



Performance Comparison of Facial Skin Type Classification Using the Segment Anything Model

Nisrina Nur Kumala^{1*}, Asfan Muqtadir², Amaludin Arifia³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknik,
Universitas PGRI Ronggolawe, Indonesia

E-Mail: ¹nisrinanurkumala@gmail.com, ²asfanme@unirow.ac.id, ³amaludinarifia@gmail.com

Received Feb 26th 2026; Revised Mar 15th 2026; Accepted Mar 26th 2026; Available Online Apr 19th 2026

Corresponding Author: Nisrina Nur Kumala

Copyright ©2026 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

The facial skin is the first area to often experience various problems. Knowing one's skin type is an important step in choosing the right skincare routine, but it can be difficult to determine accurately without a specialist's help, which can be costly. To address this, a deep learning approach can be applied to help automatically classify skin types. In this study, several combinations of CNN, MobileNetV3, and SAM models were applied and compared for facial skin type classification. The dataset used, sourced from the figshare platform, consists of 2,250 facial images representing 5 skin types: normal, dry, oily, sensitive, and combination. The dataset was divided into three parts: training (80%), validation (10%), and testing (10%). Each model was evaluated using a confusion matrix, with accuracy, precision, recall, and F1-score metrics used to determine and compare model performance. The results show that the CNN performed worst, while the MobileNetV3-based CNN was the best-performing model, achieving an accuracy of 97%. Meanwhile, adding SAM did not improve performance and actually decreased accuracy. This study demonstrates that using MobileNetV3 without segmentation is more effective than adding SAM segmentation for facial skin type classification.

Keywords: CNN, Deep Learning, MobileNetv3, Segment Anything Model, Skin Type Classification

1. INTRODUCTION

The skin is the first defense against external factors and threats [1]. Indonesia's geographical conditions are in the tropics with high sun intensity, which makes the skin susceptible to sun exposure [2]. Sunlight contains ultraviolet (UV) radiation that is beneficial to the skin, but excessive exposure to UV rays can cause skin problems [3]. In addition, other factors such as pollution and hormonal changes can also cause changes in facial skin [4].

The face is the first area that is most visible. Clean and well-groomed facial skin can increase self-confidence and have a great influence on self-appearance, especially for women [5]. In order for facial care to work well in dealing with the impact of the environment as well as the impact from within, an important first step is to know the type of skin that each individual has [6]. There are five types of facial skin, namely normal skin, sensitive skin, combination skin, oily skin, and dry skin [7]. The use of care products should be appropriate for each individual's facial skin type so as not to cause problems such as acne, irritation, or excessive dry skin [8].

Most individuals have difficulty in correctly recognizing their own skin type [9]. A beauty specialist and a skin care place are needed to determine the type of facial skin as the basis for using the appropriate care products for each individual. All of this requires relatively high consultation costs and obstacles for each individual to take time in the midst of a busy schedule [10].

In the modern era, the development of artificial intelligence (AI) technology is driving innovation across various fields. One branch of AI is deep learning, which can automatically learn and recognize features of complex images with greater accuracy [11]. Deep learning-based image classification is an effective solution for addressing challenges in understanding facial skin type and can support beauty analysis in the digital era [12]. One approach for studying image features is the convolutional neural network (CNN). CNNs can detect or recognize objects in digital images by understanding the structure and patterns of the images [13]. Despite excelling at feature extraction, CNNs require substantial computational resources [14]. In addition, the main challenges include differences in lighting, similarities in texture between classes, and the model's limited ability to generalize across diverse datasets [15]. Therefore, more efficient architectures,

such as MobileNetV3, were developed to produce models that are lightweight yet high-performing in image recognition [14]. In addition, the addition of segment anything model (SAM) as a segmentation stage can help the model to focus on important parts of the image through an automated segmentation process [16], but it is necessary to evaluate the influence on model performance.

Several previous studies have shown the effectiveness of this approach. Classification of bladder cancer images with CNN without the addition of other architectures was proven to be able to achieve high performance [17], then effective classification of corn leaf disease with high accuracy of 93% showed that CNN was capable of image data classification, but the performance of the model could still be improved through comparison with modern CNN architectures [18], then the development of facial skin type detection applications achieved 87% accuracy but was limited to 3 facial skin types [19], then CNN comparative performance analysis achieved 66% accuracy and MobileNetV3 achieved 84% accuracy which proves that MobileNetV3 is superior and has higher computational efficiency in classifying facial skin disorders [14], then classification of plant diseases with SAM as a segmentation stage in the CNN model can increase the accuracy by more than 10% [20], while in the classification of medical images with the SAM model it is 15-30% lower compared to the specific domain model [21].

Based on the results of several previous studies, CNN has proven to be effective in processing and classifying image data, including in the medical and agricultural fields with varying degrees of accuracy with limitations in computational efficiency. However, performance improvements are still possible through the addition of other architectures, such as MobileNetV3. In addition, the integration of segmentation stages using SAM can improve model performance in the classification of plant diseases. However, in the classification of medical images, the effectiveness of SAM use also leads to a decrease in performance. This requires the appropriateness of segmentation in the dataset so as not to have an impact on the decline in the performance of the classification model.

Although there has been considerable research on the classification of facial skin types, most still use a limited number of classes or focus only on one model architecture without making comparisons with different architectures. In addition, the use of SAM segmentation and its influence in the classification of facial skin types has not been widely studied in previous studies. Therefore, further research is needed to examine the performance with segmentation stages on the facial skin type dataset. These limitations are the basis for this study with the aim of comparing the performance of the combination of CNN, MobileNetV3, and SAM architectures in classifying facial skin types. Comparisons were made to determine the best-performing model based on the confusion matrix. This is expected to be the basis for the development of a facial skin type determination system for the proper skin care process.

2. MATERIALS AND METHOD

This study applies the structured stages in Figure 1, namely data splitting, pre-processing, training, validation, and evaluating. In the training and evaluating stage, CNN, MobileNetV3 and SAM were tested in 3 different model variations to produce the best performing model. The first model, conventional CNN, is a basic model for measuring performance without modification of architecture or segmentation techniques. The second model, CNN is based on MobileNetV3 to produce a lighter model but still has optimal feature extraction capabilities. Meanwhile, the third model combines MobileNetV3-based CNN with SAM as a segmentation stage so that the model can learn from more specific features of the facial area. At the evaluation stage, the model was assessed using a confusion matrix with several metrics such as accuracy, precision, recall, and F1-score.

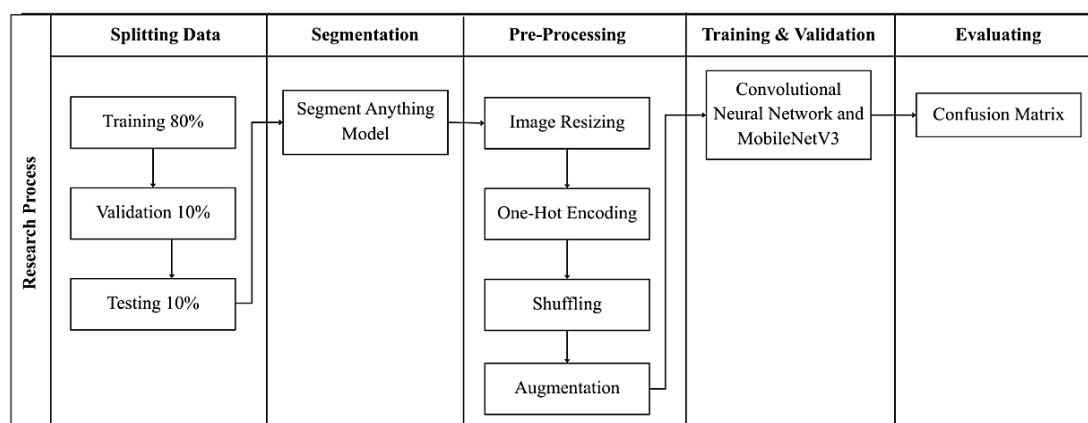


Figure 1. Research Process

2.1 Dataset

The dataset used is a public dataset "Skin Type Identification Dataset" https://figshare.com/authors/Sanjana_Solanga_Arachchi/21778199 obtained from the Figshare platform published by Arachchi and in .jpg format [22]. The dataset totals 2,250 images with 5 classes, namely normal, dry, oily, sensitive, and combination, which have been divided into 3 parts, namely training (80%), validation (10%), and testing (10%). Data splitting is shown in Table 1, and examples of skin types are shown in Figure 2.

Table 1. Data Splitting

Class	Training Data	Validation Data	Testing Data
Normal	360	45	45
Dry	360	45	45
Oily	360	45	45
Sensitive	360	45	45
Combination	360	45	45
Total	1.800	225	225

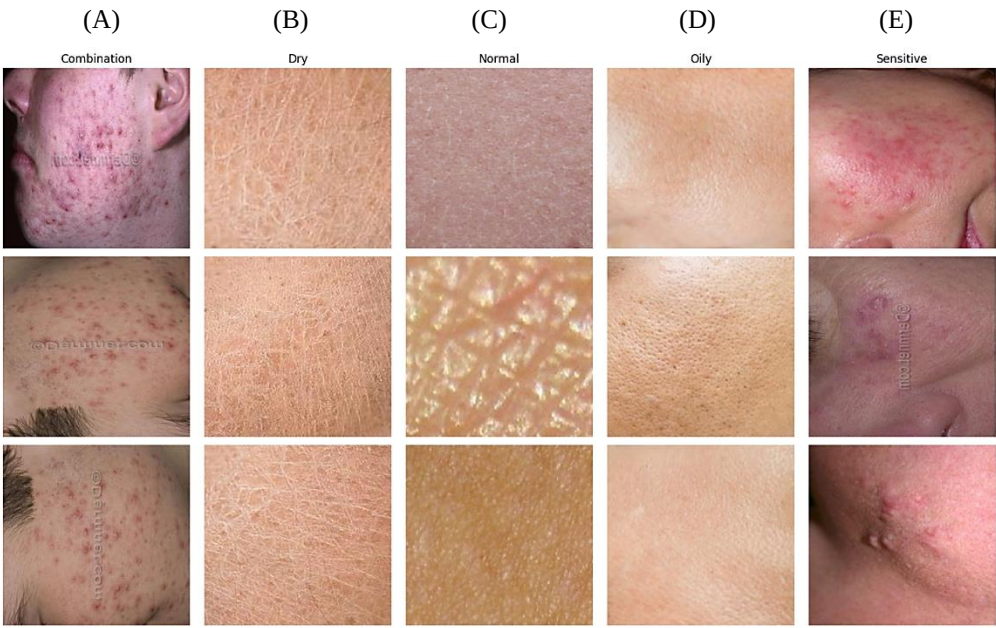


Figure 2. Examples of Skin Types (A) Combination (B) Dry (C) Normal (D) Oily (E) Sensitive

2.2 Pre-Processing Data

The dataset is uploaded to the drive and connected to the Google Colaboratory. The pre-processing stage aims to convert the raw data into a suitable format so that it can be processed by the model [23]. Pre-processing starts with image resizing so that all images are the same size (img_size = (224, 224, 3)), then one hot-encoding to change class labels (normal, dry, oily, sensitive, and combination) to five-dimensional binary vectors to avoid sequence bias (label_mode='categorical'), then shuffling to eliminate the influence of data sequence and maintain balance between classes during training (train_ds = train_ds.shuffle(256).prefetch(autotune)).

To increase the amount and diversity of data without collecting new images, a layer-based augmentation process is applied to generalize the model and prevent overfitting. Rescaling(1./255), layers. RandomFlip("horizontal"), layers. RandomRotation(0.08), layers. RandomZoom(0.08), layers. RandomContrast(0.08), layers. RandomTranslation(0.1, 0.1). Augmentation helps the model adapt to changes in position, visibility, and lighting conditions without altering the basic shape of the face in the image. Examples of skin types after SAM are shown in Figure 3.

2.3 Training, Validation, and Testing

This stage includes training, validation, and testing of the model in the research. The goal is to obtain the best model for classifying facial skin types from imagery. All three models use the same dataset and go through the same pre-processing stage. The main parameters applied to training are img_size = (224, 224, 3), batch_size = 32, epochs = 50, optimizer = 'adam', class_weight to address class imbalance, and early_stopping with patience = 5 to prevent overfitting during training.



Figure 3. Examples of Skin Types After SAM (A) Combination (B) Dry (C) Normal (D) Oily (E) Sensitive

2.3.1 Model 1 - CNN Baseline

The first model processes the pre-processing dataset using the baseline CNN to extract features and perform classification. The first model serves as a basis for determining initial performance without additional architecture or specialized segmentation techniques. In this method, the model is processed by multiple layers. The CNN model takes $224 \times 224 \times 3$ input images and consists of multiple conv2d layers, followed by batch normalization, ReLU activation, and max pooling to extract image features. The number of filters on each convolutional block increases gradually i.e. 32, 64, 128, and 256. The resulting features are then summarized using global average pooling and processed by the Dense layer before being classified into 5 classes using the Softmax layer. The CNN Architecture used for facial skin type classification is shown in Table 2.

Table 2. CNN Architecture Used for Facial Skin Type Classification

Stage	Layer	Output Shape	Parameters
1	Conv2D (32, 3×3) + BN + ReLU	$224 \times 224 \times 32$	896
2	MaxPooling (2×2)	$112 \times 112 \times 32$	0
3	Conv2D (64, 3×3) + BN + ReLU	$112 \times 112 \times 64$	18,496
4	MaxPooling (2×2)	$56 \times 56 \times 64$	0
5	Conv2D (128, 3×3) + BN + ReLU	$56 \times 56 \times 128$	73,856
6	MaxPooling (2×2)	$28 \times 28 \times 128$	0
7	Conv2D (256, 3×3) + BN + ReLU	$28 \times 28 \times 256$	295,168
8	MaxPooling (2×2)	$14 \times 14 \times 256$	0
9	Global Average Pooling	256	0
10	Dense (256) + BN + Dropout	256	65,792
11	Dense (Softmax – 5 classes)	5	1,285

2.3.2 Model 2 - MobileNetV3-based CNN

The second model uses a MobileNetV3-Small-based CNN as a feature extractor with image input ($224 \times 224 \times 3$). The features generated from MobileNetV3 have a size of $7 \times 7 \times 576$, then are summarized using global average pooling to reduce the dimension of the feature. Furthermore, batch normalization and dense layer are applied for the classification process, as well as dropouts to reduce overfitting. In the final layer, a dense layer with a softmax activation function is used to classify the image into 5 classes. The use of MobileNetV3 aims to produce a lighter model with optimal feature extraction capabilities. The MobileNetV3-based CNN architecture used for facial skin type classification can be seen in Table 3.

Table 3. MobileNetV3-Based CNN Architecture Used for Facial Skin Type Classification

Stage	Layer	Output Shape	Parameters
1	Input Layer	$224 \times 224 \times 3$	0
2	MobileNetV3Small (Feature Extractor)	$7 \times 7 \times 576$	939,120

Stage	Layer	Output Shape	Parameters
3	Global Average Pooling	576	0
4	Batch Normalization	576	2,304
5	Dense	256	147,712
6	Dropout	256	0
7	Dense (Softmax – 5 classes)	5	1,285

2.3.3 Model 3 – SAM and MobileNetV3-based CNN

The third model uses SAM as the image segmentation stage before classification, focusing on relevant image areas. SAM generates a segmentation mask of the main object in the image using a single foreground-point prompt. The segmented mask is then combined with the original image as an additional channel to produce an image with four channels: RGB and the mask. Before MobileNetV3 processes the image, a 1x1 convolution layer is applied to reduce the channel dimension. The segmented image was used as input for training and evaluation with a MobileNetV3-based CNN, which uses the same pipeline as model 2.

2.4 Evaluation

To evaluate and compare the performance of each model, a confusion matrix was used, and the results were analyzed using accuracy, precision, recall, and F1-score.

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

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

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

$$\text{F1 – Score} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

True Positive (TP) is the number of data points that are actually positive and correctly classified as positive by the system. True Negative (TN) is the number of data points that are actually negative and correctly classified as negative by the system. False Positive (FP) is the number of data points that are actually negative but are misclassified as positive by the system. False Negative (FN) is the amount of data that is actually positive but is incorrectly classified as negative by the system [24].

3. RESULTS AND DISCUSSION

This section describes the results from training and testing to determine the model with the best performance. To strengthen the interpretation, the results of the research are given in the form of comparison tables, confusion matrix visualizations, ROC curves, and performance graphs of each model.

Table 4 compares performance results based on the confusion matrix: accuracy, precision, recall, and F1-score. The test results for model 1 had the lowest performance, with an accuracy of 75% and an F1-score of 74%, indicating a higher error rate than other models. Then, model 2 performed best, with all metrics reaching 97%, indicating strong classification performance and consistency. Meanwhile, the model 3 performs quite well, with an accuracy of 79% and precision, recall, and F1-score values of 84% and 85%, respectively, which are almost the same as those of model 1.

Based on the results, Model 2 has the best performance. This shows that the MobileNetV3-based CNN is effective in extracting complex features in skin images. MobileNetV3 has the advantage of a residual inverted architecture and linear bottlenecks, which allow models to capture very fine skin texture details while remaining computationally efficient. Compared to model 1, the CNN baseline is still lacking in pattern generalization, especially in skin imagery. Meanwhile, model 3 SAM segmentation actually resulted in a 79% decrease in performance on MobileNetV3-based CNN due to the loss of important features in skin image, such as color degradation or the boundary between healthy skin and infected areas. The results are in line with previous research that states that segmentation does not always provide benefits in the classification of images that have less clear object boundaries and subtle texture differences, as they have the potential to eliminate important contextual information [26].

Table 4. Performance Results of 3 Models

	Accuracy	Precision	Recall	F1 Score
Model 1	75%	78%	75%	74%
Model 2	97%	97%	97%	97%
Model 3	79%	85%	84%	84%

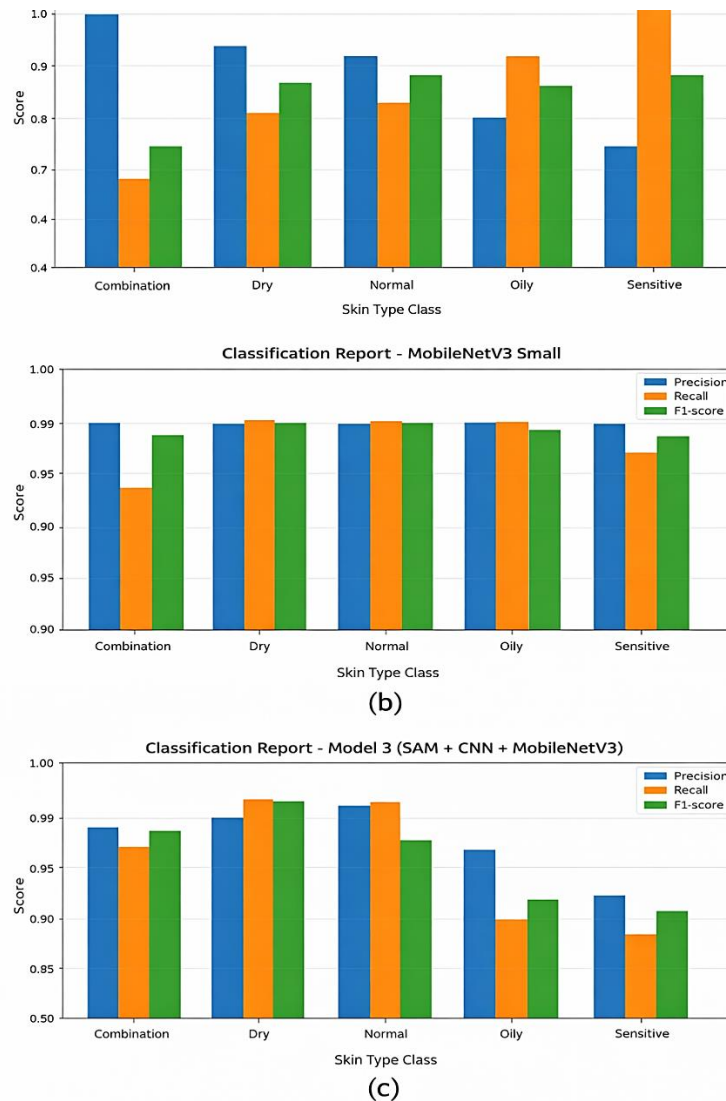


Figure 4. Graph of Performance Comparison (a) Model 1 (b) Model 2 (c) Model 3

The results of the performance evaluation of each model can be seen from the graph in Figure 4. Model 1 achieved an accuracy of 75%, with the highest accuracy of 100% in the sensitive class and the lowest recall of 31% in the combination class, while the sensitive class had a low precision of 55%, indicating that the model had visual confusion in identifying other skin types as sensitive classes and a high recall of 98%, indicating an imbalance in classification between classes. In the dry and oily classes, it has an F1-score of 91% and 86%. Model 1 has difficulty distinguishing subtle features between classes, such as the difference between oily textures and combination textures that have visual similarities in certain parts. As a result, the model is biased towards classes that have more dominant features and fails to generalize to classes with overlapping visual characteristics.

The Model 2 showed a significant improvement with an accuracy of 97% with a more balanced average precision, recall, and F1-score. This is due to the use of depthwise separable convolution which can extract skin texture features more efficiently without losing spatial features. The high F1-score in all classes also proves that the MobileNetV3-based CNN is better able to distinguish similar visual variations between classes. Meanwhile, the Model 3 only achieved an accuracy of 79% with the highest average value in the dry class. The normal and oily classes show an imbalance of precision and recall that indicates a tendency for prediction errors due to similar visuals. Based on the results in model 3, SAM can improve the ability to recognize certain classes, but it has not been able to fully address the ambiguity in classes that have similar visuals.

A confusion matrix is used to evaluate the proposed method by visualizing the test results [25]. In Figure 5 Model 1 still has misclassification in some classes, especially in classes with similar visual characteristics as in the combination class that has the lowest results, because 30 data is incorrectly predicted as sensitive classes. This shows that models still have difficulty distinguishing between combination and sensitive skin, as both classes have similar visual shapes such as uneven redness or shine. Model 2 results in

better performance by showing a minimal number of misclassification errors in almost all classes, which means that the model is able to map the important features of each class and can distinguish visual similarities. Meanwhile, model 3 experienced an increase in errors in certain classes compared to model 2.

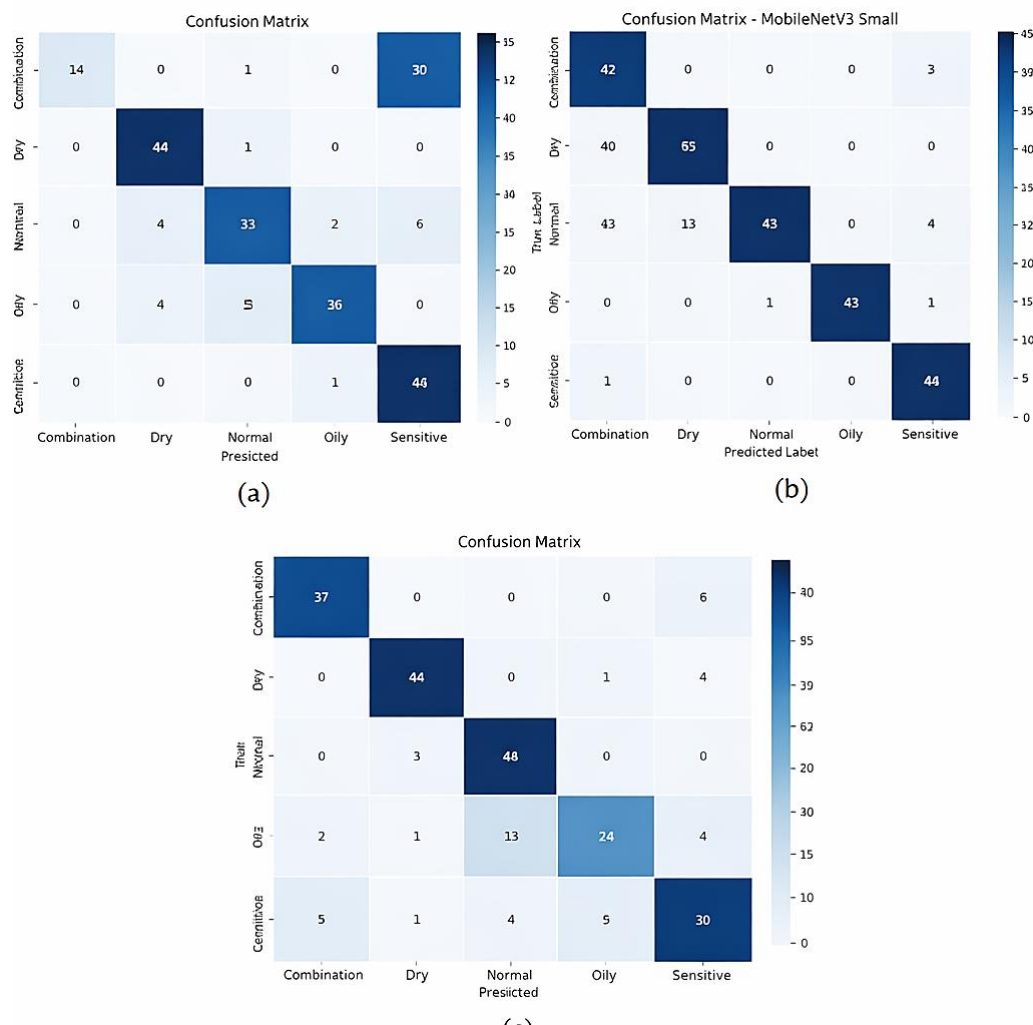


Figure 5. Confusion Matrix of 3 Models (a) Model 1 (b) Model 2 (c) Model 3

The most errors in the oily class were only 24 data correct, while 15 data were incorrectly classified as normal classes. Sensitive classes also decreased with 30 data correct and data misclassification scattered as combination, normal, and oily classes. The addition of SAM resulted in model confusion in the oily and normal classes. It shows that the oil or shine features in the combination class may have been removed or altered in texture during the segmentation process, resulting in the loss of important features as visual clues. ROC curves of the 3 models can be seen in Figure 6.

4. CONCLUSION

In this study, a combination of CNN, MobileNetV3, and SAM models was applied to classify facial skin type as an initial step toward determining appropriate skin care. The architecture was tested using 3 models to determine the best-performing model. Based on the test results, the CNN baseline model has the lowest performance. Although CNN can extract features from facial imagery, the CNN baseline still has limitations in capturing more complex features in facial skin imagery. Model 2, a MobileNetV3-based CNN, achieved the highest accuracy of 97% compared to other models. Meanwhile, adding SAM at the segmentation stage did not improve performance, indicating that automatic segmentation did not always yield better results, especially for facial skin types with distinct characteristics across classes. Not so obvious in the shape of the object, but more in the small texture on the surface of the skin.

Based on the results of the research and the limitations found, the following are suggestions for the development of further research such as the use of a larger and more diverse number of datasets, the use of optimization techniques or other architectures as well as segmentation techniques that are more suitable for

facial skin type imagery, and the results of this research can also be developed in the form of a skin care recommendation application system to make it easier to use in various times and places.

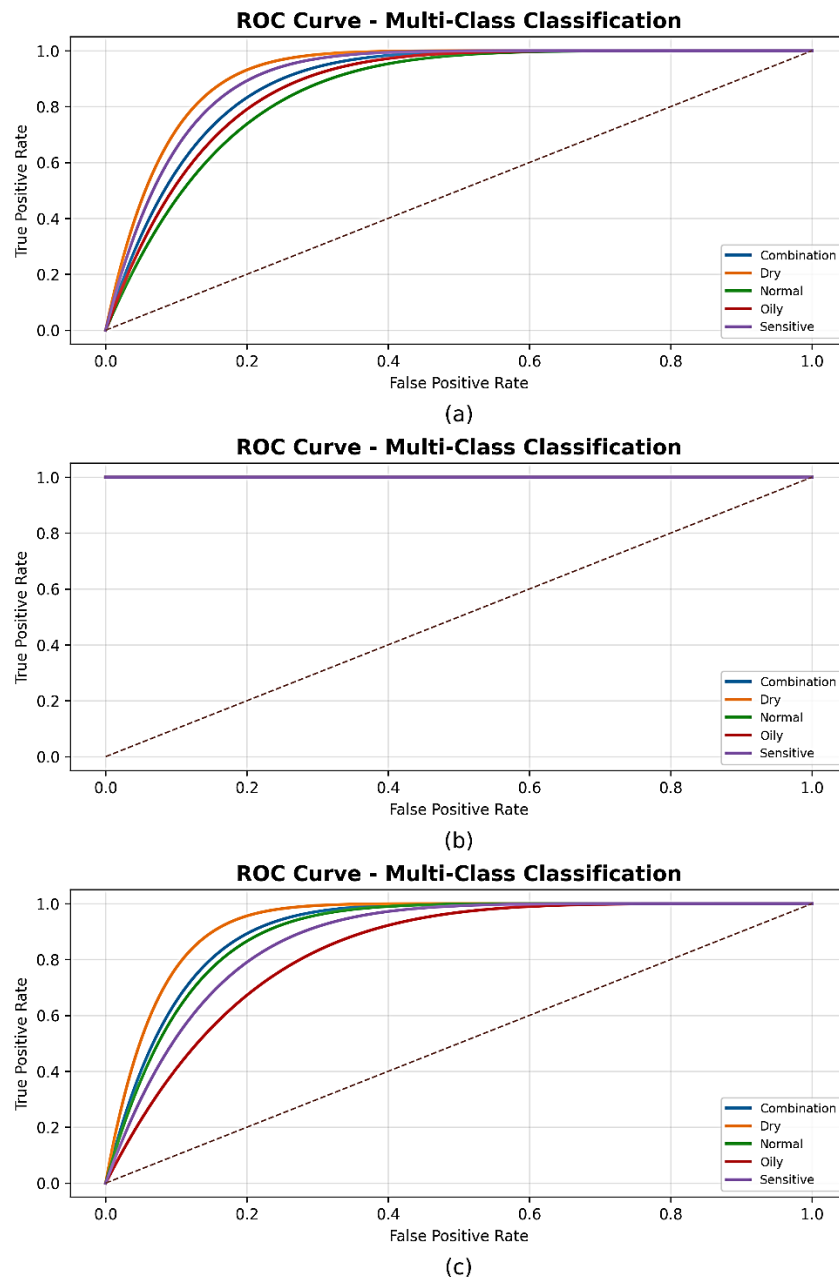


Figure 6. ROC Curve of 3 Model (a) Model 1 (b) Model 2 (c) Model 3

REFERENCES

- [1] R. Oliveira, J. Ferreira, L. F. Azevedo, dan I. F. Almeida, "An Overview of Methods to Characterize Skin Type: Focus on Visual Rating Scales and Self Report Instruments," *MDPI*, vol. 10, no. 1, hal. 1–15, 2023, doi: 10.3390/cosmetics10010014.
- [2] J. Ananta, "Gak Hanya Jerawat, Kulit Kusam Jadi Masalah Serius Perempuan Indonesia," *IDN TIMES*. Tersedia pada: <https://www.idntimes.com/life/women/kulit-kusam-jadi-masalah-perempuan-indonesia-00-p9v6l-2mqlbw>
- [3] riska sahla Nafiah, E. Fitraneti, Y. Rizal, I. Primawati, dan ayu dian Hamama, "Pengaruh Paparan Sinar Ultraviolet terhadap Kesehatan Kulit dan Upaya Pencegahannya : Tinjauan Literatur," *Sci. J.*, vol. 3, no. 3, hal. 185–194, 2024, doi: 10.56260/sciena.v3i3.147.
- [4] anisa dian Agustina, "Klasifikasi Citra Jenis Kulit Wajah Dengan Algoritma Convolutional Neural Network (Cnn) Resnet-50," *JISSI*, vol. 1, no. 3, hal. 2–8, 2024.
- [5] sarah nabila Hafni, Yuliana, dan A. Yulastri, "Literature Riview : Hubungan Tingkat Pengetahuan

- Dengan Gaya Hidup Terhadap Kesehatan Kulit Wajah,” *J. Penelit. imliah multidisiplin*, vol. 8, no. 7, hal. 285–293, 2024.
- [6] L. Khodijah, R. Latifah, dan Y. Adharani, “Identifikasi Jenis Kulit Wajah Menggunakan K-Nearest Neighbor,” *Sintaks*, 2022.
 - [7] P. S. Parraga *et al.*, “Skin type classifications : does the perfect assessment exist ?,” *J. Dermatolog. Treat.*, vol. 35, no. 1, hal. 1471–1753, 2024, doi: 10.1080/09546634.2024.2427325.
 - [8] shinta dewi Kusumaningrum dan I. Muhimmah, “Analisis Faktor dan Metode Untuk Menentukan Tipe Kulit Wajah : Tinjauan Literatur,” *JTIHK*, vol. 10, no. 4, hal. 753–762, 2023, doi: 10.25126/jtiik.2023106955.
 - [9] T. Anoushkaa, K. V Dikshitha, P. lakshmi Pranitha, dan K. Vaidehi, “Deep Learning in Dermatology: A Review on Skin Type Classification and Disease Detection,” *Interantional J. Sci. Res. Eng. Manag.*, vol. 09, no. 03, hal. 1–9, 2025, doi: 10.55041/ijrsrem42607.
 - [10] S. Cahyaningsih, A. Triayudi, dan I. D. Sholihati, “Kombinasi Metode Certainty Factor dan Forward Chaining untuk Identifikasi Jenis Kulit Wajah Berbasis Android,” *J. Media Inform. Budidarma*, vol. 5, no. 1, hal. 74, 2021, doi: 10.30865/mib.v5i1.2591.
 - [11] R. A. Dianto, T. F. Anggraeny, dan H. Maulana, “Analisis Efektifitas Algoritma Mobilenetv3-large dan efficientnet-b0 untuk klasifikasi citra penyakit daun jeruk,” *JITET*, vol. 13, no. 3, 2025.
 - [12] D. P. Sinaga dan N. Khairina, “Penerapan Mobilenetv3 Untuk Klasifikasi Jenis Bahan Pakaian,” *INCODING J. Inform. Comput. Sci. Eng.*, vol. 4, no. 1, hal. 1–8, 2024, doi: 10.34007/incoding.v4i1.469.
 - [13] S. Dewi, F. Ramadhani, dan S. Djasmayena, “View of Klasifikasi Jenis Jerawat Berdasarkan Gambar Menggunakan Algoritma CNN (Convolutional Neural Network),” 2024.
 - [14] Herimanto, A. S. Dharma, J. Amalia, D. Largo, dan C. A. P. Sihite, “Kinerjance Analysis of MobileNetV3-based Convolutional Neural,” *RESTI*, vol. 5, no. 158, hal. 701–709, 2024, doi: <https://doi.org/10.29207/resti.v8i6.5982>.
 - [15] J. Zhang, F. Zhoun, K. He, M. Ji, dan C. Li, “Recent Advancements and Perspectives in the Diagnosis of Skin Diseases Using Machine Learning and Deep Learning : A Review,” 2023.
 - [16] A. Baroroh dan H. Harintaka, “Ekstraksi Otomatis Tapak Bangunan (Building Footprint) Pada Ortofoto Menggunakan Segment Anything Model (Sam),” *Elipsoida J. Geod. dan Geomatika*, vol. 8, no. 1, hal. 52–59, 2025, doi: 10.14710/elipsoida.2025.25145.
 - [17] L. Lutviana, Rian Ardianto, dan Purwono, “CNN-based Classification of Bladder Tissue Lesions from Endoscopy Images,” *IT J. Res. Dev.*, vol. 9, no. 2, hal. 95–107, 2025, doi: 10.25299/itjrd.2025.17867.
 - [18] M. Fadhillah, D. Suryani, A. Labellapansa, dan H. Gunawan, “Corn Leaf Diseases Recognition Based on Convolutional Neural Network,” *IT J. Res. Dev.*, vol. 8, no. 1, hal. 14–21, 2023, doi: 10.25299/itjrd.2023.13904.
 - [19] prima A. Gani dan E. Poerwandono, “Pengembangan Aplikasi Android Berbasis Machine Learning Untuk Pendeteksi Jenis Kulit Wajah Menggunakan Deep Learning,” *JATI*, vol. 9, no. 6, hal. 9551–9557, 2025, doi: <https://doi.org/10.36040/jati.v9i6.15674>.
 - [20] E. Moupojou, F. Retraint, H. Tapamo, M. Nkenlifack, C. Kacpah, dan A. Tagne, “Segment Anything Model and Fully Convolutional Data Description for Plant Multi-Disease Detection on Field Images,” *IEEE Access*, vol. 12, no. July, hal. 102592–102605, 2024, doi: 10.1109/ACCESS.2024.3433495.
 - [21] Y. Huang, X. Yang, L. Hiu, H. Zhou., A. Chang, dan X. Zhou, “Medical Image Analysis,” *elsevier*, 2024.
 - [22] S. S. Arachchi, “skin type identification,” FIGSHARE. [Daring]. Tersedia pada: https://figshare.com/authors/Sanjana_Solanga_Arachchi/21778199
 - [23] J. Deng, “The Data Preprocessing Technique in Machine Learning,” *Sci. Technol. Eng. Chem. Environ. Prot.*, vol. 1, no. 10, hal. 1168–1172, 2024, doi: 10.61173/0hfqqz07.
 - [24] R. Yohannes dan ezar M. Al Rivan, “Klasifikasi Jenis Kanker Kulit Menggunakan CNN-SVM,” *J. Algoritma.*, vol. Vol.2, no. 2, hal. 1, 2022.
 - [25] A. Luque, M. Mazzoleni, A. Carrasco, dan A. Febramosca, “Visualizing Classification Results : Confusion Star and Confusion Gear,” *IEEE Access*, vol. 10, hal. 1659–1677, 2022, doi: 10.1109/ACCESS.2021.3137630.
 - [26] M. Talaat, X. Si, dan J. Xi, “applied sciences Multi-Level Training and Testing of CNN Models in Diagnosing Multi-Center COVID-19 and Pneumonia,” *MDPI*, 2023.