

Monk Skin Tone Classification: RMSprop vs Adam Optimizer in MobileNetV2

Firman Naufal Aryaputra ¹, Christy Atika Sari ², Eko Hari Rachmawanto ³

^{1,2,3} Informatics Engineering Study Program, Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang
Jl. Imam Bonjol 207, Pendrikan Kidul, Semarang, 50131, Indonesia

Info Artikel

Riwayat Artikel:

Received 2025-05-30

Revised 2025-06-26

Accepted 2025-06-28

Abstract – The lack of accurate and accessible skin tone classification systems poses significant challenges in personalized fashion recommendations and inclusive technology development. This study aims to develop a skin tone classification system utilizing the Monk Skin Tone (MST) scale through the implementation of Convolutional Neural Network with MobileNetV2 architecture enhanced by transfer learning techniques. The MST scale encompasses ten distinct categories providing comprehensive representation of human skin color diversity. The methodology leverages efficient MobileNetV2 architecture suitable for web deployment, transfer learning to enhance accuracy despite limited training data, and strategic dataset balancing. A dataset of 1,729 facial photographs representing the complete MST spectrum was utilized. Preprocessing involved scaling images to 224×224 pixels, normalization, and augmentation through various transformations to address class imbalance challenges. The dataset was partitioned using a 70:15:15 ratio for training, validation, and testing respectively. The system was implemented as a web platform called SkinToneAI that enables users to upload facial images for skin tone analysis and receive personalized clothing color recommendations. Evaluation demonstrated classification accuracy of 97.83% on the test dataset with a loss value of 0.1166 when using Adam optimizer, while RMSprop optimizer achieved better performance with 98.26% accuracy and 0.0548 loss value. The implemented web application successfully translates technical capabilities into practical fashion assistance. The system provides users with customized apparel color suggestions based on their identified skin tone category, effectively connecting advanced AI technology with everyday fashion needs.

Keywords: Classification; Convolutional Neural Network; Deep Learning; MobileNetV2; Monk Skin Tone.

Corresponding Author:

Christy Atika Sari

Email:

christy.atika.sari@dsn.dinus.ac.id



This is an open access article under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license.

Abstrak – Kurangnya sistem klasifikasi warna kulit yang akurat dan mudah diakses menimbulkan tantangan signifikan dalam rekomendasi fashion yang dipersonalisasi dan pengembangan teknologi yang inklusif. Penelitian ini bertujuan mengembangkan sistem klasifikasi warna kulit menggunakan skala Monk Skin Tone (MST) melalui implementasi Convolutional Neural Network dengan arsitektur MobileNetV2 yang ditingkatkan dengan teknik transfer learning. Skala MST mencakup sepuluh kategori yang berbeda untuk memberikan representasi komprehensif terhadap keragaman warna kulit manusia. Metodologi memanfaatkan arsitektur MobileNetV2 yang efisien untuk deployment web, transfer learning untuk meningkatkan akurasi meskipun data training terbatas, dan penyeimbangan dataset strategis. Dataset yang digunakan terdiri dari 1.729 foto wajah yang mewakili spektrum MST lengkap. Preprocessing melibatkan penskalaan gambar ke 224×224 piksel, normalisasi, dan augmentasi melalui berbagai transformasi untuk mengatasi tantangan ketidakseimbangan kelas. Dataset dibagi menggunakan rasio 70:15:15 untuk training, validasi, dan testing. Sistem diimplementasikan sebagai platform web bernama SkinToneAI yang memungkinkan pengguna mengunggah gambar wajah untuk analisis warna kulit dan menerima rekomendasi warna pakaian yang dipersonalisasi. Evaluasi menunjukkan akurasi klasifikasi 97,83% pada dataset test dengan nilai loss 0,1166 saat menggunakan optimizer Adam, sedangkan optimizer RMSprop mencapai performa better dengan akurasi 98,26% dan nilai loss 0,0548. Aplikasi web yang diimplementasikan berhasil menerjemahkan kemampuan teknis menjadi bantuan fashion praktis. Sistem memberikan saran warna pakaian yang disesuaikan berdasarkan kategori warna kulit yang teridentifikasi, secara efektif menghubungkan teknologi AI canggih dengan kebutuhan fashion sehari-hari.

Kata Kunci: Convolutional Neural Network, Deep Learning, Klasifikasi, MobileNetV2, Monk Skin Tone.

I. INTRODUCTION

Recent innovations in artificial intelligence and deep learning have catalyzed significant transformations in image analysis methodologies, with particular implications for the fashion industry. A fundamental challenge within fashion contexts involves the formulation of complementary clothing color recommendations that correspond to individual skin tone characteristics [1]. The selection of optimal clothing colors not only enhances aesthetic presentation but also potentially elevates self-confidence. Nevertheless, this selection process remains

inherently subjective, influenced by the heterogeneous spectrum of human skin pigmentation, environmental lighting conditions, and diverse individual preferences, which collectively contribute to inconsistent recommendation outcomes [2].

Concurrent with technological advancements, deep learning-based digital image processing techniques, particularly Convolutional Neural Networks (CNN), have demonstrated remarkable efficacy in visual feature extraction and complex pattern classification. CNN implementation has been extensively documented across various recognition applications, including skin tone classification systems. This technology exhibits high precision in feature recognition, rendering it particularly suitable for skin tone detection and classification across fashion, cosmetic, and medical domains [1], [3]. CNN-based systems facilitate the identification of skin color variations with better accuracy compared to conventional methodologies reliant on subjective human assessment.

Empirical research indicates that CNN implementation, particularly utilizing efficient architectures such as MobileNetV2, exhibits substantial compatibility with web-based applications due to its minimal computational requirements and inference efficiency [4], [5]. MobileNetV2 has been specifically engineered for mobile and web-based deployment, incorporating depthwise separable convolution techniques that substantially enhance inference speed and reduce parametric complexity without compromising classification accuracy [6], [7]. Furthermore, the application of transfer learning methodologies to MobileNetV2 has demonstrated significant performance enhancements, particularly in contexts characterized by limited dataset availability. Transfer learning leverages pre-trained weight parameters from models previously trained on extensive datasets such as ImageNet [8], thereby accelerating training convergence and improving model generalization capabilities for skin tone classification tasks.

In this investigation, the researchers developed a MobileNetV2-based skin tone classification system integrated within the SkinToneAI web application. This platform enables users to upload facial photographs and receive skin tone classification results according to the Monk Skin Tone (MST) scale. The system provides not only accurate skin color classification but also generates clothing color recommendations compatible with the user's identified skin tone, based on established color theory principles commonly applied in fashion industry contexts [2]. This process incorporates automated facial detection technology, skin region segmentation, and the mapping of classification outcomes to optimal clothing color palettes for each skin tone category.

The Monk Skin Tone (MST) scale implemented in this research was developed to provide a more inclusive and comprehensive representation of skin color classification. The MST scale encompasses ten distinct categories and represents human skin color variations more holistically compared to the more limited Fitzpatrick classification system [9]. MST was conceptualized with consideration for cultural and geographical factors that influence skin pigmentation variations, thereby enhancing its representative capacity for global applications. Consequently, the implementation of MST enables this application to accommodate a broader demographic spectrum representing diverse ethnic backgrounds [10].

Prior research by Mbatha et al. [11] has explored skin tone estimation under diverse lighting conditions using the MST scale with CNN regression model approaches, achieving accuracy ranges of 85.45–97.16%. However, that study employed a larger dataset (21,375 images) with primary focus on lighting condition variations, without in-depth exploration of optimizer algorithm optimization and specific transfer learning techniques for the MST scale. The position of this research is to examine the performance comparison of RMSprop versus Adam optimizers on MobileNetV2 architecture, which has not been comprehensively explored in the context of skin tone classification using the MST scale.

The skin tone detection technology employed in this research offers more objective and personalized recommendations. CNN, representing an evolution of Multi-layer Perceptron (MLP) architectures, is specifically designed for processing two dimensional data structures such as images. CNN is categorized as a Deep Neural Network with enhanced architectural complexity and is widely implemented for image data processing applications. CNN demonstrates high capability in pattern and feature detection, enabling the system to recognize diverse skin tone variations with optimal accuracy. Based on the identified challenges, this research formulates the primary problems of how to develop an objective and consistent skin tone classification system using the MST scale, how to optimize CNN MobileNetV2 model performance through comparison of RMSprop versus Adam optimizer algorithms, and how to address dataset imbalance challenges in skin tone classification to achieve optimal accuracy. This research aims to develop a deep learning-based facial skin tone classification model utilizing the CNN algorithm, specifically adopting the MobileNetV2 architecture and applying transfer learning techniques calibrated for the Monk Skin Tone (MST) scale. The system is designed to deliver more accurate and consistent skin tone classification outcomes. Additionally, the system incorporates functionality to provide recommendations for harmonious clothing colors based on individual skin tone characteristics, thereby assisting users in color selection processes with enhanced objectivity.

The primary contributions of this research to the Informatics domain encompass comprehensive implementation and evaluation of RMSprop versus Adam optimizer comparisons on MobileNetV2 architecture for skin tone classification, development of dataset balancing strategies combining undersampling and data

augmentation techniques to address class imbalance in MST datasets, realization of a web-based SkinToneAI system integrating MobileNetV2 models with real-time clothing color recommendations, and demonstration that despite utilizing a smaller dataset compared to Mbatha et al research, appropriate optimization in preprocessing and normalization can achieve better accuracy. Consequently, this research demonstrates potential to enhance user self-confidence and aesthetic presentation through a more personalized and analytically-driven approach, while contributing to the development of more inclusive and fair AI technology in recognition systems.

II. METHOD

This research implements a deep learning approach utilizing Convolutional Neural Network (CNN) algorithms with a modified pre-trained MobileNetV2 architecture based on the Monk Skin Tone (MST) scale. MST represents a contemporary paradigm in Skin Tone categorization designed to provide more accurate and inclusive representation in visual classification processes based on skin color characteristics [9]. Transfer learning techniques are applied to leverage knowledge from previously trained models, reducing the need for extensive datasets and accelerating the model training process [12], [13], [14]. To validate research results, an experimental methodology is systematically implemented according to the research process flow depicted in Figure 1.

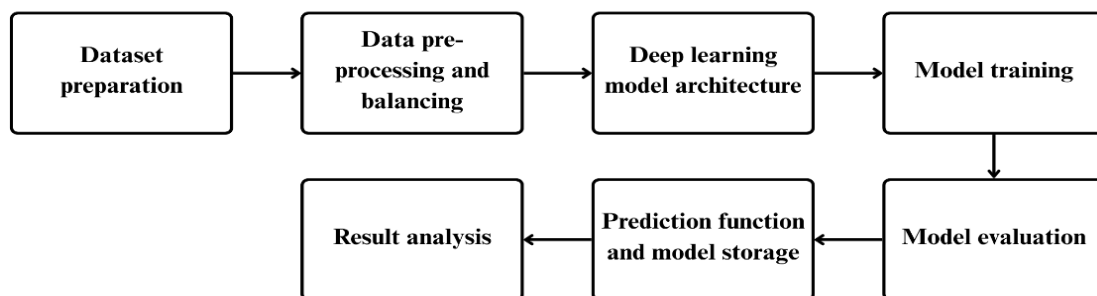


Figure 1. Research Flow Chart

As illustrated in Figure 1, the process begins with the "Dataset Preparation" stage for acquiring and organizing data. Subsequently, "Preprocessing and Data Balancing" is executed, encompassing normalization, augmentation, and balancing strategies to address class imbalance. The next stage involves "Deep Learning Model Architecture" using MobileNetV2, followed by "Model Training" using the balanced dataset. "Model Evaluation" is then conducted to measure performance with metrics such as accuracy, precision, and recall. To facilitate inference, the trained model is implemented in "Prediction Function and Model Storage." Finally, "Website Implementation" is completed to integrate the model into an interactive web interface.

A. Dataset Preparation

This research utilizes the Monk Skin Tone (MST) dataset consisting of 10 Skin Tone categories (MST1 through MST10). The dataset is annotated using the MST scale developed by Google researchers in collaboration with Dr. Ellis Monk. All model images were acquired within a limited time interval to minimize Skin Tone variations due to seasonal or temporal factors. Models were photographed in various poses, facial expressions (smiling, neutral, sad/frowning, and silly expressions), and lighting conditions. This research employs the open-access MST-E dataset encompassing 19 models across the spectrum of 10 MST categories [10]. The initial dataset comprised 1,729 images with uneven distribution among the 10 MST classes, as illustrated in Figure 2.

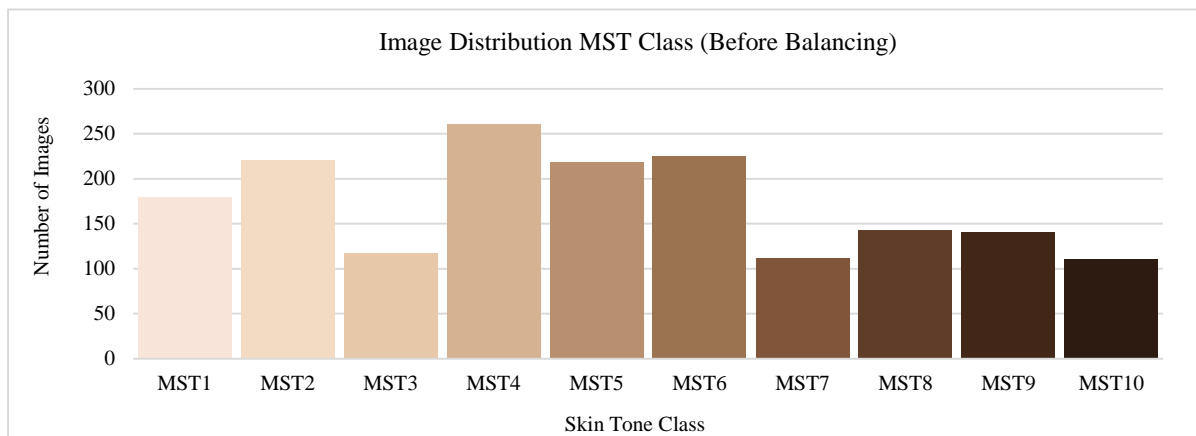


Figure 2. Dataset Distribution Before Balancing

Based on Figure 2, there is a disparity in the number of images between classes. The MST4 class has the highest representation (261 images), while the MST10 class has the lowest representation (111 images). This imbalance potentially induces bias in the model learning process, where the model tends to demonstrate better performance in classifying classes with greater representation. The implemented data augmentation strategy addresses class imbalance through systematic transformations including rotation (± 20 degrees), horizontal and vertical shifts ($\pm 20\%$), zoom operations ($\pm 20\%$), brightness adjustments (0.8-1.2 range), horizontal flipping, and shear transformations ($\pm 20\%$). These parameters were selected based on preliminary experiments to maintain facial structure integrity while introducing sufficient variation for model generalization.

Parameter tuning involved systematic evaluation of learning rates (0.001 for Adam, 0.0005 for RMSprop), batch sizes (32), and dropout rates (0.0001). The MobileNetV2 architecture utilized represents a modified implementation where the pre-trained ImageNet weights are frozen, with custom classification layers comprising GlobalAveragePooling2D, BatchNormalization, and three Dense layers (256, 128, and 10 units respectively) with ReLU activation and minimal dropout to prevent overfitting while maintaining computational efficiency.

B. Preprocessing and Data Balancing

In the subsequent stage, preprocessing is performed on the acquired data. Preprocessing in this research includes standardizing image sizes to 224×224 pixels, followed by dataset balancing using techniques proposed by [5]. The dataset balancing strategy encompasses:

- Undersampling classes with image counts exceeding 150, extracting random subsets of 150 images. This technique effectively reduces samples from dominant classes to prevent learning bias.
- Data augmentation for classes with fewer than 150 images, until reaching 150 images per class. The augmentation techniques adopt methodologies proposed by [15], including rotation (± 20 degrees), horizontal and vertical shifts ($\pm 20\%$), zoom ($\pm 20\%$), brightness modifications (0.8-1.2 range), horizontal flipping, and shear transformation ($\pm 20\%$). These parameters were selected based on preliminary experiments to maintain facial structure integrity while introducing sufficient variation for model generalization.

The implementation of data augmentation not only contributes to dataset balancing but also enhances the model's generalization capabilities by exposing it to variations of images representing real-world conditions. After the balancing process, the dataset is partitioned into three segments with proportions of 70% training set (105 images per class), 15% validation set (22 images per class), and 15% testing set (23 images per class). The total dataset after balancing amounts to 1,500 images (150 images per class) with balanced distribution as illustrated in Figure 3.

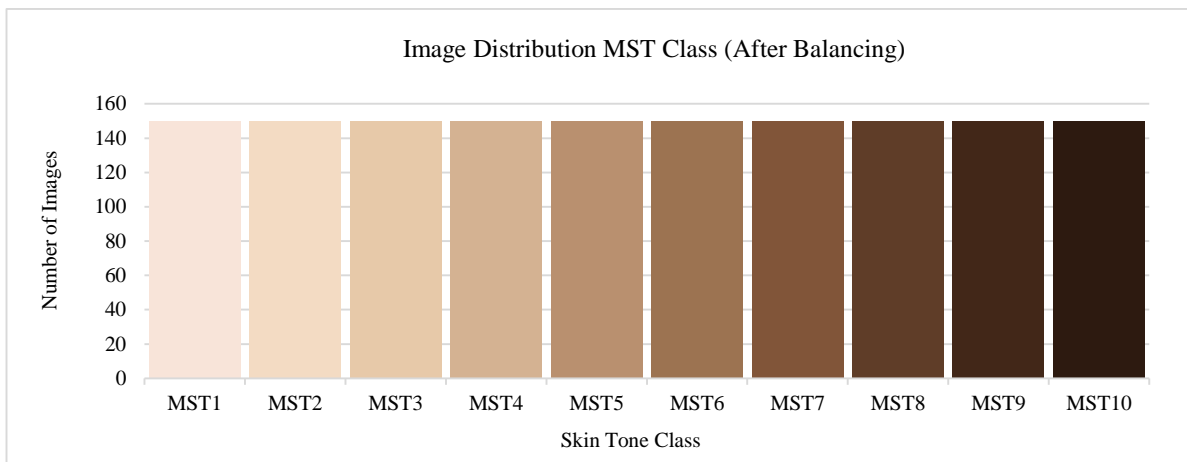


Figure 3. Image distribution per MST class after balancing

The implemented data augmentation strategy addresses class imbalance through systematic transformations including rotation (± 20 degrees), horizontal and vertical shifts ($\pm 20\%$), zoom operations ($\pm 20\%$), brightness adjustments (0.8-1.2 range), horizontal flipping, and shear transformations ($\pm 20\%$). These parameters were selected based on preliminary experiments to maintain facial structure integrity while introducing sufficient variation for model generalization.

C. Deep Learning Model Architecture

The model architecture implementation adopts a transfer learning approach with MobileNetV2 as the base model, pre-trained on the ImageNet dataset with an input shape of $224 \times 224 \times 3$. MobileNetV2 is selected as the optimal MobileNet architecture component for mobile and web applications, with an equilibrium between accuracy and computational efficiency [16]. According to research [4], the MobileNet architecture is specifically designed for applications with limited resources while maintaining better performance, making it an ideal option for implementation in web applications. This approach aligns with the methodology of [1], applying CNN models for Skin Tone classification. The MobileNetV2 architecture utilized represents a modified implementation where the pre-trained ImageNet weights are frozen, with custom classification layers comprising GlobalAveragePooling2D, BatchNormalization, and three Dense layers (256, 128, and 10 units respectively) with dropout regularization between layers, using ReLU activation and minimal dropout to prevent overfitting while maintaining computational efficiency. Table 1 illustrates the summary of the model implemented in this research

TABLE 1
 MODEL SUMMARY

Layer (type)	Output Shape	Param #
mobilenetv2_1.00_224 (Functional)	(None, 7, 7, 1280)	2257984
global_average_pooling2d (GlobalAveragePooling2D)	(None, 1280)	0
batch_normalization (BatchNormalization)	(None, 1280)	5120
dense (Dense)	(None, 256)	327936
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 128)	32896
dropout_1 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 10)	1290
Total params: 2,625,226		
Trainable params: 364,682		
Non-trainable params: 2,260,544		

Based on Table 1, the model implemented for Skin Tone classification employs a sequential architecture with MobileNetV2 as the base model. The architecture consists of multiple layers organized sequentially to optimize feature extraction capabilities and minimize computational complexity. The model utilizes

MobileNetV2 with input dimensions of 224×224, producing feature representations with dimensions (7, 7, 1280). The MobileNetV2 layer contributes significantly with 2,257,984 parameters. These feature representations are then processed through Global Average Pooling to reduce spatial dimensions, resulting in a 1280-dimensional feature vector without additional parameters [4], [17]. Subsequently, Batch Normalization is implemented to normalize activation distributions by adding 5,120 parameters. The model continues with three fully connected (Dense) layers with ReLU activation interconnected with Dropout layers to prevent overfitting. The first Dense layer transforms the input into 256 units (327,936 parameters), followed by the second Dense layer with 128 units (32,896 parameters), and the output layer with 10 units (1,290 parameters) representing the target classes. In aggregate, the model has 2,625,226 parameters, with only 364,682 parameters being trainable, while 2,260,544 parameters are non-trainable from the frozen MobileNetV2 layer. This transfer learning approach enables the model to leverage knowledge from pre-training on ImageNet dataset, while optimizing only the custom classification layers for the specific domain of Skin Tone classification.

D. Model Training

The model training process is executed with an iterative approach using multiple techniques to optimize performance and prevent overfitting. The use of callback functions such as ReduceLROnPlateau and ModelCheckpoint can optimize the training process of CNN models for Skin Tone classification. The implemented training parameters include Batch Size (32), Maximum Epochs (25), Steps per Epoch (32), and Validation Steps (6). The applied techniques include:

- **Learning Rate Reduction:** utilizing the ReduceLROnPlateau callback to reduce the learning rate when validation loss does not improve for 5 consecutive epochs, with a reduction factor of 0.5. This value is determined based on preliminary experiments showing optimal convergence.
- **ModelCheckpoint:** Saving models with better validation performance. This checkpoint practice is an implementation of fundamental principles in efficient deep learning model development, maximizing computational resource effectiveness and ensuring maximum accuracy in Skin Tone classification results [18].

Parameter tuning involved systematic evaluation of learning rates (0.001 for Adam, 0.0005 for RMSprop), batch sizes (32), and dropout rates (0.0001). The learning rate values were determined through preliminary experiments to achieve optimal convergence while maintaining training stability across both optimizer configurations.

E. Model Evaluation

Model evaluation is conducted on an independent testing set from the training process to obtain an objective performance estimate, particularly for multi-class classification according to the research methodology [19]. The implemented evaluation metrics include accuracy (percentage of accurate predictions from total predictions), categorical cross-entropy loss function value, precision, recall, and F1-score for each class, as well as a confusion matrix for model performance evaluation at the class level. Additionally, qualitative evaluation is performed with prediction visualization on sample images from the testing set. The entire evaluation process uses the scikit-learn library to ensure consistency and reliability of metric calculations. The evaluation stage is finalized with visualization of model predictions against sample images from the test dataset, with color differentiation for accurate predictions (green) and inaccurate predictions (red), as well as displaying confidence values, providing qualitative insights about the model's generalization capabilities in classifying various tones on the Monk Skin Tone Scale. Figure 4 illustrates model predictions on test data.

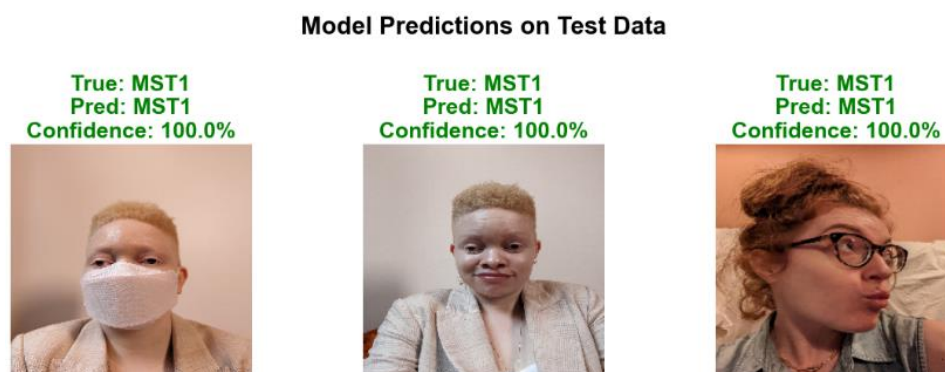


Figure 4. Model prediction on test data

F. Prediction Function and Model Storage

The developed Skin Tone prediction system implements MobileNetV2-based deep learning techniques for Skin Tone classification based on the Monk Skin Tone (MST) scale. The Skin Tone prediction function applies the model inference paradigm to input images with a standard preprocessing approach, including pixel value normalization in the range (0,1) and image dimension transformation for compatibility with the model architecture [17]. The prediction stage uses conventional softmax-based classification techniques in multi-class classification problems, consistent with research [20] in Skin Tone classification. Visualization of results is presented through graphical representation based on prediction categories on the Monk Skin Tone (MST) scale, supporting the interpretability of Skin Tone recognition system evaluation results. The significance of visual representation and color reproduction accuracy on the MST scale for validation and interpretation of results has been emphasized in previous research [21].

The implementation of model storage using the HDF5 (h5) format is a standard practice in the TensorFlow ecosystem, where models can be efficiently saved and reloaded using the official TensorFlow API [22]. Meanwhile, class metadata is stored in a structured JSON format, which has been extensively adopted to enhance interoperability and accessibility in production systems [23]. This approach aligns with methodologies emphasizing the importance of standardization and modularization in developing Skin Tone analysis systems for beauty and fashion applications. The separation of model storage and metadata formats not only enhances interoperability but also facilitates model reuse in various downstream applications such as Skin-Tone-based clothing color recommendation systems.

G. Website Implementation

The trained model is integrated into a web application, facilitating users to obtain Skin Tone classification and clothing color recommendations in real-time. The web application is developed with a responsive client-server architecture, using Python web frameworks such as Flask for the backend, and HTML, CSS, and JavaScript for the frontend, allowing the application to be optimally accessed through various devices [24], [25]. The clothing color recommendation system adopts an ontology approach to map relationships between the user's Skin Tone and clothing color palettes [26]. The system is implemented with the following feature architecture:

- Frontend: the user interface is constructed using HTML, CSS, and JavaScript. The frontend provides an area to upload facial images, display Skin Tone classification results, and visualize clothing color recommendations.
- Backend: the backend is constructed using Flask, a Python web framework. The backend manages facial detection processes, Skin Tone classification using the trained model, and generation of clothing color recommendations.
- Facial Detection: the system implements the Haar Cascade algorithm from OpenCV to detect facial regions in images. This allows the system to focus on facial areas and minimize background influence in Skin Tone classification [27].
- Skin Region Extraction: After facial detection, the system performs Skin Tone segmentation into 5 categories (Light, Light Medium, Medium, Medium Deep, Deep) and produces optimal clothing color recommendations along with color chip visualizations.

III. Result and Discussion

A. Model Training Results

The training procedure was executed across 25 epochs, incorporating a callback mechanism to retain the optimal model configuration based on validation accuracy performance. The training protocol utilized a batch size of 32, encompassing 32 iterations per epoch alongside 6 validation iterations. The Adam optimization algorithm was deployed with an initial learning rate parameter of 0.001, enabling dynamic gradient-based adjustments throughout the training process.

Multiple regularization strategies were incorporated into the model architecture to prevent overfitting. Batch normalization was implemented following the GlobalAveragePooling2D layer to normalize activations and expedite convergence. A dropout rate of 0.0001 was introduced after each Dense layer to enhance the model's generalization capabilities. Furthermore, a learning rate scheduler utilizing ReduceLROnPlateau (with a reduction factor of 0.5 and patience threshold of 5 epochs) was employed to enable progressive fine-tuning toward optimal convergence. Figure 5 presents the model's performance throughout the training cycle, depicting accuracy and loss metrics

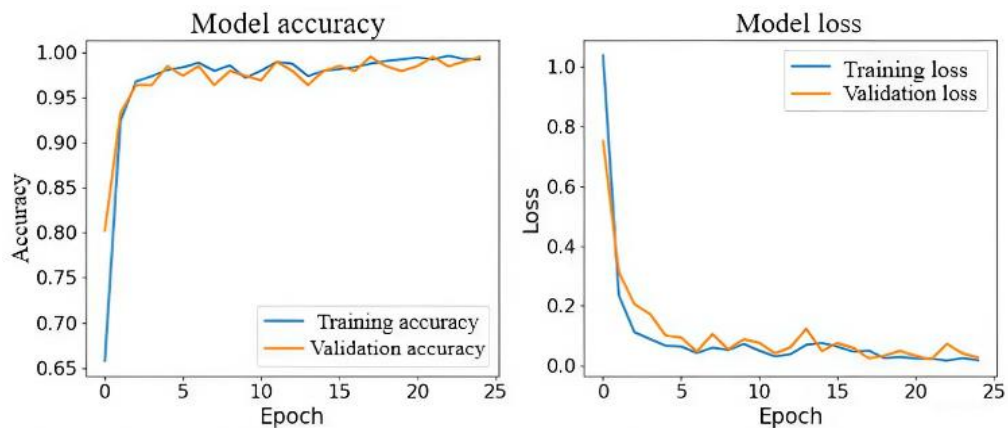


Figure 5. Accuracy and Loss Graphs during the Training Process

Analysis of Figure 5 reveals that model convergence was achieved at epoch 5, with training accuracy surpassing 90%. This finding suggests effective learning efficiency resulting from the synergistic application of the Adam optimizer and data augmentation strategies. The initial training period demonstrated marked improvements in accuracy coupled with substantial loss reduction, signifying efficient knowledge acquisition from the dataset. Following epoch 7, the validation accuracy trajectory exhibited stabilization, though minor variations persisted, reflecting ongoing refinement processes. Learning rate modifications occurred at multiple intervals during training, evidenced by alterations in the loss reduction gradient following specific epochs, validating the effective operation of the ReduceLRonPlateau callback in guiding the model toward global minimum optimization. Upon training completion, the model attained a training accuracy of approximately 99.44% and validation accuracy of 95.83%. The discrepancy between training and validation metrics suggests minimal overfitting, which remains acceptable considering the inherent complexity of multi-class classification involving 10 categories with nuanced visual distinctions.

B. Optimizer Comparison Analysis

This investigation evaluated two prevalent deep learning optimizers: RMSprop and Adam. Each optimizer underwent testing with parameters calibrated for optimal performance assessment RMSprop configured with a learning rate of 0.0005 and rho value of 0.9, while Adam utilized a learning rate of 0.001. Experimentation was performed using the Monk Skin Tone image dataset. Training procedures spanned 25 epochs with a batch size of 32 to maintain comparative objectivity. The Monk Skin Tone dataset selection was motivated by its comprehensive representation of human skin color diversity, presenting significant challenges for deep learning-based image classification. Both optimizers were assessed regarding their abilities to address convergence challenges in complex datasets and model generalization capacity across skin color variations. Table 2 presents the comparative analysis outcomes for both optimizers, detailing accuracy and loss metrics across training, validation, and testing datasets.

TABLE 2
 OPTIMIZER PERFORMANCE COMPARISON

Optimizer	Training Accuracy	Validation Accuracy	Testing Accuracy	Testing Loss
RMSprop	99.71%	97.40%	98.26%	0.0548
Adam	99.44%	95.83%	97.83%	0.1166

The findings in Table 2 reveal that RMSprop achieved a testing accuracy of 98.26%, exceeding Adam's performance by 0.43%, which reached 97.83%. This differential performance is attributable to RMSprop's proficiency in stabilizing gradient variations during training, consequently expediting model convergence. RMSprop additionally exhibited better performance over Adam across training and validation accuracy metrics, attaining 99.71% training accuracy compared to Adam's 99.44%, alongside 97.40% validation accuracy surpassing Adam's 95.83%. Convergence analysis reveals distinct behavioral patterns between optimizers throughout the training process. RMSprop demonstrates better convergence stability, achieving optimal validation accuracy at epoch 23 with minimal fluctuation during training progression. The learning rate reduction mechanism activated at epochs 14 and 19, indicating controlled adaptation to gradient variations and systematic optimization refinement. Conversely, Adam exhibits faster initial convergence, reaching 90% accuracy by epoch

5, but displays greater susceptibility to validation loss oscillations in later epochs, suggesting less stable long-term training dynamics.

Regarding overfitting tendencies, RMSprop maintains a smaller gap between training (99.71%) and testing accuracy (98.26%), suggesting better generalization capabilities and reduced overfitting risks. Adam shows a marginally larger generalization gap (99.44% training vs 97.83% testing), though both remain within acceptable bounds for practical applications. The loss differential further supports this observation, with RMSprop achieving significantly lower testing loss (0.0548) compared to Adam (0.1166), indicating more effective optimization convergence. The performance advantage of RMSprop over Adam can be attributed to their respective operational mechanisms. RMSprop's architecture specifically addresses oscillation and fluctuation challenges commonly encountered during weight updates, enabling more effective adaptation to substantial gradients and resulting in enhanced stability throughout the training procedure [28], [29]. This optimizer's adaptive learning rate mechanism effectively normalizes gradients based on recent magnitudes, providing more stable weight updates particularly beneficial for skin tone classification where subtle feature distinctions require precise parameter adjustments. While Adam offers advantages in learning rate modification and adaptive moment estimation, this algorithm exhibits greater susceptibility to instability in complex classification scenarios [30]. These findings indicate RMSprop's enhanced effectiveness in predicting previously unseen data, representing a crucial factor in managing overfitting and optimizing generalization capabilities [31]. The evidence supporting RMSprop better underscores the critical impact of appropriate optimizer selection on model performance, demonstrating enhanced efficiency in achieving optimal performance on test and validation datasets while reducing training duration and yielding better outcomes.

C. Dataset Balancing Analysis

Initial dataset examination revealed substantial class imbalance, with certain categories significantly overrepresented relative to others. The MST4 category contained the highest representation (261 images), whereas MST10 exhibited minimal representation (111 images). Such imbalance, if unaddressed, could potentially introduce model bias. This research implemented a comprehensive dataset balancing methodology employing a hybrid strategy combining undersampling and data augmentation to address this challenge. Undersampling was executed for categories exceeding 150 representations (MST2, MST4, MST5, and MST6), randomly selecting 150 images from each overrepresented class. Concurrently, data augmentation was applied to underrepresented categories (MST3, MST7, MST8, MST9, and MST10) utilizing various transformation techniques including rotation ($\pm 20^\circ$), translation ($\pm 20\%$), zoom ($\pm 20\%$), brightness adjustments, horizontal flipping, and shear transformations. The dataset balancing strategy resulted in substantial modifications to class distribution. Post-balancing, each category maintained precisely 150 image representations. Subsequently, the dataset underwent equitable partitioning into training (70%, 105 images), validation (15%, 22 images), and testing datasets (15%, 23 images). Models developed using imbalanced datasets potentially exhibit bias toward dominant categories (MST4, MST5, and MST6), whereas models trained on balanced datasets demonstrate more uniform performance across all categories.

D. Model Evaluation

Following training completion, the research team assessed the model utilizing a test dataset comprising 230 images, with uniform distribution of 23 images per MST category. These evaluation outcomes provide essential insights into the model's practical application capabilities. The assessment protocol involved computing various performance indicators including accuracy, precision, recall, and F1-score for individual MST categories, alongside constructing confusion matrices to identify specific classification error patterns. The uniform distribution of 23 images per MST category ensures unbiased model evaluation across the complete skin color spectrum, representing a critical requirement for inclusive and equitable facial recognition technology implementations. Comprehensive analysis of classification errors was undertaken to identify factors influencing accuracy, including lighting variations, image capture perspectives, and diverse facial characteristics across MST categories.

Model assessment on the test dataset yielded high performance with 97.83% accuracy, demonstrating robust generalization capabilities. Testing with previously unseen data produced notable outcomes. Classification report analysis revealed that most categories achieved outstanding performance with precision and recall metrics exceeding 91%. The classification report indicated average precision, recall, and F1-score values of 97.86%, 97.83%, and 97.82% respectively, demonstrating optimal balance between precision and recall metrics. Table 3 presents detailed precision, recall, and F1-score data.

TABLE 3
 CLASIFICATION REPORT

Class	Classification Report			
	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Support
MST1	1,0	1,0	1,0	23
MST2	1,0	0,9565	0,9778	23
MST3	0,9565	0,9565	0,9565	23
MST4	0,9545	0,913	0,9333	23
MST5	0,9583	1,0	0,9787	23
MST6	0,9583	1,0	0,9787	23
MST7	0,9583	1,0	0,9787	23
MST8	1,0	1,0	1,0	23
MST9	1,0	1,0	1,0	23
MST10	1,0	0,9565	0,9778	23
accuracy			0,9783	230
macro avg	0,9786	0,9783	0,9782	230
weighted avg	0,9786	0,9783	0,9782	230

Table 3 demonstrates that the MobileNetV2-based skin color classification model achieves optimal performance across all MST categories. High precision, recall, and F1-score values for each category indicate the model's proficiency in effective classification and comprehensive relevant data identification. Precision, recall, and F1-score metrics all approximate 0.978 for both macro and weighted averages. These findings confirm that the model functions with high efficiency across multiple categories while demonstrating better performance throughout all MST classifications. The macro average signifies optimal model performance across all categories irrespective of class distribution variations, while the weighted average indicates optimal performance considering category proportions within the dataset. MST1, MST8, and MST9 achieved excellent performance metrics with precision, recall, and F1-score values of 1.0. This result indicates highly accurate sample identification without classification errors. These findings demonstrate the model's capability to reliably identify skin colors for these specific categories while substantially minimizing prediction errors.

Performance metrics show relative decline for MST4, recording an F1-score of 0.9333, precision of 0.9545, and recall of 0.913. Despite maintaining satisfactory performance, these scores remain marginally lower than other categories, suggesting elevated classification error rates for this particular class. This performance reduction may stem from two primary factors. Firstly, red color variations within MST4 introduce recognition complexity for the model. Secondly, residual data imbalance in this category may compromise the model's red color pattern recognition capabilities. Achieving 97.83% accuracy on 230 test samples, this model establishes its reliability for skin color classification utilizing the MST scale. These metrics confirm the model's capability to classify skin colors with high accuracy across diverse applications. The model demonstrates suitability for fashion color recommendation systems and various skin color recognition applications. Most categories exhibit consistent performance with excellent results.

The MobileNetV2 classification model developed in this research achieved 97.83% accuracy, surpassing Mbatha study outcomes, which reported accuracy ranges of 85.45–97.16% utilizing CNN regression models for skin color assessment employing the Monk Skin Tone scale. This performance differential can be attributed to several methodological distinctions. Primarily, this research implemented comprehensive dataset balancing strategies through augmentation and undersampling techniques, effectively addressing class imbalance issues prevalent in skin tone datasets. Additionally, the emphasis on optimizing MobileNetV2 architecture with transfer learning enables the model to extract more nuanced skin color features compared to Mbatha multi-model approach. Despite this study's more limited dataset (1,729 images) compared to Mbatha (21,375 images), the specialized preprocessing and normalization techniques implemented contribute to enhanced accuracy. While Mbatha research provides significant insights regarding skin tone estimation under varied lighting conditions, this study demonstrates that appropriately optimized MobileNetV2 can achieve better accuracy in skin tone classification despite dataset limitations [11].

For comprehensive understanding of model classification error patterns, the research team conducted confusion matrix analysis. This matrix elucidates the model's classification mechanisms for each category while visualizing accurate and inaccurate prediction percentages across categories. This matrix enables researchers to identify various recurring error types, determining whether the model exhibits tendencies toward false positives or false negatives. Additionally, this analysis enables more precise model performance measurement, allowing researchers to identify categories requiring optimization during enhancement processes. The confusion matrix serves as an invaluable instrument for analyzing model limitations and formulating modification strategies to comprehensively optimize accuracy. Figure 6 presents the confusion matrix utilized in this research.

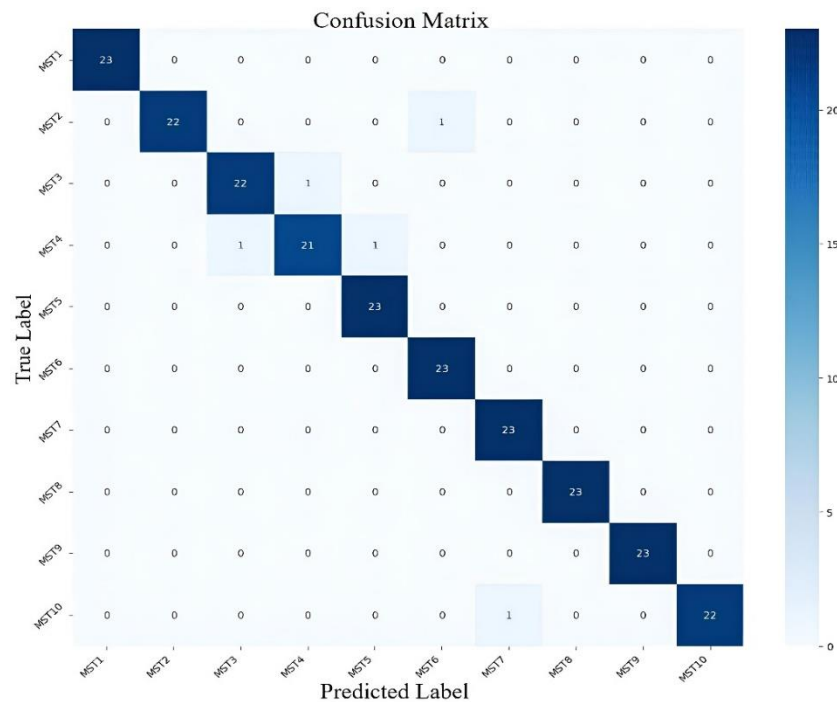


Figure 6. Confusion Matrix

According to Figure 6, the confusion matrix depicts the model's classification behavior across ten categories (MST1 through MST10). Most diagonal matrix values range between 21-23, indicating high accuracy levels across most categories, as the visualization confirms the model's robust predictive capabilities. Several misclassification instances yielded noteworthy observations. The MST4 category demonstrated the highest misclassification error frequency, with one instance misclassified as MST3 and another as MST5 this pattern potentially indicates feature similarities among these three categories. Furthermore, prediction errors occurred for MST2 classified as MST6 and MST10 classified as MST7. Comprehensively, the model exhibits optimal performance for MST1, MST5, MST6, MST7, MST8, and MST9 categories. These findings suggest that these categories possess adequately distinctive features, enabling effective model discrimination. Nevertheless, certain misclassifications were observed in MST2, MST3, MST4, and MST10 categories. Despite relatively minor occurrence frequency, these instances represent focal points for optimization in subsequent research. Misclassification patterns within the confusion matrix highlight areas requiring model optimization. Prediction errors frequently occur between categories sharing similar characteristics. Consequently, strategies enhancing the model's discriminatory capabilities for these categories warrant consideration. Overall, evaluation outcomes demonstrate the model's substantial generalization capabilities and suitability for Skin Tone web application implementation in practical contexts. Better performance across all MST categories confirms that the implemented dataset balancing techniques successfully addressed class imbalance challenges in the original dataset.

The MST classification task presents several inherent challenges that merit consideration. Dataset imbalance constituted a primary obstacle, with initial distribution ranging from 111 images (MST10) to 261 images (MST4), necessitating systematic balancing through strategic undersampling and augmentation. This imbalance reflects real-world demographic variations but introduces bias toward overrepresented categories during training. Classification complexity emerges from the subjective nature of skin tone boundaries, particularly between adjacent MST categories. The confusion matrix reveals most misclassifications occur

between neighboring classes (MST4 → MST3/MST5, MST2 → MST6), suggesting perceptual similarity challenges that extend beyond algorithmic limitations. These errors highlight the continuous spectrum nature of human skin pigmentation, where discrete categorical boundaries may not fully capture natural variations. Environmental factors pose additional challenges, including lighting conditions, image quality, and photographic settings that can significantly influence skin tone perception. The MST scale's development for standardized conditions may not fully encompass the variability encountered in practical applications, requiring robust preprocessing and normalization strategies to maintain classification consistency across diverse image sources.

E. Web Implementation Evaluation and Clothing Color Recommendations

The trained model underwent successful integration into the SkinToneAI web application, enabling users to easily upload facial photographs and receive skin tone classifications alongside clothing color recommendations. These recommendations are meticulously aligned with the user's identified skin color characteristics. The application architecture was carefully designed to prioritize both usability and processing efficiency, ensuring that users receive real-time analysis results without requiring any technical expertise in image processing or skin tone classification. The design of the application interface places a strong emphasis on intuitiveness and ease of navigation. It features layouts that are optimized for user experience, making it simple for users to interact with the system. The primary interface offers a seamless image upload functionality, accompanied by clear system guidelines that explain how the application works. Once the image is processed, the classification results are visually presented in an easily understandable format. These results are paired with suggested clothing color palettes that are optimized according to the identified skin tone categories, giving users personalized and relevant fashion advice. Figure 7 illustrates the web application interface.

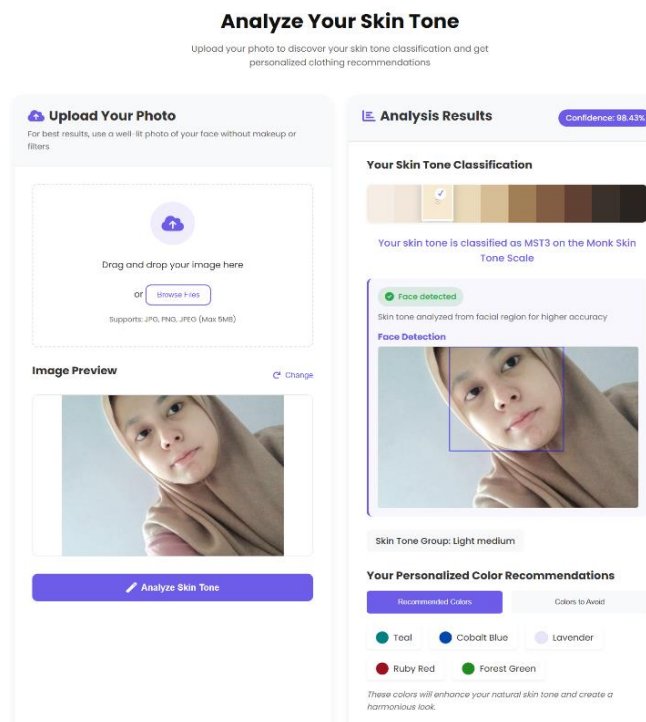


Figure 7. SkinToneAI Web App Interface

As depicted in Figure 7, the application interface incorporates simple, intuitive design principles, enabling users to effortlessly upload photographs and access analysis results. Key features of this web application include photo upload functionality, allowing users to submit facial photographs with guidelines specifying optimal lighting conditions and makeup-free requirements for enhanced accuracy. The application delivers Skin Tone analysis, wherein the system processes photographs and presents classification results based on the MST scale, including prediction confidence levels. For visualization purposes, the application displays the comprehensive MST scale with highlighted classification results, facilitating user understanding of their Skin Tone position within the complete spectrum. Based on classification outcomes, the system additionally provides recommendations for optimal clothing colors alongside colors to avoid.

The application's clothing color recommendations derive from complementary color theory and established fashion principles. Researchers developed mappings between MST categories and corresponding clothing colors for the recommendation system. This mapping foundation incorporates color theory and previous

research findings. Seasonal color theory implementation (spring, summer, autumn, and winter) proves essential for determining clothing colors compatible with users' skin characteristics, resulting in harmonious appearances and enhanced self-confidence [2]. Table 4 presents examples of clothing color recommendations across various MST categories. These recommendations emerge from color theory and contrast principles, which establish that specific colors can enhance or diminish particular Skin Tone characteristics [32]. For instance, lighter skin tones (MST1-MST2) benefit from dark, saturated colors such as Navy Blue and Burgundy creating appealing contrast, while darker skin tones (MST9-MST10) are complemented by bright colors like Yellow and Fuchsia, which accentuate the natural beauty of the Skin Tone.

TABLE 4
 CLOTHING COLOR RECOMMENDATIONS BASED ON MST CLASSIFICATION

MST Category	Clothing Color Recommendations	Colors to Avoid
MST1-MST2 (Light)	Navy Blue, Royal Purple, Emerald Green, Burgundy, Sapphire Blue	Orange, Bright Yellow, Pastel
MST3-MST4 (Light Medium)	Teal, Cobalt Blue, Lavender, Ruby Red, Forest Green	Brown, Khaki, Olive
MST5-MST6 (Medium)	Coral, Turquoise, Olive Green, Royal Blue, Magenta	Neon Color, White, Black
MST7-MST8 (Medium Deep)	Gold, Mustard Yellow, Orange, Kelly Green, Electric Blue	Pastel Colors, Beige, Silver
MST9-MST10 (Deep)	Bright Yellow, Fuchsia, Lime Green, Bright Orange, Aqua	Dark Color, Brown, Navy

This implementation demonstrates the feasibility of translating deep learning models into practical applications, though it requires further evaluation under diverse real-world conditions. Real-time processing optimization represents a priority, requiring model compression techniques and edge computing implementations for mobile deployment. Fairness evaluation mechanisms warrant development to assess potential bias across demographic groups, ensuring equitable performance regardless of user characteristics. Dataset expansion should prioritize geographic and demographic diversity, incorporating skin tone variations across different populations and environmental conditions to enhance global applicability. Technical enhancements should explore advanced augmentation strategies specific to skin tone classification, investigate ensemble methods combining multiple CNN architectures, and develop adaptive threshold systems that account for classification uncertainty. Integration of multi-modal approaches, combining facial analysis with contextual information such as lighting conditions and user preferences, could improve recommendation accuracy and personalization. While this research achieved high classification accuracy, several limitations must be acknowledged. The dataset size (1,729 images) remains relatively small compared to large-scale computer vision datasets. Environmental factors such as lighting conditions and image quality variations in real-world deployment scenarios require further investigation. Cross-platform compatibility studies and longitudinal user experience research would provide valuable insights for sustained practical implementation.

IV. CONCLUSION

This research has successfully developed a skin tone image classification system utilizing the MobileNetV2 architecture, incorporating the Monk Skin Tone Scale dataset obtained from 'Monk, Ellis. "Monk Skin Tone Scale," 2019. (<https://skintone.google>). The dataset selection was predicated on its better image quality captured across diverse lighting environments and systematic categorization into 10 distinct skin tone gradations. The MST dataset demonstrates a comparative advantage through its capacity to represent skin phenotypic diversity with greater inclusivity compared to the traditional Fitzpatrick scale [10]. The architectural configuration employed input dimensions of 224×224 pixels, an initial learning rate of 0.001, training for 25 epochs, and integration of GlobalAveragePooling2D followed by Dense layers with Dropout mechanisms. Dataset balancing was accomplished through strategic undersampling and data augmentation techniques to establish equitable representation of 150 samples per class. Data augmentation methodologies encompassed rotation, translation, zoom operations, horizontal flipping, and brightness adjustments, contributing substantially to data distribution equilibration while preserving information integrity. The dataset balancing procedure yielded a 70:15:15 distribution, allocating 105 images per class for training, 22 for validation, and 23 for testing. Training outcomes demonstrated final training accuracy of 99.44% and validation accuracy of 95.83%. Model assessment using an independent test dataset produced 97.83% accuracy with 0.1166 loss value, signifying robust generalization capabilities. Comprehensive analysis revealed highly satisfactory classification performance across the complete MST class spectrum, with precision, recall, and F1-score metrics consistently exceeding 0.93. Several classes (MST1, MST8, and MST9) achieved excellent scores of 1.0 across all evaluation parameters. Comparative analysis between optimization algorithms revealed that RMSprop outperformed Adam across all evaluation metrics, achieving 98.26% test accuracy with 0.0548 loss compared to Adam's 97.83%

accuracy with 0.1166 loss. This 0.43% improvement demonstrates RMSprop better stability in gradient optimization for skin tone classification tasks. The model was integrated into the SkinToneAI web application for practical validation. The system enables facial photograph analysis and generates clothing color recommendations based on identified skin tone classifications. This investigation demonstrates that MobileNetV2 architecture provides effective performance for skin tone classification tasks with computational efficiency suitable for web applications. The study contributes to the growing body of work on inclusive AI systems while acknowledging the complexity of human skin tone classification. Results indicate promise for consumer applications, with the understanding that continued refinement and evaluation remain essential for responsible deployment in diverse user environments.

BIBLIOGRAPHY

- [1] F. W. Prabowo, A. Homaidi, and A. Lutfi, "Deteksi Warna Kulit Menggunakan Metode Deep Learning Dengan CNN (Convolutional Neural Network) Untuk Menentukan Kecocokan Warna Kulit dan Warna Busana," *E-Link: Jurnal Teknik Elektro dan Informatika*, vol. 19, no. 2, p. 186, Jul. 2024, doi: 10.30587/e-link.v19i2.8128.
- [2] X. Su *et al.*, "Personalized clothing recommendation fusing the 4-season color system and users' biological characteristics," *Multimed Tools Appl*, vol. 83, no. 5, pp. 12597–12625, Jul. 2023, doi: 10.1007/s11042-023-16014-4.
- [3] R. Fayyadhila, A. Junaidi, and N. A. Prasetyo, "Implementasi Deep Learning Untuk Klasifikasi Citra Undertone Menggunakan Algoritma Convolutional Neural Network," *Journal of Dinda : Data Science, Information Technology, and Data Analytics*, vol. 1, no. 2, pp. 52–62, Aug. 2021, doi: 10.20895/dinda.v1i2.366.
- [4] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.
- [5] S. Saiwao, S. Arwatchananukul, L. Mungmai, W. Preedalikit, and N. Aunsri, "Human skin type classification using image processing and deep learning approaches," *Heliyon*, vol. 9, no. 11, p. e21176, Nov. 2023, doi: 10.1016/j.heliyon.2023.e21176.
- [6] V. D. Br Sebayang and I. G. N. L. W. Kusuma, "Klasifikasi Jenis Jerawat Berdasarkan Citra Menggunakan Convolutional Neural Network dengan Arsitektur MobileNetV2," *JURNAL FASILKOM*, vol. 14, no. 3, pp. 766–774, Dec. 2024, doi: 10.37859/jf.v14i3.8202.
- [7] Y. Gulzar, "Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique," *Sustainability*, vol. 15, no. 3, p. 1906, Jan. 2023, doi: 10.3390/su15031906.
- [8] M. G. Somoal and A. R. Dzirkillah, "Komparasi MobileNETV2 dengan Kustomisasi Transfer Learning dan Hyperparameter untuk Identifikasi Tumor Otak," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 1, pp. 229–240, Feb. 2025, doi: 10.25126/jtiik.2025129582.
- [9] V. P. Matias and J. Batista Neto, "Enhancing Fairness in Machine Learning: Skin Tone Classification Using the Monk Skin Tone Scale," in *Anais Estendidos da XXXVII Conference on Graphics, Patterns and Images (SIBGRAPI Estendido 2024)*, Sociedade Brasileira de Computação - SBC, Sep. 2024, pp. 76–81. doi: 10.5753/sibgrapi.est.2024.31648.
- [10] C. Schumann, G. O. Olanubi, A. Wright, E. Monk, C. Heldreth, and S. Ricco, "Consensus and Subjectivity of Skin Tone Annotation for ML Fairness," Jan. 2024. [Online]. Available: <http://arxiv.org/abs/2305.09073>
- [11] S. K. Mbatha, M. J. Booyesen, and R. P. Theart, "Skin Tone Estimation under Diverse Lighting Conditions," *J Imaging*, vol. 10, no. 5, p. 109, Apr. 2024, doi: 10.3390/jimaging10050109.
- [12] A. Mahbod, G. Schaefer, C. Wang, G. Dorffner, R. Ecker, and I. Ellinger, "Transfer learning using a multi-scale and multi-network ensemble for skin lesion classification," *Comput Methods Programs Biomed*, vol. 193, p. 105475, Sep. 2020, doi: 10.1016/j.cmpb.2020.105475.
- [13] P. N. Srinivasu, J. G. SivaSai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM," *Sensors*, vol. 21, no. 8, p. 2852, Apr. 2021, doi: 10.3390/s21082852.
- [14] V. A. O. Nancy, P. Prabhavathy, M. S. Arya, and B. S. Ahamed, "Comparative study and analysis on skin cancer detection using machine learning and deep learning algorithms," *Multimed Tools Appl*, vol. 82, no. 29, pp. 45913–45957, Dec. 2023, doi: 10.1007/s11042-023-16422-6.
- [15] K. Alomar, H. I. Aysel, and X. Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J Imaging*, vol. 9, no. 2, p. 46, Feb. 2023, doi: 10.3390/jimaging9020046.
- [16] R. J. Hendri Butar-Butar and N. L. Marpaung, "Deep Learning untuk Identifikasi Daun Tanaman Obat Menggunakan Transfer Learning MobileNetV2," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 8, no. 2, pp. 142–148, May 2023, doi: 10.30591/jpit.v8i2.5217.
- [17] A. Furqon, K. Malik, and F. N. Fajri, "Detection of Eight Skin Diseases Using Convolutional Neural Network with MobileNetV2 Architecture for Identification and Treatment Recommendation on Android Application," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 10, no. 2, pp. 373–384, Jul. 2024, doi: 10.26555/jiteki.v10i2.28817.
- [18] Tensorflow, "ModelCheckpoint," Tensorflow. Accessed: Jun. 25, 2025. [Online]. Available: https://www.tensorflow.org/api_docs/python/tf/keras/callbacks/ModelCheckpoint
- [19] T. Ahmed, F. S. Mou, and A. Hossain, "SCCNet: An Improved Multi-Class Skin Cancer Classification Network using Deep Learning," in *2024 3rd International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)*, IEEE, Apr. 2024, pp. 1–5. doi: 10.1109/ICAEEE62219.2024.10561672.
- [20] Qorry Aina Fitroh and Shofwatul 'Uyun, "Deep Transfer Learning untuk Meningkatkan Akurasi Klasifikasi pada Citra Dermoskopi Kanker Kulit," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 12, no. 2, pp. 78–84, May 2023, doi: 10.22146/jnteti.v12i2.6502.
- [21] W.-C. Cheng, "Monastic Color Reproduction: A Software Tool for Printing and Assessing the Monk Skin Tone Scale," *Journal of Imaging Science and Technology*, vol. 68, no. 5, pp. 1–13, Sep. 2024, doi: 10.2352/J.ImagingSci.Technol.2024.68.5.050404.
- [22] K. Struniawski and R. Kozera, "TfELM: Extreme Learning Machines framework with Python and TensorFlow," *SoftwareX*, vol. 27, p. 101833, Aug. 2024, doi: 10.1016/j.softx.2024.101833.
- [23] G. Kruseman, "A Flexible, Extensible, Machine-Readable, Human-Intelligible, and Ontology-Agnostic Metadata Schema (OIMS)," *Front Sustain Food Syst*, vol. 6, Mar. 2022, doi: 10.3389/fsufs.2022.767863.

- [24] S. S. N. Challapalli, P. Kaushik, S. Suman, B. D. Shivahare, V. Bibhu, and A. D. Gupta, "Web Development and performance comparison of Web Development Technologies in Node.js and Python," in *2021 International Conference on Technological Advancements and Innovations (ICTAI)*, IEEE, Nov. 2021, pp. 303–307. doi: 10.1109/ICTAI53825.2021.9673464.
- [25] L. Albeshier and R. Alfayez, "An Observational Study on Flask Web Framework Questions on Stack Overflow (SO)," *IET Software*, vol. 2024, no. 1, p. 1905538, Jan. 2024, doi: 10.1049/sfw2/1905538.
- [26] G. Kaur, H. Malhotra, and T. Gupta, "Choosing the Right Model: A Comprehensive Analysis of Outfit Recommendation Systems," *Int J Comput Appl*, vol. 183, no. 12, pp. 13–20, Jun. 2021, doi: 10.5120/ijca2021921413.
- [27] S. S. Ali, J. H. Al' Ameri, and T. Abbas, "Face Detection Using Haar Cascade Algorithm," in *2022 Fifth College of Science International Conference of Recent Trends in Information Technology (CSCTIT)*, IEEE, Nov. 2022, pp. 198–201. doi: 10.1109/CSCTIT56299.2022.10145680.
- [28] M. Taqiyuddin, K. Adi, O. Dwi Nurhayati, and H. Ochi, "Comparison of Optimizers for Drone Signal Detection Using Convolutional Neural Networks (CNN)," *E3S Web of Conferences*, vol. 448, p. 02025, Nov. 2023, doi: 10.1051/e3sconf/202344802025.
- [29] G. A. Saputra and I. M. A. Agastya, "Betta Fish Identification System Based On Convolutional Neural Network," *Journal of Applied Informatics and Computing*, vol. 8, no. 2, pp. 443–452, Nov. 2024, doi: 10.30871/jaic.v8i2.8449.
- [30] A. Murthy, P. S. Rao, N. S. Pallavi, N. Kharvi, B. R. Neha, and K. Poojary, "Optimizing Convolutional Neural Networks: A Comparative Study of Gradient-Descent, Adam, and RMSprop Optimizers for Accuracy and Loss in Apple Leaf Disease Detection," in *2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)*, IEEE, Aug. 2024, pp. 1–6. doi: 10.1109/NMITCON62075.2024.10699138.
- [31] Q. Zhang, Y. Zhou, and S. Zou, "Convergence Guarantees for RMSProp and Adam in Generalized-smooth Non-convex Optimization with Affine Noise Variance," *Transactions on Machine Learning Research (TMLR)*, Mar. 2025, Accessed: May 02, 2025. [Online]. Available: <http://arxiv.org/abs/2404.01436>
- [32] D. I. Perrett and R. Sprengelmeyer, "Clothing Aesthetics: Consistent Colour Choices to Match Fair and Tanned Skin Tones," *Iperception*, vol. 12, no. 6, Nov. 2021, doi: 10.1177/20416695211053361.