs for skin lesion segmentation with geometry regularization," *Comput Methods Programs Biomed*, vol. 238, 2023, doi: 10.1016/j.cmpb.2023.107601.

[8] Agustina and Dian Anisa, "JURNAL RISET SISTEM INFORMASI (JISSI) KLASIFIKASI CITRA JENIS KULIT WAJAH DENGAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN) RESNET-50," vol. 1, no. 3, pp. 2–8, 2024.

[9] R. Agustina, R. Magdalena, and N. K. C. Pratiwi, "Klasifikasi Kanker Kulit menggunakan Metode Convolutional Neural Network dengan Arsitektur VGG-16," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 10, no. 2, p. 446, 2022, doi: 10.26760/elkomika.v10i2.446.

[10] Y. N. Fu'adah, N. C. Pratiwi, M. A. Pramudito, and N. Ibrahim, "Convolutional Neural Network (CNN) for Automatic Skin Cancer Classification System," *IOP Conf Ser Mater Sci Eng*, vol. 982, no. 1, 2020, doi: 10.1088/1757-899X/982/1/012005.

[11] S. Khan, H. Ali, and Z. Shah, "Brain Hemorrhage Detection Using Improved AlexNet with Inception-v4," *2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things, AIBThings 2023 - Proceedings*, no. September, pp. 1–5, 2023, doi: 10.1109/AIBThings58340.2023.10292461.

[12] I. R. Agustin and M. B. N. Putra, “Prediction of Skin Diseases using Convolutional Neural Networks as an Effort to Prevent Their Spread in Islamic Boarding School Environments,” *Khazanah Journal of Religion and Technology*, vol. 1, no. 2, pp. 49–53, 2023, doi: 10.15575/kjrt.v1i2.296.