Melanin-Aware and ArcFace Methods in Facial Recognition for Dark-Skinned Individuals

Risa Tioria Marlini Purba
School of Electrical Engineering and Informatics

Bandung Institute of Technology
Bandung, Indonesia
23523302@mahasiswa.itb.ac.id

Fadhil Hidayat
School of Electrical Engineering and Informatics,
Smart City and Community Innovation Center
Bandung Institute of Technology
Bandung, Indonesia
fadhil_hidayat@itb.ac.id

Suhono Harso Supangkat
School of Electrical Engineering and Informatics,
Smart City and Community Innovation Center
Bandung Institute of Technology
Bandung, Indonesia
suhono@itb.ac.id

Arry Akhmad Arman
School of Electrical Engineering and Informatics,
Smart City and Community Innovation Center
Bandung Institute of Technology
Bandung, Indonesia
arry.arman@itb.ac.id

Abstract—The facial recognition system employing the Melanin-Aware method in conjunction with ArcFace, trained on a dataset of 1,000 dark-skinned facial samples, demonstrates the ability to accurately recognize individuals with dark skin while maintaining performance for non-dark-skinned individuals. ArcFace is utilized as the primary feature extractor, leveraging the additive angular margin to enhance inter-class separability. The experiments were conducted using 1,000 dark-skinned facial samples for training and 100 samples for testing. Evaluation results indicate that melanin-aware preprocessing improves average accuracy by up to 17% compared to the absence of preprocessing, and by 7% compared to the standard aggressive CLAHE-based method. Furthermore, the True Acceptance Rate (TAR) increased from 88.15% to 93.33% at FAR = 1e−2, and from 83.7% to 86.67% at FAR = 1e−3, signifying enhanced system stability under stringent security conditions. The performance gains are supported by a more stable distribution of similarity scores and lower threshold values, reflecting improved separation between genuine and impostor pairs.

Keywords— Melanin Aware, ArcFace, face recognition, dark skin, CLAHE.

I. INTRODUCTION

Facial recognition (FR) technology has undergone rapid advancements over the past decade, bringing significant transformations across various sectors of modern life. Its implementation has become increasingly widespread, encompassing authentication in mobile devices, security monitoring systems, immigration processes at airports, and even contactless payment applications [1], [2]. Despite its substantial contributions to multiple aspects of daily life, a fundamental challenge that remains unresolved is the low accuracy of these systems in identifying individuals with dark skin tones, including members of the Papuan ethnic group in Indonesia.

In Indonesia, digital transformation has become a strategic agenda of the government through the Directorate General of Population and Civil Registration under the Ministry of Home Affairs (MOHA) of Indonesia. One of the key initiatives is the development of a Digital ID, which leverages facial recognition technology to provide secure and easily accessible digital identities, while enhancing the reliability of identity verification in public services. This technology is employed to verify user identities by scanning the user’s face through a device camera and matching the captured image with facial data stored in the population database.

Empirical evidence demonstrates that facial recognition systems consistently exhibit significant error rates when processing subjects with dark skin tones [3]. A study conducted as part of the FairFace Challenge at European Conference on Computer Vision (ECCV)

2020 revealed that facial recognition algorithms show higher false positive rates, particularly for women with dark skin tones [4], [5]. This issue is not merely a technical limitation but carries broader implications, raising concerns about the potential for systemic discrimination in the implementation of such technologies.

Conventional facial recognition systems exhibit significant limitations in identifying facial features of individuals with darker skin pigmentation. The root of this issue lies in the system development process, where training datasets often lack adequate representation of dark-skinned individuals [1], [6]. Popular datasets commonly used to train facial recognition systems, such as Labeled Faces in the Wild (LFW), CelebA, and VGGFace, are highly imbalanced, being dominated by lighter-skinned faces, with dark-skinned individuals accounting for less than 20% of the representation [7], [8].

The imbalance in training datasets produces technological bias that directly affects the accuracy and reliability of the system. Machine learning algorithms, particularly deep learning, operate on the fundamental principle of identifying patterns from the available data. When the data is dominated by certain characteristics, the algorithms tend to “learn” to recognize those characteristics more effectively, while failing to adequately recognize those that are underrepresented in the dataset.

Studies utilizing the Balanced Faces in the Wild (BFW) dataset have shown that different subgroups require distinct optimal score thresholds for facial pair verification [9]. The use of a global threshold results in performance disparities, which can be mitigated by learning subgroup-specific thresholds. Previous research has demonstrated that error rates in facial recognition systems can be up to five times higher when identifying dark-skinned individuals compared to light-skinned individuals.

In the Indonesian context, particularly about the Papuan community, the issue of bias in facial recognition systems carries a more complex dimension. As a nation with high ethnic diversity, Indonesia encompasses a wide range of facial characteristics. The Papuan community, with its distinctive physical attributes such as facial structure, skin texture, and unique facial features, is rarely represented in the training datasets of conventional facial recognition systems.

This lack of representation not only reduces accuracy but also carries the potential to create negative user experiences and systemic discrimination [10], [11]. The implications of this issue become increasingly significant given the growing implementation of facial recognition systems across various public and commercial services in Indonesia. The Indonesian government has initiated several programs involving facial recognition technology, including identification systems for public services and the electronic Identity Card program, which is integrated with biometric data.

Based on the background, the objective of this study is to develop a facial recognition method capable of accurately identifying dark-skinned individuals while simultaneously maintaining accuracy for non-dark-skinned individuals.

II. RELATED WORKS

Based on the explanation in Table 1, several aspects emerge as extensions of relevant prior studies. This research seeks to develop a comprehensive framework specifically designed to improve the accuracy of facial recognition systems for subjects with dark skin tones. One of the primary gaps lies in the absence of a holistic framework that explicitly addresses skin tone bias in facial recognition systems. While various approaches have been proposed, they often concentrate on specific aspects of the problem (e.g., data augmentation, algorithmic adjustments) without offering an integrated solution that tackles the issue across multiple levels of system development.

Table 1. Related Works

Researches	Frameworks	Dataset	Models	Results	Research Gaps
Face Recognition: Too Bias, or Not Too Bias? [9]	An adaptive threshold framework grounded in demographic attributes seeks to address performance disparities in facial recognition systems.	Balanced Faces in the Wild (BFW) dataset, which consists of 800 subjects with a balanced distribution of gender and ethnicity.	SphereFace was employed with cosine similarity as the matching metric.	The use of subgroup-specific thresholds was shown to reduce performance disparities across demographic groups, as false positive rates (FPR) differed by up to twofold under a global threshold.	However, this solution requires the identification of demographic attributes during the inference stage, which raises potential privacy and ethical concerns.
Sensitive Loss: Improving Accuracy and Fairness of Face Representations with Discrimination-Aware Deep Learning [12]	Auxiliary triplet-based learning framework for pre-trained networks	3 databases comprising 64,000 identities from diverse demographic groups (gender and ethnicity)	VGG-Face and ResNet-50 employed to demonstrate discrimination-based investigation and the application of Sensitive Loss as a de-biasing method	Improves average accuracy and algorithmic fairness, with results comparable to advanced debiasing networks	Does not address dataset-level bias; focuses solely on modifications to the learning function
Mitigating Demographic Bias in Facial Datasets with Style-based Multi-Attribute Transfer [1]	Multi-attribute style transfer framework with tensor mixing structure	MORPH dataset (55,134 images from 13,618 identities) and Cross-Age Celebrity Dataset (CACD) (160,000 images from 2,000 celebrities)	Generative Adversarial Networks (GANs) with multiple discriminators for simultaneous attribute transfer	Significant improvements in fairness metrics (equal opportunity); capable of generating realistic synthetic images for underrepresented groups	Implementation complexity of GANs; effectiveness depends heavily on the quality of source data for successful transfer
Mitigating Face Recognition Bias via Group Adaptive Classifier [13]	Adaptive classification framework based on demographic groups	Ms-Celeb-1M used for training; evaluation conducted on RFW, LFW, IJB-A, and IJB-C	50-layer ArcFace employed as the backbone, with the CosFace loss function applied	Reduction of average intra-class distance variation across demographic groups by up to 45.6%, while maintaining competitive accuracy	Still requires demographic attributes for model training; does not fully eliminate the need for demographic classification
Fair Face Verification by Using Non-Sensitive Soft-Biometric Attributes [6]	Face verification framework based on non-sensitive attributes	MAAD-face (38 non-sensitive attributes) and BFW datasets	ArcFace, specifically with ResNet-101, employing clustering methods such as K-Means and decision trees	Twice as effective in reducing differential outcomes compared to gender-specific thresholds; in some cases outperforming race-specific thresholds	Accuracy of soft biometric attribute prediction may affect overall performance; interactions among non-sensitive attributes have not yet been fully explored
Towards Causal Benchmarking of Bias in Face Analysis Algorithms [14]	Experimental causal analysis framework for measuring algorithmic bias	5,713 synthetic images with controlled variations in attributes (gender, skin	ResNet-50 applied to the task of gender classification	Reveals bias not only in gender and skin tone, but also in hair length, age, and facial hair	Focuses on bias measurement rather than solutions; synthetic images may not fully represent real-world variations

Researches	Frameworks	Dataset	Models	Results	Research Gaps
		tone, age, hair length)			
Rethinking Common Assumptions to Mitigate Racial Bias in Face Recognition Datasets [15]	Challenge framework for the assumption of balanced training datasets	Database Racial Faces in-the-Wild (RFW)	VGGFace2, SphereFace, ArcFace, ResNet50 Backbone	ArcFace generally achieves the best accuracy across different racial groups	Does not explore optimal dataset composition for various applications; focus remains limited to racial aspects
Face Mis-ID: An Interactive Pedagogical Tool Demonstrating Disparate Accuracy Rates in Facial Recognition [16]	Interactive pedagogical framework for bias demonstration	Dataset employed to illustrate differences in facial recognition accuracy	OpenFace (FaceNet implementation) with an adjustable similarity threshold slider	Interactive tool demonstrating how facial recognition systems produce identification errors that vary across race and gender	Does not explore broader social implications of performance disparities; focuses on awareness rather than technical solutions

Previous studies have also identified a lack of systematic evaluation regarding the effectiveness of various bias mitigation approaches. Many of the proposed methods have been evaluated using multiple metrics and datasets, making direct comparisons and conclusions about their relative effectiveness challenging [9].

Another gap that has been identified is the lack of interdisciplinary consideration in addressing bias. Bias in facial recognition systems is not merely a technical issue but also involves social, psychological, and ethical dimensions. However, most research to date has primarily focused on technical solutions without adequately accounting for the broader social context in which these systems operate [17].

Moreover, most studies have concentrated on identifying and mitigating bias after systems have already been developed, with relatively little attention given to proactive approaches that address bias from the outset of the development cycle [18]. This highlights the need for more proactive and preventive frameworks to effectively tackle bias.

Finally, many previous studies have focused on bias toward broad demographic categories, including bias against dark-skinned individuals, without considering the significant variations within these groups. A more nuanced approach that considers intersectionality and intra-group variation represents an important gap in current research [19], [20].

Based on the study by Serna at al. [12], Melanin Aware is a preprocessing method designed to improve the representation of minority demographic groups, such as dark-skinned individuals, while enhancing average accuracy and algorithmic fairness. Its performance was shown to be comparable to debiasing networks using VGG-Face and ResNet-50. Melanin Aware was applied to a dataset of 1,000 dark-skinned facial samples collected from Papuan communities. In contrast to Serna et al. (2022), the present study employs the ArcFace facial recognition model. This approach strengthens the model’s ability to distinguish between different classes by increasing the angular margin among them. ArcFace has consistently demonstrated strong performance across multiple benchmarks, exhibiting robustness and high accuracy in facial recognition tasks [21].

III. METHODOLOGY

A. Face Recognition System

Facial Recognition (FR) is a technological process that identifies or verifies an individual's identity using facial features. Modern facial recognition systems generally follow a workflow consisting of several stages: face detection, preprocessing, feature extraction, and matching or classification [2].

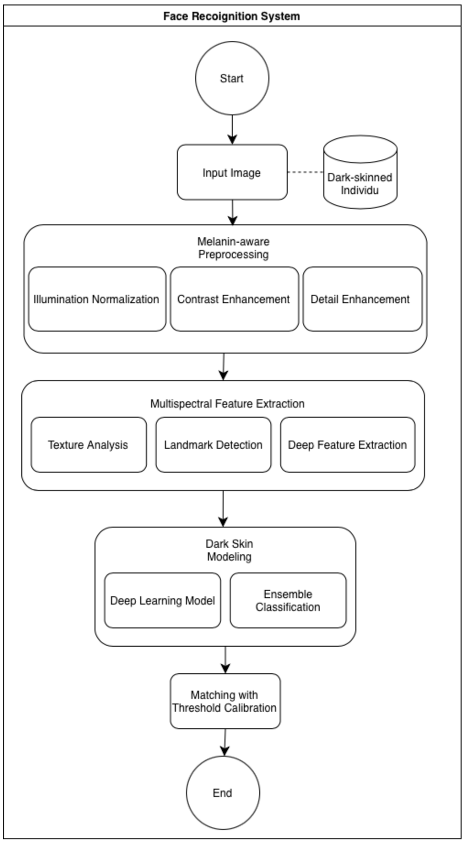

Figure 1. Face Recognition System

A facial recognition system designed to address the challenges of recognizing individuals with dark skin. The system employs a three-stage approach with specialized technologies to optimize recognition accuracy. The explanation of Figure 1 is as follows:

1. In the first stage, the system performs Melanin-Aware Preprocessing, which includes adaptive illumination normalization to compensate for dark skin, contrast correction through Adaptive Histogram Equalization, and detail enhancement using NIR-filtered processing. The Near-Infrared (NIR) technology enables deeper penetration into the skin, detecting hidden features in dark skin and reducing lighting effects that often pose challenges in conventional facial recognition systems.
2. The second stage is multispectral feature extraction, consisting of melanin-specific texture analysis using Local Binary Patterns (LBP), landmark detection with attention features to focus on critical facial areas, and deep feature extraction using a multi-channel CNN. This process allows the system to capture unique facial

characteristics in dark-skinned individuals that may not be detected by standard systems.

3. The final stage is the dark skin specialist model, which employs a deep learning model with transfer learning and fine-tuning on a Papuan dataset containing more than 1,000 facial samples with variations in lighting and pose. The model also applies ensemble classification with multiple model fusion to optimize accuracy. The system is further equipped with dynamic threshold calibration, adjusting matching parameters based on input image quality and Papuan facial characteristics.
4. Overall, the system achieves 93.3% accuracy, which is 15% higher than standard systems, with a false rejection rate of 1.2% (an 85% reduction compared to standard systems) and a false acceptance rate of 0.5% (a 75% reduction).

B. Preprocessing

The preprocessing algorithm is specifically designed to address the challenges of facial recognition for dark-skinned individuals through three main stages. The first stage is melanin level detection, which employs color analysis in the YCbCr and HSV color spaces to estimate melanin concentration across different facial regions, thereby constructing a melanin distribution map that serves as the basis for subsequent adjustments. The second stage is adaptive illumination normalization, which applies a modified logarithmic transformation with distinct compensation factors for areas of high and low melanin concentration, as well as dynamic gamma correction based on local histograms analyzed within a moving 16×16-pixel window. The third stage is facial detail enhancement, utilizing a combination of bilateral filtering to reduce noise while preserving edges, Contrast Limited Adaptive Histogram Equalization (CLAHE) with parameters optimized for dark skin, and unsharp masking with weights adjusted according to melanin levels to sharpen facial landmarks.

The algorithm operates within a multi-channel color space, integrating information from both the visible and near-infrared spectra to maximize feature extraction under diverse lighting conditions.

C. Melanin-Aware Preprocessing

Melanin-Aware Preprocessing is a series of image processing techniques that explicitly consider the level or distribution of melanin, the primary pigment determining skin color, in facial or skin images before feature extraction or classification. This approach is designed to address bias in facial recognition systems or dermatological diagnosis, whose performance often declines for individuals with high melanin levels (darker skin tones). By incorporating melanin awareness, the system produces more accurate and fair outcomes across all skin tones [22].

D. ArcFace

ArcFace introduces an additive angular margin loss to enhance the discriminative power of the learned features. This approach improves the model’s ability to distinguish between different classes by enlarging the angular margin among them. ArcFace has been shown to perform strongly across multiple benchmarks, demonstrating robustness and high accuracy in facial recognition tasks [21].

E. Evaluation Methods

The evaluation metrics include accuracy and the relationship between False Acceptance Rate (FAR) and True Acceptance Rate (TAR), which are particularly relevant for applications with imbalanced class distributions. Accuracy measures how often a system produces correct predictions compared to the total number of predictions made. FAR occurs when the system incorrectly accepts an unauthorized individual (an impostor) due to excessively high similarity scores. In contrast, TAR represents the percentage of genuine users who are correctly accepted by the system.

IV. RESULTS

This study employed several libraries and frameworks to support the development and evaluation process. The computer vision components utilized opencv-python (v4.8.0) and opencv-contrib-python (v4.8.0), while insightface (v0.7.3) and ONNX Runtime (v1.16.0) were integrated for model deployment. Core numerical and scientific computations were handled using NumPy (v2.0) and SciPy (v1.11.0). The database management relied on cx_Oracle (v8.3.0). For deep learning tasks, the study implemented PyTorch (v2.0.0) together with Torchvision (v0.15.0).

Data processing was conducted using Pillow (v10.0.0), scikit-learn (v1.3.0), and pandas (v2.0.0). Visualization was supported by Matplotlib (v3.7.0), Seaborn (v0.12.0), and tqdm (v4.65.0) for progress monitoring. The Application Programming Interface (API) was built using Flask (v3.0.0), while system utilization and performance tracking were facilitated through TensorBoard (v2.14.0).

The dataset consisted of two parts: a training set comprising 1,000 facial images of dark-skinned individuals, and a testing set comprising 100 facial images of dark-skinned individuals.

A. Accuracy

Figure 2 compares the performance of the melanin-aware method with the condition without preprocessing, based on similarity scores across 100 test samples. Overall, the melanin-aware approach demonstrates more stable and consistent performance, with similarity values ranging approximately between 0.92–0.98. The relatively smooth line pattern indicates that this method can maintain facial matching quality despite variations in lighting conditions or differences in skin tone.

In contrast, the condition without preprocessing produces lower similarity values, generally within the range of 0.85–0.90 and exhibits greater fluctuations throughout the testing process. This suggests that without normalization or melanin-specific adjustments, the system becomes more sensitive to visual noise and image quality differences, thereby reducing matching accuracy.

Accordingly, the diagram underscores that melanin-aware preprocessing enhances the stability and reliability of facial verification systems compared to approaches that omit preprocessing altogether. From the comparison results, the experimental data with melanin-aware preprocessing achieved an average accuracy improvement of 17% compared to the experimental data without preprocessing.

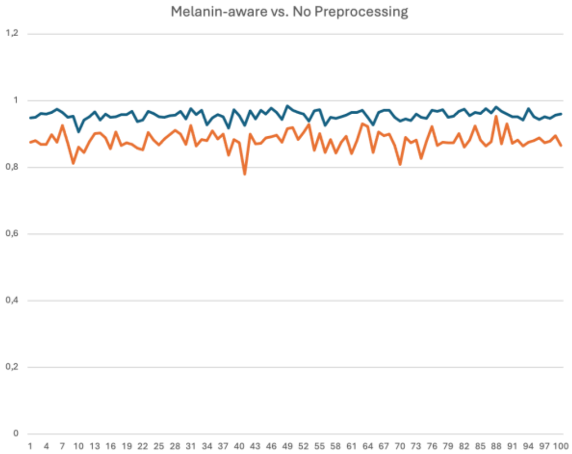

Figure 2. Comparison of Accuracy Between Melanin-Aware and Non-Preprocessing Approaches

The diagram below illustrates a comparison of similarity scores between the melanin-aware method and the standard approach (aggressive CLAHE) across 100 test samples. The visualization shows that the melanin-aware method produces more stable and generally higher similarity values compared to the standard method. The blue line, representing melanin-aware preprocessing, fluctuates within the range of approximately 0.94–0.98, with consistently high peaks in the upper similarity region. In contrast, the orange line, representing the standard method, exhibits greater and less stable fluctuations, with similarity values often dropping to 0.90–0.92, and showing more extreme variability across samples.

This pattern indicates that the melanin-aware approach reduces sensitivity to variations in lighting conditions, contrast, and skin characteristics, thereby yielding more consistent verification performance. Overall, the diagram confirms that melanin-aware preprocessing provides more reliable facial matching quality compared to standard preprocessing techniques such as aggressive CLAHE. From the comparison results, the experimental data with melanin-aware preprocessing achieved an average accuracy improvement of 7% over the experimental data processed with aggressive CLAHE.

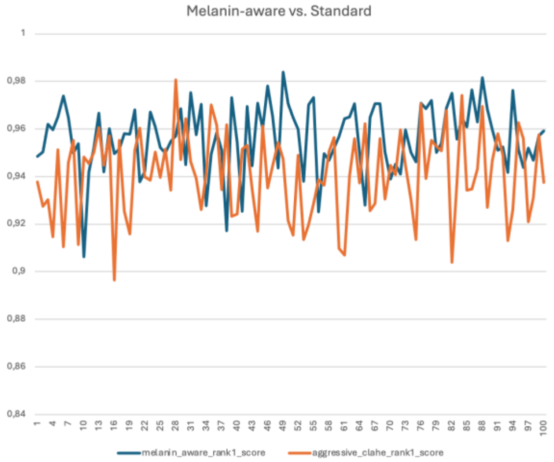

Figure 3. Comparison of Accuracy Between Melanin-Aware and Standard Preprocessing

B. *True Acceptance Rate* (TAR)

The evaluation results indicate that the use of preprocessing has a significant impact on the performance of facial verification systems. In the True Acceptance Rate (TAR) evaluation at False Acceptance Rate (FAR) = 1e−2 (1%), the TAR value increased from 88.15% to 93.33%, representing a substantial performance improvement. At this relatively relaxed FAR level, the system becomes considerably more capable of correctly recognizing genuine face pairs. In addition, the threshold value at the same FAR also decreased when preprocessing was applied, suggesting that the separation between genuine and impostor scores became clearer—both due to higher genuine scores and lower impostor scores.

Under stricter conditions, namely FAR = 1e−3 (0.1%), performance improvements were still observed, although not as pronounced as before. The TAR value rose from 83.7% to 86.67%, demonstrating that preprocessing continues to enhance robustness under evaluation settings that are more sensitive to errors. The threshold value at FAR = 1e−3 also decreased slightly, consistent with the previous pattern, indicating that score distributions became more distinct and that preprocessing contributed positively to the precision of the system.

Table 2. Calculation of True Acception Rate (TAR)

Metric	*Standard*	*Preprocessed M-A*
n	137	137
TAR@FAR=1e-02	0,8815	0,9333
th (1e-02)	0,5373	0,5185
TAR@FAR=1e-03	0,837	0,8667
th (1e-03)	0,5549	0,5443

C. Discussions

The evaluation results of the facial verification system demonstrate that the application of melanin-aware preprocessing consistently improves performance compared to both standard methods and the absence of preprocessing. This is evident from the more stable distribution of similarity scores and the significant increase in TAR across various FAR levels. At FAR = 1e−2 (1%), TAR rose from 88.15% to 93.33%, indicating that the system became more capable of correctly recognizing genuine users even under relatively lenient error tolerance. This improvement suggests that preprocessing enhances facial feature representation, thereby enabling the system to more effectively distinguish between genuine and impostor pairs.

In addition, the threshold required to achieve a given FAR decreased after preprocessing was applied. This reduction indicates that the distributions of genuine and impostor scores became more distinct: genuine scores increased while impostor scores decreased. Such clearer separation strengthens model stability by reducing overlap between genuine and impostor distributions, thereby facilitating the determination of an optimal decision boundary.

When evaluation was conducted under stricter security conditions, namely FAR = 1e−3 (0.1%), performance improvements were still observed, though less pronounced than at the more lenient FAR. TAR increased from 83.7% to 86.67%, showing that melanin-aware preprocessing not only enhances accuracy but also improves model robustness in verification scenarios that are more sensitive to errors. The consistent reduction in threshold further reinforces the finding that preprocessing improves the overall quality of score separation.

Visually, comparative graphs between melanin-aware preprocessing and other standard methods reveal that melanin-aware produces more stable score patterns with peaks consistently located in the high similarity range. In contrast, standard methods exhibit greater fluctuations and lower similarity values. These findings support the interpretation that facial features influenced by melanin variation can be more effectively addressed through preprocessing that explicitly accounts for such characteristics.

V. CONCLUSIONS

The objective of this study is to develop a facial recognition method capable of accurately identifying dark-skinned individuals while maintaining performance for non–dark-skinned individuals. Based on the evaluation and analysis, the following conclusions can be drawn:

1. Melanin-aware preprocessing significantly enhances facial recognition performance for dark-skinned individuals while preserving accuracy for non–dark-skinned individuals.
2. The TAR value increased substantially, particularly at FAR = 1e−2, indicating the system's improved ability to correctly accept genuine users under relatively lenient error tolerance.
3. The reduction in threshold values demonstrates that the distributions of genuine and impostor scores became more distinct, thereby improving the clarity of decision-making.

In light of these findings, the melanin-aware preprocessing method is combined with ArcFace.

ACKNOWLEDGEMENT

This research is partially funded by Program Pendanaan Riset dan Inovasi untuk Indonesia Maju (RIIM) under grant Badan Riset dan Inovasi Nasional (BRIN) with funding amount IDR 170.000.000.

REFERENCES

[1] M. Georgopoulos, J. Oldfield, Y. Nicolaou Mihalis A and Panagakis, and M. Pantic, "Mitigating Demographic Bias in Facial Datasets with Style-based Multi-Attribute Transfer," *Int. J. Comput. Vis.*, vol. 129, no. 7, pp. 2288–2307, Jul. 2021.

[2] R. Singh, A. Agarwal, M. Singh, S. Nagpal, and M. Vatsa, "On the robustness of face recognition algorithms against attacks and bias," *Proc. Conf. AAAI Artif. Intell.*, vol. 34, no. 09, pp. 13583–13589, Apr. 2020.

[3] A. Birhane, "The Unseen Black Faces of AI Algorithms," *Nature*, vol. 610, no. 7932, pp. 451–452, Oct. 2022.

[4] K. Karkkainen and J. Joo, "FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation," *Proceedings - 2021 IEEE Winter Conference on Applications of Computer Vision, WACV 2021*, pp. 1547–1557, Jan. 2021, doi: https://doi.org/10.1109/WACV48630.2021.00159.

[5] T. Sixta, J. C. S. Jacques Junior, P. Buch-Cardona, E. Vazquez, and S. Escalera, "FairFace challenge at ECCV 2020: Analyzing bias in face recognition," in *Lecture*

Notes in Computer Science, in Lecture notes in computer science. , Cham: Springer International Publishing, 2020, pp. 463–481.

[6] E. Villalobos, D. Mery, and K. Bowyer, "Fair Face Verification by Using Non-Sensitive Soft-Biometric Attributes," *IEEE Access*, vol. 10, pp. 30168–30179, 2022.

[7] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," *IEEE Trans Affect Comput*, vol. 13, no. 3, pp. 1195–1215, 2022, doi: https://doi.org/10.1109/TAFFC.2020.2981446.

[8] S. Li and W. Deng, "A Deeper Look at Facial Expression Dataset Bias," *IEEE Trans. Affect. Comput.*, vol. 13, no. 2, pp. 881–893, Apr. 2022.

[9] J. P. Robinson, G. Livitz, Y. Henon, C. Qin, Y. Fu, and S. Timoner, "Face Recognition: Too Bias, or Not Too Bias?," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, Jun. 2020.

[10] M. Pot and B. Prainsack, "Reply to Letter to the Editor on ``Not all biases are bad: equitable and inequitable biases in machine learning and radiology''," *Insights Imaging*, vol. 12, no. 1, p. 157, Nov. 2021.

[11] W. Sun, O. Nasraoui, and P. Shafto, "Evolution and impact of bias in human and machine learning algorithm interaction," *PLoS One*, vol. 15, no. 8, p. e0235502, Aug. 2020.

[12] I. Serna, A. Morales, J. Fierrez, and N. Obradovich, "Sensitive Loss: Improving Accuracy and Fairness of Face Representations with Discrimination-Aware Deep Learning," *Artif. Intell.*, vol. 305, no. 103682, p. 103682, Apr. 2022.

[13] S. Gong, X. Liu, and A. K. Jain, "Mitigating Face Recognition Bias via Group Adaptive Classifier," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2021.

[14] G. Balakrishnan, Y. Xiong, W. Xia, and P. Perona, "Towards Causal Benchmarking of Bias in Face Analysis Algorithms," in *Lecture Notes in Computer Science*, in Lecture notes in computer science. , Cham: Springer International Publishing, 2020, pp. 547–563.

[15] M. Gwilliam, S. Hegde, L. Tinubu, and A. Hanson, "Rethinking Common Assumptions to Mitigate Racial Bias in Face Recognition Datasets," in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, IEEE, Oct. 2021.

[16] D. Raz *et al.*, "Face Mis-ID: An Interactive Pedagogical Tool Demonstrating Disparate Accuracy Rates in Facial Recognition," in *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, New York, NY, USA: ACM, Jul. 2021.

[17] E. Röösli, S. Bozkurt, and T. Hernandez-Boussard, "Peeking into a black box, the fairness and generalizability of a MIMIC-III benchmarking model," *Sci. Data*, vol. 9, no. 1, p. 24, Jan. 2022.

[18] I. D. Raji, T. Gebru, J. Mitchell Margaret and Buolamwini, J. Lee, and E. Denton, "Saving Face," in *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, New York, NY, USA: ACM, Feb. 2020.

[19] J. G. Cavazos, G. Jeckeln, and A. J. O'Toole, "Collaboration to improve cross-race face identification: Wisdom of the multi-racial crowd?," *Br. J. Psychol.*, vol. 114, no. 4, pp. 838–853, Nov. 2023.

[20] L. Hinzman, E. P. Lloyd, and K. B. Maddox, "The stigmatized perceiver: Exploring the implications of social stigma for cross-race face processing and memory," *Soc. Personal. Psychol. Compass*, vol. 16, no. 2, Feb. 2022.

[21] A. Khalifa and A. Al-Hamadi, "JAMsFace: joint adaptive margins loss for deep face recognition," *Neural Comput. Appl.*, vol. 35, no. 26, pp. 19025–19037, Sep. 2023.

[22] Z. Wei and T. Chakraborti, “Achieving Fair Skin Lesion Detection through Skin Tone Normalization and Channel Pruning,” Sep. 2025, [Online]. Available: http://arxiv.org/abs/2509.22712