



# Object Segmentation in Stunted Face Images using Deeplabv3+ with Resnet-50

Yunidar Yunidar \*, Melinda Melinda, Muhammad Irhamsyah

<sup>1</sup> Department of Electrical Engineering and Computer, Universitas Syiah Kuala, Banda Aceh, 23111, Indonesia.

## ARTICLE INFORMATION

Received: October 10, 2024

Revised: November 6, 2024

Accepted: November 25, 2024

Available online: November 30, 2024

## KEYWORDS

Stunting, segmentation, ResNet-50, accuracy

## CORRESPONDENCE

Phone: +6281397404546

E-mail: [yunidar@usk.ac.id](mailto:yunidar@usk.ac.id)

## A B S T R A C T

Stunting is the impaired growth and development that children experience from poor nutrition, repeated infection, and inadequate psychosocial stimulation. This study explores the impact of data preprocessing, specifically using DeepLabV3+ segmentation, on the performance of ResNet-50 in classifying stunting and non-stunting facial images. Initially, ResNet-50 achieved 99% accuracy and a 3.22% loss with the unsegmented dataset. By applying DeepLabV3+ to remove irrelevant features and backgrounds, the model's performance improved to a perfect 100% accuracy and a reduced loss of 0.45%. These results underscore the importance of high-quality data preprocessing in enhancing model precision and reliability. The findings have significant implications for practical applications, particularly in medical imaging, where improved diagnostic accuracy can benefit patient outcomes. Further research is recommended to explore additional preprocessing methods and their effects on model performance across diverse domains. This study highlights the transformative potential of effective data preprocessing in optimizing deep learning models for more accurate and reliable machine learning solutions.

## INTRODUCTION

Stunting is the impaired growth and development that children experience from poor nutrition, repeated infection, and inadequate psychosocial stimulation. If a child's height for age is more than two standard deviations below the median of the WHO Child Growth Standards, they are considered stunted [1]. Accurate identification and classification of stunting through facial image analysis can significantly aid early intervention and treatment. Our previous research successfully presented the classification results of stunting and normal facial images in children using a CNN architecture. This study investigates the role of advanced data preprocessing techniques, specifically. In recent years, deep learning techniques have demonstrated significant potential in automating classification tasks based on facial images [2]. Our previous research successfully presented the classification results of stunting and normal facial images in children using a CNN architecture. Facial recognition technology has advanced significantly, and identification applications such as convolutional neural networks (CNN) have demonstrated outstanding performance in facial recognition applications [3]. Here, we aim to increase the percentage of accuracy obtained in our previous research by using the segmentation method to see the facial characteristics of stunted children based on several facial recognition studies [4] [5] [6].

We created a facial recognition system to identify stunting in children using deep learning and segmentation methods to recognize stunting characteristics through facial images. The

<https://doi.org/10.25077/jnte.v13n3.1253.2024>

research utilizes the CRISP-DM methodology for data processing, as indicated in the study [3]. Various classification models were tested on the data, including Support Vector Machine - Linear Kernel (SVM-Linear), Support Vector Machine - RBF Kernel (SVM-RBF), Random Forest Classifier (RFC), AdaBoost (ADA), Logistic Regression (LR), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), K-Nearest Neighbor (KNN), Classification and Regression Trees (CART), Naive Bayes (NB) and Neural Network (MLPC) Study aims to optimize machine learning (ML) performance through classification analysis to determine nutritional status. The Multilayer Perceptron (MLP) algorithm and the Artificial Neural Network (ANN) approach are employed in the ML classification analysis process as detailed in the study [7].

The segmentation approach method used is the ResNet-50 and DeepLabV3+ architecture [8] [9] [10]. This strategy was chosen due to this architecture's proven robustness and performance in various computer vision tasks. ResNet-50 is one of the most widely used deep-CNN models for extracting attributes from facial images [11]. At the same time, DeepLabV3+ has been successful in tasks such as fabric defect detection and semantic segmentation. The model uses different backbone networks, such as ResNet-18, ResNet-50, and Mobilenetv2, and illustrates high performance on PASCAL VOC 2012 and Cityscape test sets [12]. This research carefully analyzes the performance of ResNet-50 and DeepLabV3+ in stunting and non-stunting recognition. The results will be examined to determine whether the Deeplabv3+ segmentation method can improve accuracy. The hope is that

[Attribution-NonCommercial 4.0 International](https://creativecommons.org/licenses/by-nc/4.0/). Some rights reserved

these methods can become a reference in classifying stunting and non-stunting. The decision to consider algorithms other than DeepLabv3+ for segmenting the stunting spectrum facial dataset arose from the need to identify the most appropriate method, considering factors such as algorithm compatibility, performance evaluation criteria, feature extraction requirements, and potential originality. Although DeepLabv3+ is effective in semantic segmentation tasks, its structure and specific features may not align with the dataset's characteristics [13]. This requires exploring alternative algorithms that may offer superior performance or new insights. By exploring uncharted territory using algorithms not previously used in this context, there is an opportunity to discover innovative methodologies and potentially improve our understanding of facial spectrum attributes in stunting, thereby driving progress in this field [14].

This research offers valuable contributions to the field by illustrating an effective segmentation strategy to maximize the accuracy of classifying stunted faces. Furthermore, it provides a comprehensive analysis of the performance of the ResNet50 model when implemented with semantic segmentation techniques.

**METHODS**

This study evaluates the performance of ResNet-50 in classifying stunting and normal datasets, both before and after implementing DeepLabv3+, as displayed in Figure 1. Starting with a comparative benchmark, where ResNet-50 serves as the baseline model for the stunting face dataset classification. ResNet-50 demonstrates commendable accuracy and efficiency in identifying facial features associated with stunting. However, to further improve the classification performance, this study incorporates DeepLabv3+ for segmentation. DeepLabv3+ segmentation enhances the model's understanding of facial structure and image spatial analysis. Ultimately, our combined approach outperforms stand-alone ResNet-50, highlighting the potential of advanced segmentation techniques for classification accuracy of stunting face datasets.

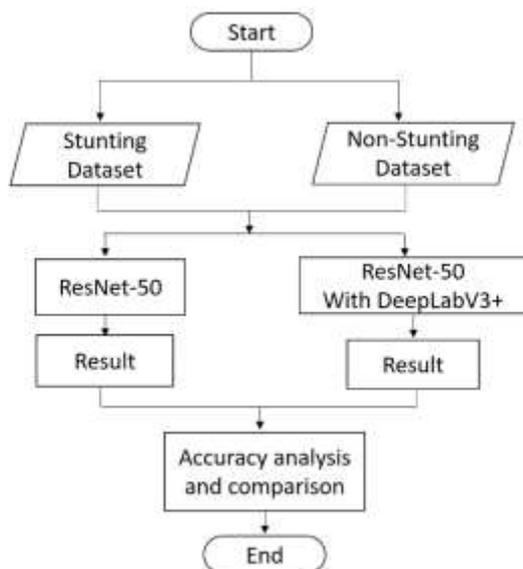


Figure.1 Proposed method

**ResNet-50**

ResNet-50 is a 50-layer convolutional neural network (CNN) architecture that is widely used in various computer vision tasks, including image classification and object detection [15]. Figure 2 provides a detailed architecture of ResNet-50. The architecture is based on residual learning, which allows the network to learn learning unreferenced functions [16]. ResNet-50 is a variant of the ResNet convolutional neural network with 50 layers. [17]. ResNet50 was previously trained on ImageNet, which is composed of a large number of around 1.5 million natural scene images [18]. This figure also demonstrates the significant layers of the major layers of ResNet-50, namely: 3 × 3 conv 64, 3 × 3 conv 128, 3 × 3 conv 256, 3 × 3 conv 512, feature vector map, etc [19]. The equations for the ResNet-50 model can be described using the following mathematical equation (1). Building Block:

$$f(X) = \sigma(W1\sigma(W1 X + B1) + B2) \tag{1}$$

where  $X$  is the input,  $W1$  and  $W2$  are the weight matrices,  $B1$  and  $B2$  are the bias terms, and  $\sigma$  is the ReLU activation function.

The residual connection can be determined by equation (2).

$$f(X) + X \tag{2}$$

where  $f(X)$  is the output of the convolutional layers in the building block, and  $X$  is the input to the block. The output value can be determined by equation (3)

$$\text{Output} = \text{Average Pooling} \times \text{Fully Connected Layer} \tag{3}$$

where Average PoolingAverage Pooling is the average pooling layer, and Fully Connected LayerFully Connected Layer is the fully connected layer for classification. These equations describe the basic components and structure of the ResNet-50 model. The specific implementation details may vary depending on the programming language and framework used.

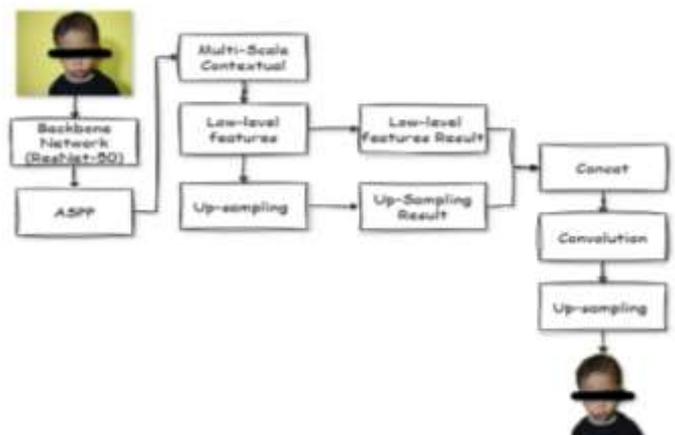


Figure 2 DeepLabV3+ framework.

**DeepLabV3+**

DeepLabv3+ is currently one of the best semantic segmentation algorithms available [20]. Deeplabv3+ is a semantic segmentation model designed to tackle the problem of segmenting complex images into different classes. It is an

advanced neural network architecture that builds upon the previous DeepLab models, offering enhanced accuracy and efficiency in segmenting fine details and object boundaries [21]. Deeplabv3+ uses ResNet-50 as its backbone network, a deep neural network architecture designed to address the vanishing gradient problem in intense networks [22]. Based on DeepLabv3, this technique forms an encoding-decoding structure by integrating a straightforward and efficient decoder. DeepLabV3+ consists of two main components: an encoder and a decoder. The encoder includes a backbone feature extraction network and an atrous spatial pyramid pooling (ASPP) structure. The decoder uses low-level features from the backbone feature extraction network and up-samples them to generate pixel-by-pixel classification results that match the size of the input image. [23].

### Dataset

The dataset was derived from the results of previous research [3] and comprises two categories: stunting and non-stunting. Examples of datasets from two categories can be seen in Figure 2. Each category contains an equal number of images, with 1.218 images in the stunting category and 1.218 images in the non-stunting category, making a total of 2.436 images. These images are focused on children's faces and were meticulously gathered by the author of this paper. The primary objective of collecting these images is to facilitate research into the physical manifestations and potential diagnostic indicators of stunting, a significant health issue characterized by impaired growth and development in children due to chronic malnutrition. By maintaining an equal number of images in both categories, the dataset ensures a balanced representation, which is crucial for training and evaluating machine learning models. This balance aids in reducing bias and enhances the accuracy and reliability of any predictive analysis or diagnostic tools developed using these datasets [24].



Figure 2 (a) Dataset of stunted children (b) Dataset of non-stunted children

### Accuracy

Scikit-learn manages the performance evaluation aspect, utilizing metrics based on the “Confusion Matrix” to assess the binary classification model’s performance. Key metrics contributing to the performance evaluation include: (i) False Positive (FP): when the data is negative but the model predicts positive; (ii) False Negative (FN): when positive data is predicted as negative by the model, and (iii) True Positive (TP): when both the data and the model predict positive. Accuracy measures the model's ability to function correctly [25], with the trained model reflecting the actual system performance. System performance is evaluated

using Accuracy, Recall, Precision, and F-Score metrics. These metrics are derived from the four possible combinations of actual and predicted values to assess the classification system's performance with two output classes [26].

- Accuracy:

$$\frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- Recall:

$$\frac{TP}{TP+FN} \quad (2)$$

- Precision:

$$\frac{TP}{TP+FP} \quad (3)$$

- F-score:

$$2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (4)$$

## RESULTS AND DISCUSSION

Previous research on object segmentation in facial images has used various approaches, such as Mask R-CNN for complex object segmentation, but is less efficient for data with high background variation [27] [28]. Additionally, methods like U-Net show good results in medical segmentation, but their performance decreases on non-medical images with non-uniform textures. In the context of stunting detection, research using CNN architecture often focuses on classification without optimizing the segmentation of relevant facial features [29] [30] [31]. In a prominent detection context, research using CNN often architecture focuses on classification without optimizing relevant facial feature segmentation. In this study, we used Deeplabv3+ with ResNet-50 algorithms. This approach allows the identification of more accurate facial features in surprising children with a complex background. These results expand previous research by providing high-quality segmentation, especially for facial analysis in a prominent context.

Here, tests will be conducted to assess the impact of data preprocessing, particularly DeepLabV3+ segmentation, on the performance of ResNet-50 in classifying stunted and non-stunted facial images, as illustrated in Figure 2. Using an unsegmented dataset, ResNet-50 achieved a high accuracy of 99% and a loss of 3.22%. These results demonstrate that ResNet-50 can effectively classify images even with extraneous features.

The model's performance improved significantly after applying the DeepLabV3+ segmentation technique, which removes irrelevant backgrounds and focuses on critical facial features. The accuracy reached 100%, and the loss dropped drastically to 0.45%. These enhancements in accuracy and loss metrics provide evidence of substantial improvements in model precision and reliability due to the refined dataset.

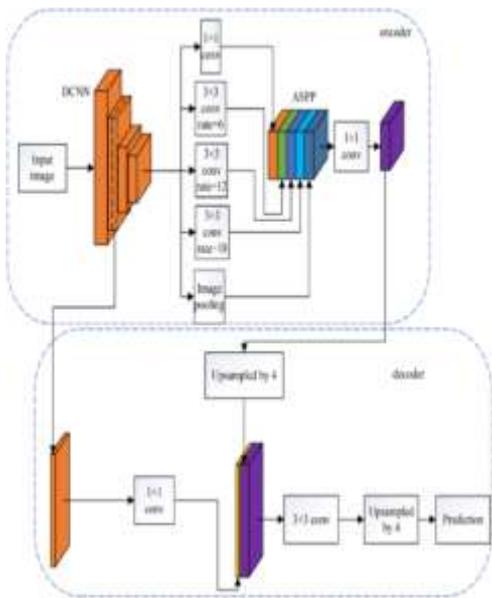


Figure 3. DeepLabV3+ architecture [32]

Table 1, Figure 1, and Figure 2 summarize the study's key findings. The results highlight the transformative impact of preprocessing on model performance, emphasizing that segmentation can significantly reduce errors and improve classification accuracy.

The primary hypothesis of this study was that preprocessing techniques, specifically DeepLabV3+ segmentation, would significantly enhance the performance of ResNet-50 in classifying stunting and non-stunting facial images. The findings support this hypothesis, demonstrating that segmentation can refine datasets by removing irrelevant features, allowing the model to focus on the critical aspects necessary for accurate classification. DeepLabV3+ has also been successfully applied in human face segmentation tasks, demonstrating high accuracy and efficiency in capturing fine details and object boundaries. ResNet-50 itself is one of the best models for image classification tasks. This study makes a unique contribution by being one of the first to apply ResNet-50 in combination with DeepLabV3+ segmentation for classifying stunting and non-stunting facial images. Previous research in this domain has typically employed other deep-learning architectures and preprocessing techniques. While these studies have demonstrated the importance of preprocessing and model selection in improving classification accuracy, none have specifically combined ResNet-50 with DeepLabV3+ segmentation for this particular task.

Earlier studies [3] have used simpler convolutional neural networks (CNNs) or other segmentation methods that do not offer the same level of sophistication or performance improvement as DeepLabV3+. These approaches have demonstrated moderate success in classifying medical images; however, they frequently require further enhancement to effectively manage the complexity and variability inherent in facial images associated with stunting and non-stunting conditions. By contrast, the approach of this study leverages the powerful residual learning capabilities of ResNet-50 and the advanced segmentation prowess of DeepLabV3+, setting a new benchmark in this field. The significant improvements in accuracy and loss metrics underscore the effectiveness of our methodology.

The implications of this study are significant both theoretically and practically. Theoretically, it reinforces the importance of advanced data preprocessing techniques in deep learning workflows. By demonstrating that DeepLabV3+ segmentation can lead to perfect classification accuracy, this study highlights the need for robust preprocessing to enhance model performance. This has broad applications across various domains where precise image classification is essential, such as medical diagnostics, facial recognition, and autonomous driving. The results suggest incorporating sophisticated preprocessing steps like segmentation can lead to more reliable and accurate machine-learning models. In medical imaging, for instance, improved model accuracy can enhance diagnostic capabilities, leading to better patient outcomes. The approach demonstrated in this study can be extended to other medical conditions and imaging techniques, potentially transforming diagnostic processes by providing more precise and reliable assessments. These practical implications underscore the potential impact of our research in real-world settings.

Table 1. Performance Metrics Before and After Segmentation

Metric	Before Segmentation	After Segmentation
Accuracy (%)	99	100
Loss (%)	3.22	0.45

### *Training and Validation using ResNet-5 before Segmentation*

The training and validation process for detecting stunting and non-stunting in facial images using ResNet-5 involved a well-prepared dataset of 1.218 images for each class. Despite its good quality, this dataset included unimportant features and unrelated elements before segmentation. These extraneous details could interfere with the model's ability to accurately learn the distinguishing characteristics of stunting versus non-stunting faces. ResNet-5, a simplified version of the deeper ResNet architectures, was chosen for this task due to its efficiency and effectiveness in handling image classification tasks. The model was trained on this dataset, and rigorous validation was conducted to ensure its performance and generalization capabilities. The training process involved feeding the images into the model, allowing it to learn the patterns and features associated with each class. The model's predictions were compared against the actual labels during validation to evaluate its accuracy and loss.

Remarkably, the training and validation yielded an impressive accuracy of 99% and a loss of 3.22%. This high accuracy indicates that the model could distinguish between stunting and non-stunting faces precisely. The relatively low loss further supports the model's reliability and minimizes prediction errors. However, it's important to note that these results were achieved before the segmentation of the images. The presence of unrelated and unimportant features in the dataset likely presented additional challenges, making the high performance even more noteworthy. Segmentation, which involves isolating the relevant parts of the images, would likely enhance the model's performance by eliminating noise and focusing on the critical features needed for accurate classification. This step would streamline the dataset, ensuring that the model is trained on the most pertinent

information, potentially improving accuracy and loss metrics further. This result's plot is displayed in Figures 4 and 5.

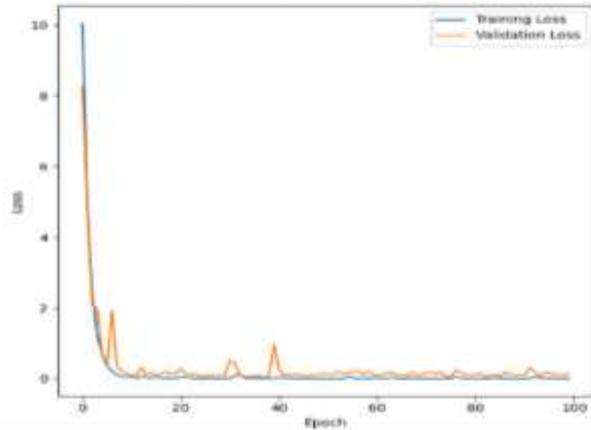


Figure 4. The loss curve before segmentation

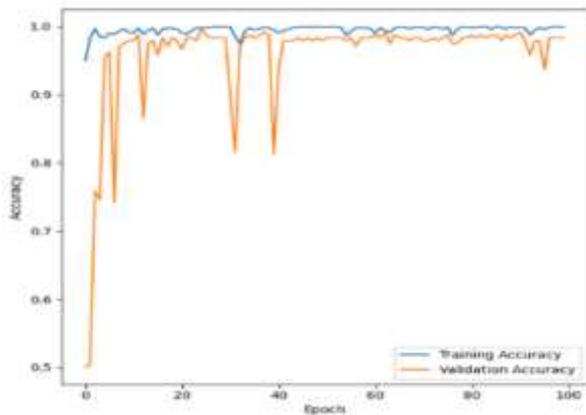


Figure 5. The accuracy curve before segmentation

### ***Training And Validation using ResNet-50 with segmentation using DeepLabV3+***

Figures 6 and 7 below show the loss and accuracy curves, which are the relationship between epoch and loss using the ResNet-50 architecture after segmentation. In the second phase of training and validation, the same ResNet-50 model and methodology were employed, but with a crucial addition: the segmentation technique using DeepLabV3+. This advanced segmentation approach was utilized to remove backgrounds and extraneous features from the dataset, isolating the relevant facial regions essential for the study. By applying DeepLabV3+, the dataset was refined to focus solely on the pertinent aspects of the images, eliminating noise and unrelated elements that could interfere with the model's learning process.

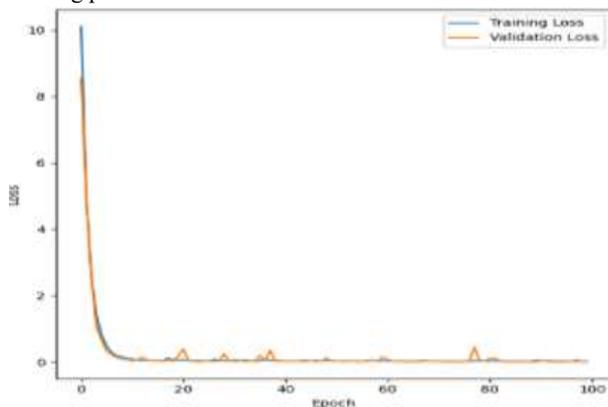


Figure 6. The loss curve after segmentation

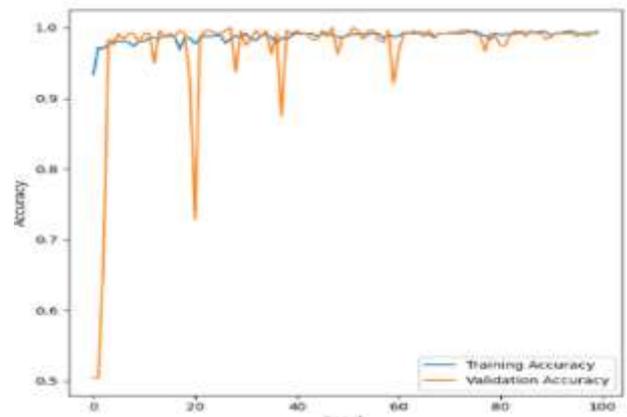


Figure 7. The accuracy curve after segmentation

As a result of this segmentation, the model's accuracy increased to an impressive 100%, and the loss decreased to just 0.45%. This significant improvement in performance underscores the importance of data preprocessing in machine learning tasks. The perfect accuracy indicates that the model could distinguish between stunting and non-stunting faces without errors. Similarly, the markedly lower loss substantially reduces prediction errors, demonstrating the model's enhanced reliability and precision.

This outcome highlights the efficacy of segmentation techniques like DeepLabV3+ in improving the dataset's quality. By removing unnecessary features, the model could focus more effectively on the critical characteristics required for accurate classification. This simplifies the learning process and enhances the model's overall performance, as evidenced by the perfect accuracy and minimal loss in this second training and validation phase. The plot of these results is illustrated in Figure 6 and Figure 7.

## **CONCLUSIONS**

Object segmentation using DeepLabV3+ with ResNet-50 on stunted children's facial images significantly improves the model's classification performance. By utilizing segmentation to highlight relevant features on facial images, the model achieves high accuracy and lower loss than data without segmentation. These results underscore the importance of segmentation-based data preprocessing in enhancing the effectiveness of the machine learning pipeline.

This approach has important implications, particularly in medical image analysis, with the potential to improve diagnostic accuracy. Additionally, this method opens up opportunities for application in various domains that require high-precision image analysis, such as facial recognition and other computer vision-based systems.

## **ACKNOWLEDGMENT**

The authors thank the Universitas Syiah Kuala for supporting this study.

## REFERENCES

- [1] C. Scheffler *et al.*, “Stunting is not a synonym of malnutrition,” *Eur. J. Clin. Nutr.*, vol. 74, no. 3, pp. 377–386, 2020, Accessed: Aug. 10, 2024. [Online]. Available: <https://www.nature.com/articles/s41430-019-0439-4>
- [2] S. K. Zhou *et al.*, “A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises,” *Proc. IEEE*, vol. 109, no. 5, pp. 820–838, 2021, Accessed: Aug. 10, 2024. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/9363915/>
- [3] 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,” *Radioelectron. Comput. Syst.*, vol. 2024, no. 1, pp. 76–86, 2024, Accessed: Jul. 17, 2024. [Online]. Available: <http://nti.khai.edu/ojs/index.php/reks/article/view/reks.2024.1.07>
- [4] Y. Liu, X. Bai, J. Wang, G. Li, J. Li, and Z. Lv, “Image semantic segmentation approach based on DeepLabV3 plus network with an attention mechanism,” *Eng. Appl. Artif. Intell.*, vol. 127, p. 107260, 2024, Accessed: Aug. 10, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0952197623014446>
- [5] B. M. Рувінська and Ю. Ю. Тімков, “Deep learning technology for video frame processing in face segmentation on mobile devices,” *Вісник Сучасних Інформаційних Технологій*, vol. 4, no. 2, pp. 185–194, 2021, Accessed: Aug. 10, 2024. [Online]. Available: <https://hait.od.ua/index.php/journal/article/view/109>
- [6] S. Patel, “Deep learning models for image segmentation,” in *2021 8th International conference on computing for sustainable global development (INDIACom)*, IEEE, 2021, pp. 149–154. Accessed: Aug. 10, 2024. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/9441388/>
- [7] A. Y. Agyeman and S. G. Tetteh, “Technical Evaluation of Machine Learning Models: An Empirical Study,” *ESP Int. J. Adv. Comput. Technol. ESP-IJACT*, vol. 2, no. 1, pp. 1–9, Jan. 2024, Accessed: Nov. 05, 2024. [Online]. Available: <https://www.espjournals.org/IJACT/ijact-v2i1p101>
- [8] Z. H. U. Runhu, X. I. N. Binjie, D. Na, and F. A. N. Mingzhu, “Semantic segmentation using deeplabv3+ model for fabric defect detection,” *Wuhan Univ. J. Nat. Sci.*, vol. 27, no. 6, pp. 539–549, 2022, Accessed: Aug. 10, 2024. [Online]. Available: <https://wujns.edpsciences.org/articles/wujns/abs/2022/06/wujns-1007-1202-2022-06-0539-11/wujns-1007-1202-2022-06-0539-11.html>
- [9] H. Polat, “A modified DEEPLABV3 + based semantic segmentation of chest computed tomography images for COVID -19 lung infections,” *Int. J. Imaging Syst. Technol.*, vol. 32, no. 5, pp. 1481–1495, Sep. 2022, doi: 10.1002/ima.22772.
- [10] S. O. Atik, M. E. Atik, and C. Ipbuker, “Comparative research on different backbone architectures of DeepLabV3+ for building segmentation,” *J. Appl. Remote Sens.*, vol. 16, no. 2, pp. 024510–024510, 2022, Accessed: Nov. 05, 2024. [Online]. Available: <https://www.spiedigitallibrary.org/journals/journal-of-applied-remote-sensing/volume-16/issue-2/024510/Comparative-research-on-different-backbone-architectures-of-DeepLabV3-for-building/10.1117/1.JRS.16.024510.short>
- [11] H. Polat, “Multi-task semantic segmentation of CT images for COVID-19 infections using DeepLabV3+ based on dilated residual network,” *Phys. Eng. Sci. Med.*, vol. 45, no. 2, pp. 443–455, Jun. 2022, doi: 10.1007/s13246-022-01110-w.
- [12] M. A. Butt and F. Riaz, “CARL-D: A vision benchmark suite and large scale dataset for vehicle detection and scene segmentation,” *Signal Process. Image Commun.*, vol. 104, p. 116667, 2022, Accessed: Aug. 11, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0923596522000224>
- [13] M. Cai *et al.*, “Image segmentation method for sweetgum leaf spots based on an improved DeeplabV3+ network,” *Forests*, vol. 13, no. 12, p. 2095, 2022, Accessed: Aug. 11, 2024. [Online]. Available: <https://www.mdpi.com/1999-4907/13/12/2095>
- [14] T. Ashine, “Image-Based Rose Leaf Diseases Detection Using Deep Learning,” Thesis, St. Mary’s University, 2024. Accessed: Aug. 11, 2024. [Online]. Available: <http://repository.smuc.edu.et/handle/123456789/7886>
- [15] B. Li and D. Lima, “Facial expression recognition via ResNet-50,” *Int. J. Cogn. Comput. Eng.*, vol. 2, pp. 57–64, 2021, Accessed: Aug. 11, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2666307421000073>
- [16] G. Singh, A. Mittal, and N. Aggarwal, “ResDNN: deep residual learning for natural image denoising,” *IET Image Process.*, vol. 14, no. 11, pp. 2425–2434, Sep. 2020, doi: 10.1049/iet-ipr.2019.0623.
- [17] A. Amelio *et al.*, “Representation and compression of Residual Neural Networks through a multilayer network based approach,” *Expert Syst. Appl.*, vol. 215, p. 119391, 2023, Accessed: Aug. 11, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0957417422024095>
- [18] H. Tahir, M. S. Khan, and M. O. Tariq, “Performance analysis and comparison of faster R-CNN, mask R-CNN and ResNet50 for the detection and counting of vehicles,” in *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, IEEE, 2021, pp. 587–594. Accessed: Aug. 11, 2024. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/9397079/>
- [19] M. M. Murimi, “Identification Of Maize Leaf Diseases Using Support Vector Machine And Convolutional Neural Networks Alexnet And Resnet50,” Phd Thesis, Karatina University, 2023. Accessed: Jul. 16, 2024. [Online]. Available: <https://karuspace.karu.ac.ke/bitstream/handle/20.500.12092/2999/FULLTEXT.pdf?sequence=1>
- [20] P. Venugopal, P. K. R. Maddikunta, T. R. Gadekallu, A. Al-Rasheed, M. Abbas, and B. O. Soufiene, “An adaptive DeepLabv3+ for semantic segmentation of aerial images using improved golden eagle optimization algorithm,” *IEEE Access*, 2023, Accessed: Aug. 11, 2024. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10261990/>
- [21] Y. Wang, C. Wang, H. Wu, and P. Chen, “An improved Deeplabv3+ semantic segmentation algorithm with multiple loss constraints,” *PLOS ONE*.
- [22] N. Siddique, “Estimation of Road Boundary for Intelligent Vehicles Based on DeepLabV3+ Architecture,” vol. 9, 2021.
- [23] C. Cai, J. Tan, P. Zhang, Y. Ye, and J. Zhang, “Determining Strawberries’ Varying Maturity Levels by Utilizing Image Segmentation Methods of Improved DeepLabV3+,” *Agronomy*, vol. 12, no. 8, p. 1875, Aug. 2022, doi: 10.3390/agronomy12081875.
- [24] Y. Gong, G. Liu, Y. Xue, R. Li, and L. Meng, “A survey on dataset quality in machine learning,” *Inf. Softw. Technol.*, vol. 162, p. 107268, 2023, Accessed: Aug. 11, 2024. [Online]. Available: <https://doi.org/10.25077/jnte.v13n3.1253.2024>

<https://www.sciencedirect.com/science/article/pii/S0950584923001222>

- [25] A. W. Salehi *et al.*, “A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope,” *Sustainability*, vol. 15, no. 7, p. 5930, 2023, Accessed: Aug. 11, 2024. [Online]. Available: <https://www.mdpi.com/2071-1050/15/7/5930>
- [26] M. Muntean and F.-D. Militaru, “Metrics for Evaluating Classification Algorithms,” in *Education, Research and Business Technologies*, vol. 321, C. Ciurea, P. Pocatilu, and F. G. Filip, Eds., in Smart Innovation, Systems and Technologies, vol. 321. , Singapore: Springer Nature Singapore, 2023, pp. 307–317. doi: 10.1007/978-981-19-6755-9\_24.
- [27] M. Melinda *et al.*, “Image Segmentation Performance using Deeplabv3+ with Resnet-50 on Autism Facial Classification,” *J. INFOTEL*, vol. 16, no. 2, Art. no. 2, May 2024, doi: 10.20895/infotel.v16i2.1144.
- [28] C. Sahaya Pushpa Sarmila Star, T. M. Inbamalar, and A. Milton, “Automatic semantic segmentation of breast cancer in DCE-MRI using DeepLabV3+ with modified ResNet50,” *Biomed. Signal Process. Control*, vol. 99, p. 106691, Jan. 2025, doi: 10.1016/j.bspc.2024.106691.
- [29] D. Xiangwu, L. Song, Q. Long, and Y. Shuting, “Method study on semantic segmentation of weeds at seedling stage in paddy fields based on DeepLabV3+ model,” *J. Chin. Agric. Mech.*, vol. 44, no. 4, p. 174, Apr. 2023, doi: 10.13733/j.jcam.issn.2095-5553.2023.04.024.
- [30] H. Polat, “A modified DeepLabV3+ based semantic segmentation of chest computed tomography images for COVID-19 lung infections,” *Int. J. Imaging Syst. Technol.*, vol. 32, no. 5, pp. 1481–1495, 2022, doi: 10.1002/ima.22772.
- [31] R. Preethi, R. Sudharsan, P. Swetha, C. Rajasekaran, and S. Kausalya, “Deep Learning-Based Semantic Segmentation of Turmeric Crop Images Using the DeepLabV3+ Architecture,” in *2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)*, IEEE, 2024, pp. 1–6. Accessed: Nov. 28, 2024. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10531416/>
- [32] J. Fu, X. Yi, G. Wang, L. Mo, P. Wu, and K. E. Kapula, “Research on ground object classification method of high resolution remote-sensing images based on improved DeeplabV3+,” *Sensors*, vol. 22, no. 19, p. 7477, 2022, Accessed: Aug. 10, 2024. [Online]. Available: <https://www.mdpi.com/1424-8220/22/19/7477>
- [33] C. S. P. S. Star, T. M. Inbamalar, and A. Milton, “Automatic semantic segmentation of breast cancer in DCE-MRI using DeepLabV3+ with modified ResNet50,” *Biomed. Signal Process. Control*, vol. 99, p. 106691, 2025, Accessed: Nov. 05, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1746809424007493>

## NOMENCLATURE

The meaning of symbols used in the equations and other symbols presented in your article must be presented in this section.

$\sigma$	ReLU activation
$W$	weight matrices
$B$	the bias terms

## AUTHORS BIOGRAPHY

Yunidar was born in Banda Aceh, Aceh, on June 10, 1974. She has been a lecturer at the Faculty of Engineering, Department of Electrical Engineering, University of Syiah Kuala, since March 2000. After completing her bachelor of science degree in Physics from Syiah Kuala University, Aceh, Indonesia, in 1997, she then acquired her master's degree in engineering (M. Eng.) in Optoelectrotechniques and Laser Applications from the University of Indonesia, Jakarta, Indonesia, in 2000. She is also a member of IEEE. Her research interests involve performing biomedical engineering, and sensors used in biomedical applications include multimedia. She can be contacted at: [yunidar@usk.ac.id](mailto:yunidar@usk.ac.id).

Melinda was born in Bireuen, Aceh, on June 10, 1979. She received B.Eng degree from the Department of Electrical Engineering, Faculty of Engineering, Universitas Syiah Kuala, Banda Aceh in 2002. She completed her master's degree at the Faculty of Electrical Department, University of Southampton, United Kingdom, with a concentration in the field study of Radio Frequency Communication Systems in 2009. She completed her Doctoral degree at the Department of Electrical Engineering, Engineering Faculty of Universitas Indonesia in February 2018. She has been with the Department of Electrical Engineering, Faculty of Engineering, Universitas Syiah Kuala since 2002. She is also a member of IEEE. Her research interests include multimedia signal processing and fluctuation processing. She can be contacted at email: [melinda@usk.ac.id](mailto:melinda@usk.ac.id).

Muhammad Irhamsyah was born in Banda Aceh, July 18, 1972. He is currently a lecturer in the Department of Electrical Engineering, Faculty of Engineering, Syiah Kuala University. Joined Syiah Kuala University since 2001 in Telecommunication Engineering. Graduated from S1 Department of Electrical Engineering ITS Surabaya in 1998 and S2 from the Department of Electrical Engineering University of Indonesia in 2008 with the field of Telekomunikasi Engineering expertise. Has a research interest in wireless telecommunication and antenna. He currently serves as Chairman of Telecommunication Engineering. He can be contacted at email: [irham.ee@usk.ac.id](mailto:irham.ee@usk.ac.id).