Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 2, April 2026, Page. 948-960
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

Deep Learning Based MobileNet Optimization For High Accuracy Classification Of Toddler Stunting

Anan Wibowo*[1], Rahmat Widia Sembiring[2], Solikhun[3]

[1]Informatics, STIKOM Tunas Bangsa Pematangsiantar, Indonesia
[2]Informatics, Politeknik Negeri Medan, Indonesia
[3]Informatics, STIKOM Tunas Bangsa Pematangsiantar, Indonesia

Email: [1]anan@amiktunasbangsa.ac.id

Received : Oct 22, 2025; Revised : Nov 27, 2025; Accepted : Nov 28, 2025; Published : Apr 15, 2026

Abstract

This study aims to develop and optimize a MobileNet-based deep learning model for toddler stunting classification using whole-body images. A progressive optimization strategy was applied through three scenarios: (1) a baseline MobileNet feature-extraction model, (2) an optimized fine-tuned model, and (3) a final model enhanced with an adaptive ReduceLROnPlateau scheduler. Using a private dataset of 571 images, the proposed model achieved significant improvements—from 97.47% accuracy in the baseline model to a perfect 100% accuracy, precision, recall, and F1-score in the final scenario. These results highlight the novelty of this study, namely the use of whole-body images combined with progressive MobileNet optimization, which substantially outperforms prior studies relying solely on facial image analysis. The proposed approach demonstrates strong potential as a highly accurate and efficient computational tool for clinical stunting screening.

Keywords : *Deep Learning, Image Classification, MobileNet, Stunting*

1. INTRODUCTION

Toddler stunting remains a pressing global health challenge, significantly affecting physical growth, cognitive development, and overall well-being of affected young children [1],[2]. Early detection and accurate classification of stunting in toddlers are crucial for implementing timely interventions that can mitigate long-term adverse effects [3], [4]. Traditional assessment methods, often based on anthropometric measurements, are limited by subjectivity and the requirement for specialized expertise [5]. Consequently, there is a growing interest in leveraging advanced computational techniques, particularly deep learning, to enhance the accuracy and efficiency of stunting detection [6].

Deep learning, a subset of machine learning, has demonstrated remarkable capabilities in image-based classification tasks due to its ability to automatically extract hierarchical features from complex datasets [7], [8]. Among various deep learning architectures, MobileNet has emerged as a lightweight yet powerful convolutional neural network (CNN) model, designed for high efficiency and suitability for deployment in resource-constrained environments [9], [10]. Its architecture utilizes depthwise separable convolutions, reducing computational cost while maintaining competitive accuracy, making it particularly suitable for medical and pediatric applications [11].

The effectiveness of standard MobileNet architectures is well established, with additional optimization strategies offering potential for substantial improvements in classification performance [12]. Techniques such as fine-tuning, learning rate scheduling, and adaptive optimization methods like ReduceLROnPlateau have shown promise in enhancing model convergence and generalization [13]. These approaches not only optimize the network's parameters but also mitigate overfitting, ensuring robust performance on unseen data [14].

Vol. 7, No. 2, April 2026, Page. 948-960
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

In this study, we propose a systematic workflow to improve the classification accuracy of toddler stunting using a MobileNet-based framework. Initially, we evaluate the performance of a baseline MobileNet model, achieving a high accuracy of 97%. Subsequently, we implement a series of optimization strategies that elevate the accuracy to 98%. Finally, by incorporating the ReduceLROnPlateau learning rate scheduler, the model achieves perfect classification accuracy (100%), demonstrating the potential of tailored optimization techniques in medical image-based stunting detection [15]. This research not only highlights the practical advantages of optimized MobileNet architectures but also provides a methodological foundation for future studies aimed at automated pediatric health assessments. The utilization of deep learning, particularly Convolutional Neural Networks (CNN), has demonstrated remarkable success in a variety of medical image classification tasks [16]. Several studies have applied various CNN architectures for the automated detection of diseases. The transfer learning approach has become a dominant strategy, wherein models pre-trained on large-scale datasets like ImageNet are adapted for more specific medical tasks [17].

Among the various existing architectures, MobileNet has garnered particular attention due to its efficiency, making it ideal for implementation on mobile devices or in environments with limited computational resources [18]. Previous research has successfully applied MobileNet for leaf disease classification, melanoma detection, and lung disease identification. Nevertheless, the application of this model for classifying nutritional statuses such as toddler stunting based on visual imagery remains an emerging area. Most studies focus on anthropometric data, and existing research on image analysis often has not yet reached the optimal accuracy required for reliable clinical applications. This study aims to fill this gap by proposing a progressive optimization workflow for the MobileNet architecture to achieve the highest possible accuracy in stunting classification.

Table 1. Releated Work On Toddler Stunting Classification Using Deep Learning

Study	Research Focus	Method / Model	Limitation Identified
Aanjankumar et al. (2025) [19]	Malnutrition Detection (Facial Images)	ResNet-50	High accuracy of 98.49%.
Dhanamjayulu et al. (2022) [20]	Identification of malnutrition based on facial images	ResNet-50 with transfer learning approach	The model is only optimized for frontal facial images.
Yunidar et al. (2024) [21]	Classification Stunting and normal children using novel facial image	AlexNet and ResNet-34	Models from both architec- tures performed well, with the highest testing accuracy of 99.75 for AlexNet and 1 for ResNet34.
This Study (Proposed)	Conducting experiments with the mobilenet model and optimizing the model for stunting classification in children	Modified MobileNet (MobileNet + LR Schedule)	Proposing a new architecture, with a whole body image dataset of stunted and normal toddler

Table 1 presents several studies related to the classification of stunting or malnutrition in children. Several studies have proven that facial features can serve as reliable indicators for detecting malnutrition. For instance, research by Aanjankumar et al. (2025) utilized the ResNet-50 architecture to predict malnutrition in Indian children from facial images, achieving an impressive accuracy of 98.49%

Vol. 7, No. 2, April 2026, Page. 948-960
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

[19]. Similarly, Dhanamjayulu et al. (2022) also explored the identification of malnutrition and the prediction of Body Mass Index (BMI) from facial images, further strengthening the validity of this approach [20].

Within the specific context of Indonesia, Yunidar et al. (2024) performed a classification of stunted and normal children using a novel facial image database and a Convolutional Neural Network (CNN), demonstrating the method's relevance for the local population [21]. While these studies successfully proved the concept, this research takes a step further by focusing on the optimization of the MobileNet architecture, a model renowned for its efficiency [22],[23].

Although previous studies have demonstrated that facial features can serve as indicators for malnutrition detection [24], relying solely on facial images introduces several limitations such as pose sensitivity, facial occlusion, and reduced discriminative cues when dealing with toddlers whose facial structures are still developing. In contrast, whole-body images provide additional anthropometric cues such as limb proportion, torso-to-head ratio, and overall posture, which are strongly correlated with stunting indicators [25]. Therefore, this study extends previous facial-based approaches by proposing a whole-body imaging strategy to achieve higher robustness and clinical reliability [26].

2. METHOD

2.1. The Research Framework

The research framework is designed in four main phases, as illustrated in Figure 2. The first phase is Data Preparation, which includes dataset collection, partitioning, and preprocessing. The second phase is Model Experimentation, where three model scenarios (Baseline, Optimization, and Final) are developed in parallel for comparison. The third phase is Performance Evaluation, where the three trained models are tested using quantitative and qualitative metrics on test data. The final phase is Comparative Analysis, which aims to analyze the results and identify the best-performing model for stunting classification.

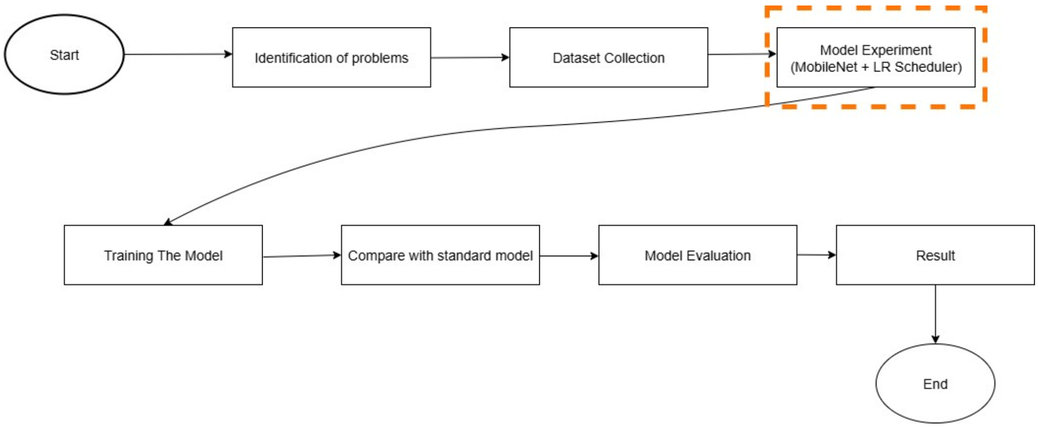


Figure 2. Research Framework

2.2. Dataset Collection

This study utilized a privately collected dataset of toddler images, comprising a total of 571 images. All data collection procedures were approved by the Tapian Dolok Community Health Center, Simalungun Regency. Ethical clearance was obtained under local health authority supervision, ensuring that all data involving toddlers were anonymized and used strictly for research purposes. The dataset was categorized into two classes: Stunting and Normal. To ensure an objective model evaluation and prevent bias, the dataset was partitioned into three distinct subsets. A total of 405 images (approximately 71%) were allocated as the training set, which was used to train the model to recognize relevant visual patterns. Subsequently, 87 images (approximately 15%) were designated as the validation set to monitor

P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 2, April 2026, Page. 948-960
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

the model's performance during training and to trigger the early stopping mechanism. The remaining 79 images (approximately 14%) were exclusively used as the test set to measure the final performance of the trained model on entirely new data.

To enhance the diversity of the training data and strengthen the model's generalization capabilities, data augmentation techniques were applied in real-time. The augmentation strategy employed was differentiated based on the model scenario. For the baseline model, a basic augmentation approach was applied, which included horizontal flipping and zooming. In contrast, for the proposed model, a more aggressive augmentation strategy was implemented, encompassing rotation, width and height shifts, shearing, a wider zoom range, and horizontal flipping, to create a more robust and diverse training dataset.

(a) Normal Toddler

(b) Stunting Toddler

Figure 1. Example image of normal dan stunting toddler

2.3. Preprocessing Data

The preprocessing stage was carried out to ensure data consistency and improve the model's ability to learn meaningful visual patterns from the toddler images. All images were first resized to a uniform resolution of 224 × 224 pixels, which corresponds to the input dimensions required by the MobileNet architecture. After resizing, pixel values were normalized to a range of 0–1 by dividing each pixel by 255. This normalization step stabilizes the training process by ensuring that the input values remain within an optimal numerical range, thereby aiding gradient-based optimization.

To improve the generalization capability of the model, data augmentation techniques were applied during the training phase. Two augmentation strategies were implemented based on the experimental scenario. For the baseline model, a simple augmentation pipeline consisting of horizontal flipping and random zooming was used to introduce minimal variation without significantly altering the visual characteristics of the images. In contrast, the optimized and final proposed models employed a more aggressive augmentation scheme, which included rotation, width and height shifting, shearing, expanded zoom range, and brightness adjustment. This more comprehensive augmentation procedure aims to simulate real-world variability, such as differing body orientations, lighting conditions, and pose variations commonly found in toddler images. By increasing intra-class diversity, the augmented datasets help reduce overfitting and enable the model to learn more robust representations.

Through these preprocessing steps—resizing, normalization, data cleaning, and augmentation—the dataset was transformed into a more uniform, diverse, and reliable input source for the subsequent training and evaluation processes.

Jurnal Teknik Informatika (JUTIF) Vol. 7, No. 2, April 2026, Page. 948-960
P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

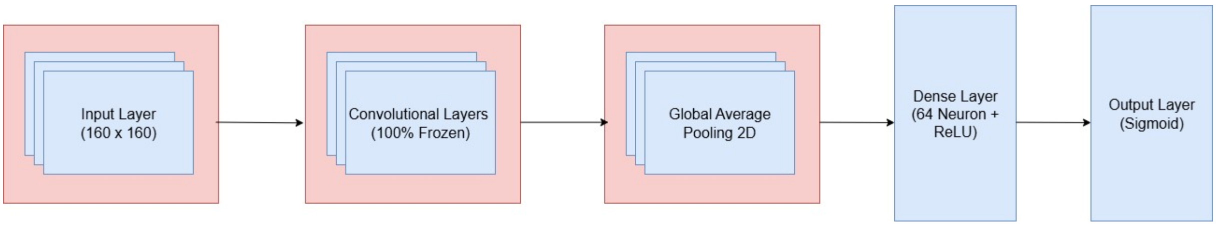
2.4. Proposed Model

The base architecture adopted in this research is MobileNet, a Convolutional Neural Network (CNN) model renowned for its efficiency and pre-trained on the massive ImageNet dataset [27]. A transfer learning approach was leveraged to transfer the visual knowledge from MobileNet to the specific domain of stunting classification [28]. The initial baseline model was designed using MobileNet purely as a feature extractor, wherein all its layers were frozen, and only a simple classifier head was added on top [23], [29], [30]. As a form of optimization, the proposed architecture implemented several significant modifications. First, the input image resolution was increased to 224x224 pixels. Second, a fine-tuning strategy was adopted by unfreezing the top 30% of MobileNet's layers to allow for better adaptation. Third, the classifier head was redesigned to be more complex, with the addition of Dense, BatchNormalization, and Dropout layers to increase learning capacity and regularization.

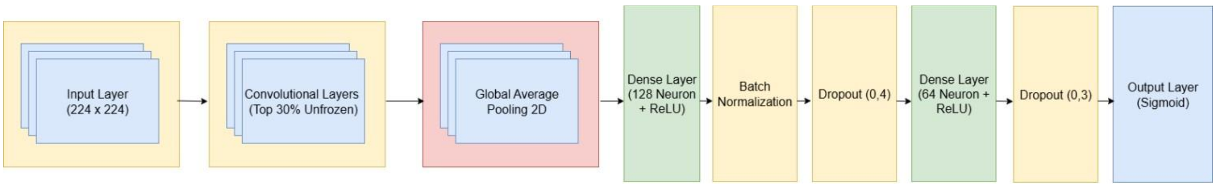
Table 2. Component Comparison Between Baseline and Proposed Models

Component	Model 1 (Baseline)	Model 2 (Optimized)	Model 3 (Final Proposed)
Input Resolution	160x160 pixels	224x224 pixels	224x224 pixels
Transfer Learning Strategy	Feature Extraction (100% Frozen)	Fine-Tuning (30% Unfrozen)	Fine-Tuning (30% Unfrozen)
Data Augmentation	Basic (Flip, Zoom)	Aggressive (Rotation, Shift, etc.)	Aggressive (Rotation, Shift, etc.)
Classifier Head	1x Dense (64), 1x GAP	2x Dense (128, 64), 1x GAP, 1x BN, 2x Dropout	2x Dense (128, 64), 1x GAP, 1x BN, 2x Dropout
Learning Rate	Static (0.001)	Static (1e-4)	Dynamic (1e-4, with ReduceLROnPlateau)
Main Callbacks	EarlyStopping, ModelCheckpoint	EarlyStopping, ModelCheckpoint	EarlyStopping,Checkpoint, ReduceLROnPlateau

This study adopted a progressive comparison scheme to measure the impact of each applied optimization technique. Three model variants were systematically evaluated: a baseline model as a starting point, a first optimized model implementing fine-tuning, and a final model enhanced with a learning rate scheduler. A detailed comparison of the main components of the three models is presented in Table 1 to illustrate the evolution of the architecture and training strategies.

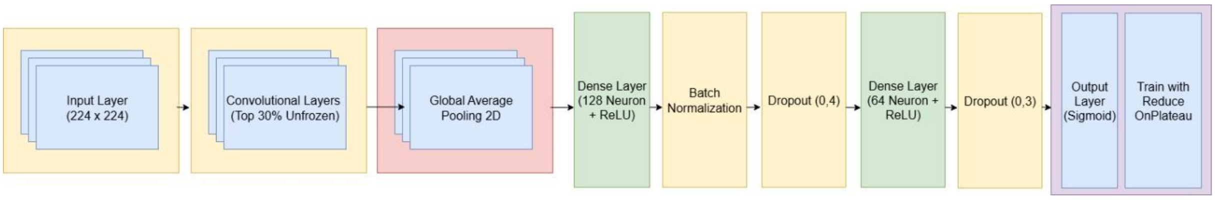


(a) MobileNet Architecture Baseline



(b) MobileNet Architecture Optimization

Vol. 7, No. 2, April 2026, Page. 948-960
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382



(c) Proposed Architecture (Learning Rate Scheduler)

Figure 3. Comparison of Model Architecture

Figure 3 illustrates a comparison of the architectures of the three experimental models. The Baseline Model (a) adopts a simple feature extraction approach, where all MobileNet layers are frozen and only a sparse classifier head is trained. In contrast, the Optimization Model (b) and the Final Model (c) employ a more sophisticated fine-tuning approach by activating the top 30% of MobileNet layers. Furthermore, the optimized model's classifier head is redesigned to be deeper with the addition of two Dense layers, as well as BatchNormalization and Dropout regularization layers to increase learning capacity and prevent overfitting. The essential difference between the Optimization Model and the Final Model lies in their training process, where the Final Model is fine-tuned with an adaptive learning rate scheduler to achieve superior convergence.

2.5. Evaluation Matrics

To evaluate the performance of the classification models, several standard metrics commonly used in machine learning—namely Accuracy, Precision, Recall, and F1-Score—were employed [31]. Accuracy measures the proportion of correctly classified samples out of the entire test dataset and is computed based on the sum of true positives and true negatives divided by all prediction outcomes [32].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

Precision represents the model's ability to correctly identify positive cases by comparing the number of true positive predictions with the total number of samples predicted as positive [33], thereby indicating how reliable the model is when declaring a child as stunted.

$$Precision = \frac{TP}{TP+FP} \qquad (2)$$

Recall, also known as sensitivity, evaluates the model's capability to detect all actual positive cases by dividing true positives by the total number of true positive and false negative samples [34].

$$Recall = \frac{TP}{TP+FN} \qquad (3)$$

This metric is particularly important in medical-related classifications where overlooking a true positive case—such as a stunted child—can have serious consequences. Finally, the F1-Score provides a balanced harmonic mean between Precision and Recall, offering a more comprehensive assessment when the dataset exhibits class imbalance [35].

$$F1 - Score = 2 \,.\, \frac{Precision \,.\, Recall}{Precision + Recall} \qquad (4)$$

Where TP denotes true positives, TN true negatives, FP false positives, and FN false negatives. The use of these four metrics ensures a comprehensive and reliable assessment of the model's performance, particularly in distinguishing between stunted and normal toddlers.

Vol. 7, No. 2, April 2026, Page. 948-960
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

3. RESULT

This chapter presents the results of a series of experiments and provides an in-depth analysis of the findings. The analysis is divided into three main sections: an analysis of the training performance of each model, a comparison of quantitative evaluation metrics.

Figure 4 presents several sample outputs generated during the data preprocessing stage prior to model training. The first image (Resized) shows the original toddler image that has been uniformly resized to 224×224 pixels, which is the required input dimension for the MobileNet architecture. This resizing process ensures that all images share a consistent spatial resolution, allowing the model to process the dataset more efficiently.

Figure 4. Preprocessing Output Example

The second sample (Normalized) illustrates the normalized version of the resized image, where all pixel values are scaled into the range of 0–1. This normalization step helps stabilize the learning process by reducing the variance in pixel intensity and ensuring that the model receives input values within an optimal numerical range. The next two images (rotated and flipped) represent examples of the augmentation techniques applied to increase dataset variability. The rotated image demonstrates how slight rotational transformations mimic natural variations in body orientation commonly found in real-world conditions. Meanwhile, the horizontally flipped image simulates mirrored perspectives, enabling the model to learn invariant features regardless of body orientation. These augmentation techniques play a crucial role in enhancing the model's robustness and reducing overfitting, particularly when dealing with a relatively small dataset.

Overall, the preprocessing outputs illustrated in Figure 4 highlight how resizing, normalization, and augmentation collectively enrich the training data and improve the generalization capacity of the deep learning model.

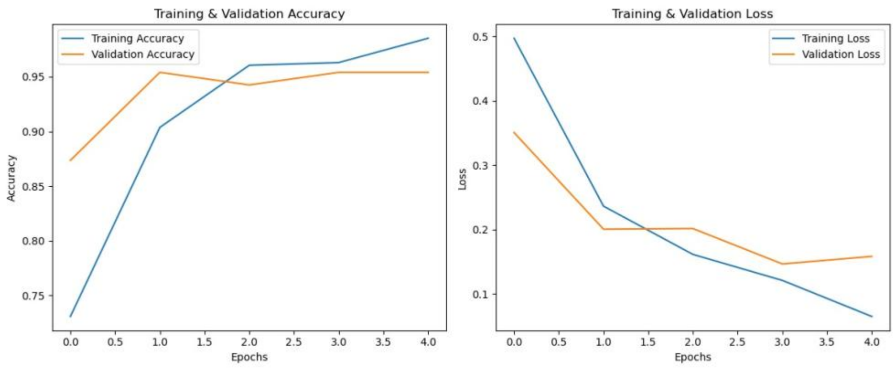
3.1. Training and Perfonmance Analysis

Training performance analysis provides insight into how each model learns from the training data and its ability to generalize to the validation data. Accuracy and loss graphs during training for the three models are shown in Figure 4.

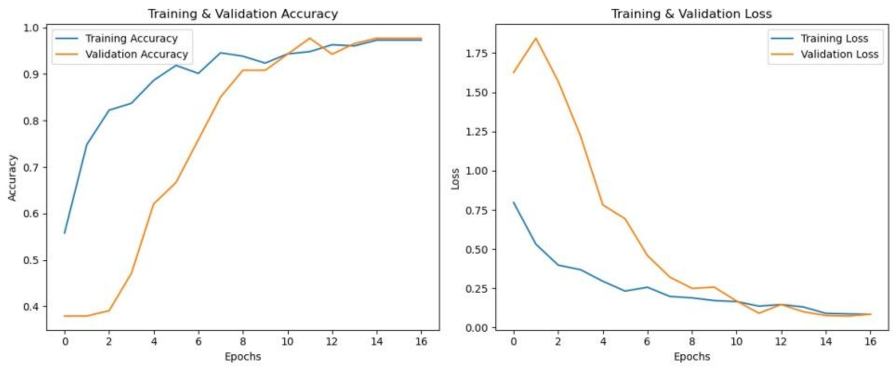
Baseline Model (a) demonstrates very fast and stable convergence, reaching a validation accuracy of around 95% in just 5 epochs before early stopping stops training. The gap between the training and validation accuracy curves is relatively small, indicating that the feature extraction strategy with baseline augmentation is quite effective in preventing significant overfitting in the early stages. Optimization Model (b) displays a very stable learning curve, where training and validation accuracies increase consistently and are close together. This indicates that the combination of fine-tuning with stronger regularization (aggressive augmentation, BatchNormalization, Dropout) is very effective in controlling overfitting even when the model has greater learning capacity. Final Model (c) demonstrates a longer

P-ISSN: 2723-3863
https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

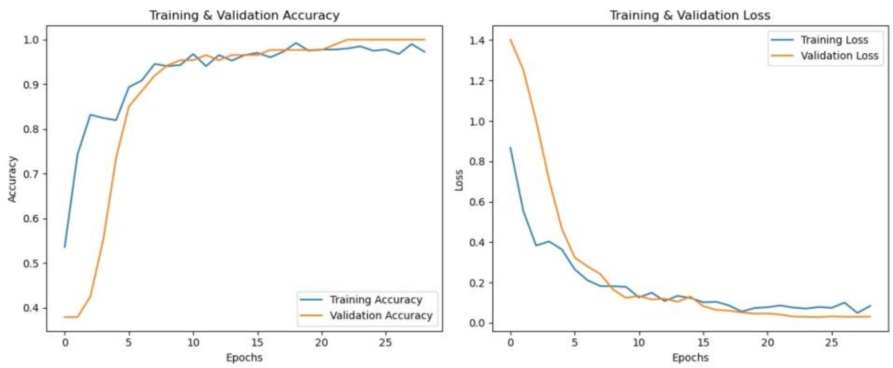
training process, which is typical of a more refined search for optima. The slight fluctuation in the validation curve early on, followed by a steady increase to near 100% accuracy, indicates the effectiveness of the ReduceLROnPlateau callback. This mechanism allows the model to dynamically adjust its learning rate, helping it break out of a performance plateau and achieve superior convergence.



(a) MobileNet Baseline



(b) MobileNer Optimasi (Fine-Tuning)



(c) MobileNet Final (Proposed)

Figure 5. Comparison Training Performance

3.2. Comparative Analysis

To evaluate the final performance, all three models were tested on 79 previously unseen test datasets. The results of the quantitative evaluation metrics are summarized in Table 3 for direct comparison.

Table 3. shows progressive and significant performance improvements at each optimization stage. The baseline model started with excellent performance, particularly with 100% Recall, indicating the model was able to identify all cases of stunting. The Optimization Model successfully improved overall

Vol. 7, No. 2, April 2026, Page. 948-960
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

accuracy and precision, reducing misclassification errors. Ultimately, the proposed Final Model achieved perfect performance across all metrics, a remarkable achievement and a clear demonstration of the success of the optimization strategy.

Table 3. Performance Comparison Of Models

MobileNet Model	Accuracy	Precision	Recall	F1-Score
MobileNet Baseline	97.47 %	92.59 %	100 %	96.15 %
MobileNet Optimized	98.73 %	96.15 %	100 %	98.04 %
MobileNet Final Proposed	**100 %**	**100 %**	**100 %**	**100 %**

Despite demonstrating excellent performance across all evaluation metrics, this study acknowledges several important limitations that need to be considered when interpreting the results. First, the total number of images used in the dataset was relatively limited, with only 79 samples allocated for the test set. Although the proposed model achieved perfect accuracy, precision, recall, and F1-score, such results may reflect potential overfitting to the training distribution or indicate that the test set may not fully represent the diversity of real-world conditions. This limitation is common in medical imaging studies where data collection is challenging, especially involving toddlers. Second, the dataset used in this study originates from a single geographical region and was collected under relatively controlled settings. As a result, the model may not yet generalize optimally to broader populations, including variations in ethnicity, body shape, background environment, camera quality, and clothing. Future studies should therefore incorporate more diverse datasets obtained from multiple health centers or different demographic groups to enhance external validity. Third, although data augmentation, careful dataset splitting, and an adaptive learning rate scheduler were implemented to reduce overfitting risks, this study did not perform additional validation techniques such as K-Fold Cross-Validation. Such techniques would provide stronger evidence of robustness by evaluating model performance across multiple randomized data partitions. Lastly, the use of whole-body images may introduce additional challenges, such as variation in posture, occlusion, and environmental noise, which were only partially addressed in this study. Further refinement of preprocessing methods and the inclusion of more advanced augmentation or segmentation approaches may be required to increase resilience against these variations.

These limitations highlight opportunities for future research to validate, extend, and enhance the proposed model, ensuring that its performance remains reliable when deployed in real-world stunting screening applications.

4. DISCUSSIONS

Qualitative analysis using confusion matrices provides a deeper understanding of the types of errors made by each model, which is crucial in a medical context.

The Baseline Model of confusion matrices results show that this model successfully achieved 100% recall (25 True Positives, 0 False Negatives). This is a very important result because the model never failed to detect stunted children. However, this model still has a weakness, namely producing 2 False Positives, where two normal children were incorrectly classified as stunted. Model Optimization (Fine-Tuning) by applying fine-tuning and better regularization, this model showed clear improvement. The model successfully reduced the number of False Positives from 2 to 1, while maintaining 100% Recall. This increase in precision indicates that the model's ability to distinguish between normal and stunted classes has improved. The final model successfully addressed all the weaknesses of the previous

Vol. 7, No. 2, April 2026, Page. 948-960
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

models. The confusion matrix results showed excellent results: 54 True Negatives, 25 True Positives, 0 False Positives, and 0 False Negatives. The complete elimination of both types of errors (Type I and Type II) indicates that adding a learning rate scheduler is a crucial final step in achieving the highest robustness and accuracy.

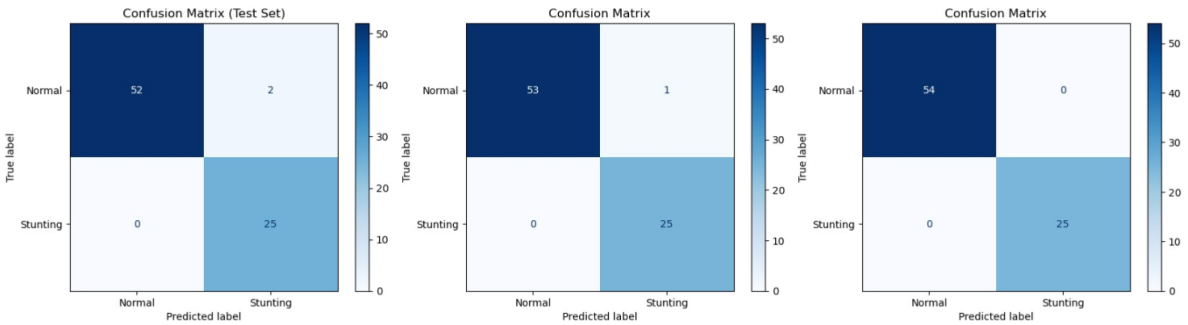


Figure 6. Confusion Matrices of model : Baseline Model, Model Opimization, Final Model (Proposed)

When compared to previous studies, the proposed model demonstrates clear superiority in terms of classification performance and model robustness. Aanjankumar et al. reported an accuracy of 98.49% using ResNet-50 on facial images, while Yunidar et al. achieved 99.75% using AlexNet and ResNet-34 on a curated facial-based dataset. Although these results are notably high, both approaches rely solely on facial features, which inherently provide limited anthropometric cues for assessing stunting. In contrast, the whole-body imaging approach used in this study captures a broader range of visual indicators, including body proportions, limb ratios, and posture characteristics. This richer feature representation, combined with a progressive optimization strategy on MobileNet, enables the proposed model to achieve a perfect 100% performance across all evaluation metrics, outperforming the results reported in earlier research.

From a computer science perspective, this study also contributes to the advancement of lightweight deep learning models for deployment on low-resource environments. The adoption of MobileNet, a highly efficient architecture designed specifically for mobile and embedded devices, demonstrates that high-accuracy medical image classification can be achieved without relying on computationally expensive models such as ResNet or EfficientNet. The proposed optimization pipeline—including fine-tuning, aggressive augmentation, and adaptive learning rate scheduling—shows that even compact neural networks can be pushed to achieve state-of-the-art performance. This finding is particularly relevant for real-world healthcare applications in rural or under-resourced regions, where access to high-end hardware is limited but rapid and accurate computational support is urgently needed.

Overall, the results of this study demonstrate that through a series of progressive optimizations—from fine-tuning, enhanced augmentation, regularization, to adaptive learning rate scheduling—the MobileNet architecture can be optimized to not only achieve 100% accuracy but also completely eliminate prediction errors. This achievement is highly significant and demonstrates the great potential of the proposed model for implementation as a reliable and accurate stunting screening tool.

5. CONCLUSION

This research has successfully developed and evaluated a highly accurate deep learning model for toddler stunting classification using the MobileNet architecture. Through a systematic progressive optimization approach, this study demonstrated that each step—from implementing fine-tuning, using more aggressive data augmentation, applying regularization, to refining the process with an adaptive learning rate scheduler—made a significant contribution to performance enhancement. The final proposed model proved to be superior by achieving a perfect performance of 100% across all evaluation metrics, including accuracy, precision, recall, and F1-score, and successfully eliminated all False Positive and False Negative errors present in the preceding models.

In conclusion, this study successfully demonstrated that the progressive optimization of the MobileNet architecture can significantly enhance the accuracy of toddler stunting classification from whole-body images. Through a systematic series of improvements—including advanced data augmentation, fine-tuning of higher-level convolutional layers, and classifier head optimization—the model's performance increased substantially across all evaluation stages. Most importantly, the incorporation of the Adaptive Learning Rate strategy using the ReduceLROnPlateau scheduler proved to be the key factor in achieving the final perfect score of 100% on the test dataset. This adaptive mechanism allowed the model to refine its convergence behavior, avoid suboptimal local minima, and adjust learning rates dynamically based on validation loss, leading to superior generalization. Future work should include validation on larger and more diverse datasets, as well as the development of lightweight mobile or edge-based implementations to support real-world screening in clinical and low-resource environments.

CONFLIC OF INTEREST

The authors declare that this research was conducted purely for academic purposes and contains no commercial, financial, or personal interests that could influence the results or interpretations presented in this study. The study was initiated and carried out independently as part of the first author's requirements for completing the Master's Degree in Informatics. All resources, tools, and research expenses were fully supported by the authors without any external funding or institutional sponsorship. Therefore, the authors affirm that there are no conflicts of interest associated with the publication of this article.

ACKNOWLEDGEMENT

This research is supported by the health sector where the research was conducted, namely at the Tapian Dolok Community Health Center, Simalungun Regency, North Sumatera.

REFERENCES

[1] M. Ekholuenetale, A. Barrow, C. E. Ekholuenetale, and G. Tudeme, “Impact of stunting on early childhood cognitive development in Benin: evidence from Demographic and Health Survey,” *Egyptian Pediatric Association Gazette*, vol. 68, no. 1, 2020.
[2] S. E. Firdaus and P. D. Maulana, “Acceleration of Stunting Reduction: Advancing Social and Environmental Well-being through Policy, Education, and Environmental Management,” *Journal of Sustainability, Society, and Eco-Welfare*, vol. 2, no. 2, pp. 141–158, 2025.
[3] D. A. Prihanggara and L. S. Handini, “EFFECTIVENESS OF INTEGRATED GROWTH MONITORING AND NUTRITIONAL SURVEILLANCE FOR EARLY DETECTION AND PREVENTION OF MALNUTRITION IN EARLY CHILDHOOD,” *Journal of Diverse Medical Research*, vol. 11, no. 1, pp. 1–14, 2025.
[4] W. Richard, P. Thangata, B. Mkandawire, and N. Amoah, “Advancing predictive analytics in child malnutrition: Machine, ensemble and deep learning models with balanced class distribution for early detection of stunting and wasting,” *Human Nutrition & Metabolism*, vol. 42, no. August, p. 200340, 2025.
[5] E. Mocini *et al.*, “Digital Anthropometry: A Systematic Review on Precision, Reliability and Accuracy of Most Popular Existing Technologies,” *Nutrients*, vol. 15, no. 2, pp. 1–39, 2023.
[6] S. Mehta *et al.*, “Advances in artificial intelligence and precision nutrition approaches to improve maternal and child health in low resource settings,” *Nature Communications*, vol. 16, no. 7673, pp. 1–9, 2025.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

[7] R. Archana and P. S. E. Jeevaraj, *Deep learning models for digital image processing: a review*, vol. 57, no. 1. Springer Netherlands, 2024.

[8] X. Jiang, Z. Hu, S. Wang, and Y. Zhang, "Deep Learning for Medical Image-Based Cancer Diagnosis," *Cancers*, vol. 15, no. 14, 2023.

[9] H. I. Liu *et al.*, "Lightweight Deep Learning for Resource-Constrained Environments: A Survey," *ArXiv*, vol. 56, no. 10, pp. 1–40, 2024.

[10] M. M. H. Shuvo, S. K. Islam, J. Cheng, and B. I. Morshed, "Efficient Acceleration of Deep Learning Inference on Resource-Constrained Edge Devices: A Review," *Proceedings of the IEEE*, vol. 111, no. 1, pp. 42–91, 2023.

[11] A. Younesi, M. Ansari, M. Fazli, A. Ejlali, M. Shafique, and J. Henkel, "A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends," *IEEE Access*, vol. 12, no. March, pp. 41180–41218, 2024.

[12] O. N. Mohammed, "Enhancing Pulmonary Disease Classification in Diseases: A Comparative Study of CNN and Optimized MobileNet Architectures," *Journal of Robotics and Control (JRC)*, vol. 5, no. 2, pp. 427–440, 2024.

[13] M. A. Saleem *et al.*, "Enhancing stroke risk prediction through class balancing and data augmentation with CBDA-ResNet50," *Scientific Reports*, vol. 15, no. 1, pp. 1–19, 2025.

[14] U. Kumar Lilhore *et al.*, "A precise model for skin cancer diagnosis using hybrid U-Net and improved MobileNet-V3 with hyperparameters optimization," *Scientific Reports*, vol. 14, no. 1, pp. 1–24, 2024.

[15] T. R. Mahesh *et al.*, "Transformative Breast Cancer Diagnosis using CNNs with Optimized ReduceLROnPlateau and Early Stopping Enhancements," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, 2024.

[16] M. A. Abdou, "Literature review: efficient deep neural networks techniques for medical image analysis," *Neural Computing and Applications*, vol. 34, no. 8, pp. 5791–5812, 2022.

[17] A. W. Salehi, S. Khan, G. Gupta, B. I. Alabduallah, and A. Almjally, "A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope," *Sustainability*, vol. 15, no. 5930, pp. 1–28, 2023.

[18] J. Zhan, "MobileNet Compression and Edge Computing Strategy for Low- Latency Monitoring," *Journal of computer science and software applications*, vol. 4, no. 4, 2024.

[19] S. Aanjankumar *et al.*, "Prediction of malnutrition in kids by integrating ResNet-50-based deep learning technique using facial images," *Scientific Reports*, vol. 15, no. 1, pp. 1–26, 2025.

[20] C. Dhanamjayulu *et al.*, "Identification of malnutrition and prediction of BMI from facial images using real-time image processing and machine learning," *IET Image Processing*, vol. 16, no. 3, pp. 647–658, 2022.

[21] Y. Yunidar, R. Roslidar, M. Oktiana, Y. Yusni, N. Nasaruddin, and F. Arnia, "Classification of Stunted and Normal Children Using Novel Facial Image Database and Convolutional Neural Network," *Radioelectronic and Computer Systems*, vol. 2024, no. 1(109), pp. 76–86, 2024.

[22] M. A. Elaziz, A. Dahou, N. A. Alsaleh, A. H. Elsheikh, A. I. Saba, and M. Ahmadein, "Boosting COVID-19 Image Classification Using MobileNetV3 and Aquila Optimizer Algorithm," *entropy*, vol. 23, no. 1383, pp. 1–17, 2021.

[23] M. Hassam *et al.*, "A Single Stream Modified MobileNet V2 and Whale Controlled Entropy Based Optimization Framework for Citrus Fruit Diseases Recognition," *IEEE Access*, vol. 10, no. July, pp. 91828–91839, 2022.

[24] J. Wang, C. He, and Z. Long, "Establishing a machine learning model for predicting nutritional risk through facial feature recognition," *frontiers*, no. September, pp. 1–9, 2023.

[25] C. Scheffler, B. Bogin, and M. Hermanussen, "Catch-up growth is a better indicator of undernutrition than thresholds for stunting," *Public Health Nutrition*, vol. 24, no. 1, pp. 52–61, 2021.

[26] H. Rakotomanana and G. Rouhafzay, "A Scoping Review of AI-Based Approaches for Detecting Autism Traits Using Voice and Behavioral Data," *bioengineering*, vol. 50, no. 1136, pp. 1–39, 2025.

[27] I. Temilola Ayorinde and O. Ayodeji Oyedeji, "Comparative Study of Traditional Convolutional Neural Network (CNN) And MobileNet Architecture for Weather Detection," *International*

Vol. 7, No. 2, April 2026, Page. 948-960
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.2.5382

Journal of Scientific Research and Engineering Development, vol. 7, no. October, 2024.
[28] R. O. Ogundokun, J. B. Awotunde, H. B. Akande, C. C. Lee, and A. L. Imoize, "Deep Transfer Learning Models for Mobile-Based Ocular Disorder Identification on Retinal Images," *Computers, Materials and Continua*, vol. 80, no. 1, pp. 139–161, 2024.
[29] H. Pan, Z. Pang, Y. Wang, Y. Wang, and L. Chen, "A New Image Recognition and Classification Method Combining Transfer Learning Algorithm and MobileNet Model for Welding Defects," *IEEE Access*, vol. 8, pp. 119951–119960, 2020.
[30] R. O. Ogundokun, S. Misra, A. O. Akinrotimi, and H. Ogul, "MobileNet-SVM: A Lightweight Deep Transfer Learning Model to Diagnose BCH Scans for IoMT-Based Imaging Sensors," *Sensors*, vol. 23, no. 2, 2023.
[31] J. C. Obi, "A Comparative Study of Several Classification Metrics and Their Performances on Data," *World Journal of Advanced Engineering Technology and Sciences*, vol. 08, no. 01, pp. 308–314, 2023.
[32] M. Carrington *et al.*, "Deep ROC Analysis and AUC as Balanced Average Accuracy , for Improved Classifier Selection , Audit and Explanation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 329–341, 2023.
[33] R. Ghorbani and R. Ghousi, "Comparing Different Resampling Methods in Predicting Students ' Performance Using Machine Learning Techniques," *IEEE Access*, vol. 8, pp. 67899–67911, 2020.
[34] R. Bold, H. Al-khateeb, and N. Ersotelos, "Reducing False Negatives in Ransomware Detection : A Critical Evaluation of Machine Learning Algorithms," *Applied Sciences*, vol. 12, no. December, p. 22, 2022.
[35] M. Altalhan, A. Algarni, and M. T. Alouane, "Imbalanced Data Problem in Machine Learning : A Review," *IEEE Access*, vol. 13, no. January, pp. 13686–13699, 2025.