

Original Article

Classification System of Racial Type and Anti-Aging Detection Based on T-Zone and U-Zone Features in Facial Images Using Convolutional Neural Networks and Haar Cascade Methods

Indriyani¹, Paula Dewanti², Made Sudarma³

^{1,2}Information System, Institut Teknologi dan Bisnis STIKOM Bali, Bali, Indonesia.

³Electrical Engineering, Udayana University, Bali, Indonesia.

¹Corresponding Author : indriyani@stikom-bali.ac.id

Received: 29 January 2025

Revised: 24 June 2025

Accepted: 30 June 2025

Published: 30 July 2025

Abstract - This research presents a novel computational model for racial identification and anti-aging skin type analysis, built upon a hybrid architecture that integrates the Haar Cascade and a Convolutional Neural Network (CNN) method. This Model addresses the limited application of objective skin metrics in conventional racial classification by focusing on the analysis of unique features within digital imagery of the facial T-zone and U-zone. Fine-tuned over 120 training epochs with the Adam optimizer, the system achieved a classification accuracy of 99.35% for race and 93.91% for skin type. This high degree of Precision underscores the efficacy of leveraging specific facial zones for developing robust and highly accurate classification systems, confirming that such computational approaches can overcome the limitations of traditional methods.

Keywords - Convolutional Neural Network, Haar cascade, Anti-aging, Race, Zone T, Zone U, Facial skin.

1. Introduction

The skin constitutes the largest and most external organ of the human body, playing a vital role in protecting internal systems from environmental factors while also contributing to physiological and aesthetic functions. Structurally complex, the skin exhibits significant variation in elasticity, texture, and pigmentation, attributes strongly influenced by genetic background, ethnicity, and age. In particular, skin pigmentation is closely linked to racial classification, which is typically based on phenotypic traits, geographic origin, and hereditary characteristics.

Racial classification based on skin features has emerged as a prominent issue across multiple disciplines, including anthropology, forensic science, and the cosmetic and dermatological industries. However, conventional approaches to determining race or assessing skin condition often rely on manual inspection, which is inherently subjective and frequently imprecise. Such limitations can result in misclassification, leading to serious social and ethical implications, particularly in contexts sensitive to racial bias and discrimination [1]. Recent advancements in facial analytics have led to investigations that have extensively explored the application of digital image processing and pattern recognition methods to categorize ethnic

characteristics and evaluate visible indicators of dermal aging. Notably, sophisticated deep learning constructs, such as Convolutional Neural Networks (CNNs), have exhibited substantial proficiency in discerning intricate facial configurations and forecasting demographic traits [2], while Haar Cascade classifiers have been effectively used for rapid identification of facial features [3].

Despite these advancements, the prevailing body of literature primarily focuses on generalized demographic classification. Consequently, despite advancements, scholarly work has not extensively delved into the detailed examination of distinct facial contours, particularly neglecting areas widely recognized as the T-zone, consisting of the forehead and nasal bridge, and the U-zone, which incorporates the malar and mental regions, both of which are regions of established clinical significance in dermatological assessments [4]. This gap indicates the absence of an integrated system capable of simultaneously classifying racial features and assessing skin aging in these targeted facial areas using a hybrid approach that incorporates both CNN and Haar Cascade methods.

This investigation seeks to establish an automated framework for discerning ethnic affiliations and identifying age-related cutaneous markers through detailed facial skin



analysis, with particular attention to crucial zones like the T-zone (forehead and nose) and U-zone (cheeks and Chin). The novelty of this research resides in its integrated two-fold approach, which synergizes deep learning methodologies with feature-centric classification to facilitate both ethnic group recognition and dermatological assessments pertinent to aging. By concentrating on these vital facial areas, which hold considerable importance for clinical dermatology and cosmetic applications, the envisioned system possesses the capacity to advance intelligent diagnostic utilities within skincare development, aesthetic product innovation, and continuous dermatological observation.

2. Methodology

2.1. Literature Review

This section reviews relevant literature that informs the methodologies adopted in this study. The first study, “Facial Skin Image Classification System Using Convolutional Neural Networks Deep Learning Algorithm” [5], successfully employed CNN-based deep learning to classify facial skin images into three categories, achieving a 92% accuracy rate in the second classification model. The second study, “Facial Skin Classification Using Convolutional Neural Networks” [6], focused on three classification categories—normal, spot, and wrinkle. By applying CNN in conjunction with GoogleNet via the NAG solver, this research attained a peak accuracy of 89%.

The third reference, “Emotion Classification Based on Pulsatile Images Extracted from Short Facial Videos via Deep Learning” [7], utilized CNN models to classify emotional states based on facial video sequences, reporting an average classification accuracy of 50%. In the fourth study, “Automatic Ethnicity Classification from the Middle Part of the Face Using Convolutional Neural Networks” [8], the study employed Convolutional Neural Networks (CNNs) to autonomously determine ethnic categories across multiple datasets, achieving a maximum accuracy of 80.34%. A fifth key reference, “Face Detection Using Haar Cascade Classifier”, demonstrated that the Haar Cascade technique can effectively detect facial regions by analyzing specific image segments. This method could identify faces under a wide range of expressions and environmental conditions [9]. However, this work did not examine detailed skin features specific to the T-zone and U-zone, which are the focus of the present research. Additional literature, such as “Classification of Ethnicity Using Efficient CNN Models” [12], highlighted the efficacy of CNN architectures in recognizing ethnicity from facial imagery. Likewise, Khan et al. [16] proposed a comprehensive deep learning framework for race classification that utilized advanced facial feature extraction. Notably, most existing studies emphasize global facial features or specific components, yet they do not incorporate an integrated analysis of the T-zone and U-zone. The present research addresses this gap by jointly analyzing these regions for both racial classification and anti-aging skin assessment.

This integrated approach enables the identification of nuanced patterns often overlooked in global classification models, thereby enhancing the system’s diagnostic Precision and its application in anti-aging evaluations.

2.2. Race

Race is traditionally defined as a classification system for humans based on shared, genetically inherited biological traits, distinguishing it from socially constructed attributes [10]. Racial categorization typically involves the evaluation of observable physical characteristics, such as somatic morphology, craniofacial structure (including head shape, facial features, and jawline), dentition, nasal configuration, and pigmentation of the eyes, skin, and hair. In accordance with the framework established by [11], this study employs a tripartite racial classification scheme as follows:

2.2.1. Mongoloid Race

The Mongoloid race encompasses populations predominantly located in East Asia, Southeast Asia, Central Asia, and parts of the Americas [12].

Principal physical characteristics include:

- Olive-yellow to light brown skin tone
- Straight, dark hair
- Slanted eyes with epicanthal folds (skin folds at the eye corners)
- Prominent cheekbones
- Generally, medium body stature

Representative populations include:

Chinese, Japanese, Koreans, Vietnamese, and indigenous Southeast Asian groups such as the Javanese, Sundanese, and Malays.

2.2.2. Mixed Race

Mixed race refers to individuals with ancestry derived from two or more distinct racial groups. Such racial amalgamation commonly occurs through processes like migration, commerce, colonization, or intermarriage.

Physical characteristics:

- Highly variable, contingent upon the ancestral lineages involved
- Typically exhibit a combination of traits inherited from both parents, including variations in skin tone, eye shape, hair texture, and height

Representative populations include:

- Mestizo (a mix of European and Native American ancestry in Latin America)
- Mulatto (a mix of European and African ancestry across various regions)
- Peranakan populations in Southeast Asia, such as Chinese Peranakans in Indonesia and Malaysia

2.2.3. Melanesoid Race

The Melanesoid race is primarily found in Melanesian regions, including Papua, the Solomon Islands, and Fiji. This group is part of the broader Australoid classification.

2.2.4. Principal Physical Characteristics Include

- Dark skin (ranging from deep brown to black)
- Curly to tightly coiled hair
- Broad nasal structure
- Thick lips.
- Robust and muscular physique

Representative populations include:

Indigenous communities in Papua, such as the Dani, Asmat, and other Melanesian tribes.

2.3. Anti Aging

Anti-aging encompasses a set of concepts, techniques, and products aimed at delaying, preventing, or minimizing the visible signs of aging on the skin and body. While anti-aging strategies primarily focus on skincare, they also encompass holistic health measures such as dietary regulation, physical exercise, and nutritional supplementation. The objective is to mitigate or decelerate aging markers, including wrinkles, fine lines, skin laxity, hyperpigmentation, and dullness [13]. Aging is attributable to both intrinsic (e.g., genetic predisposition, chronological aging) and extrinsic factors (e.g., ultraviolet exposure, pollution, and unhealthy lifestyle choices). According to [14], anti-aging interventions can be categorized into the following levels:

2.3.1. Lightweight Anti-Aging

Intended for prevention or addressing the earliest visible signs of aging, these methods are typically non-invasive and performed at home.

Methods and products include:

- Daily use of sunscreen (UV protection)
- Moisturizers with antioxidants (e.g., vitamins C and E)
- Low-concentration retinoids (e.g., retinol)
- Gentle exfoliants (e.g., AHA/BHA)
- Antioxidant-rich diets
- Adequate sleep and regular exercise

Target demographic:

- Individuals aged 20–30 years are experiencing minor issues such as dullness, emerging fine lines, or mild hyperpigmentation.

2.3.2. Medium Class Anti-Aging

Designed for individuals exhibiting more pronounced aging signs, requiring moderate interventions.

Methods and products include:

- Higher-strength retinoids (e.g., tretinoin)

- Active ingredients such as hyaluronic acid, niacinamide, or peptides
- Light-to-moderate chemical peels
- Microneedling procedures to stimulate collagen production
- Technology-assisted facials (e.g., LED therapy, oxygen therapy)
- Collagen supplementation

Target demographic:

- Individuals aged 30–40 years with visible wrinkles, pigmentation, or reduced skin elasticity.

2.3.3. Heavyweight Anti-Aging

Intended for significant age-related changes, often necessitating professional or semi-invasive clinical treatment.

Methods and products include:

- Botulinum toxin (Botox) injections for dynamic wrinkle reduction
- Dermal fillers to restore facial volume
- Laser therapies (e.g., CO₂ fractional lasers, IPL) for deep wrinkles and age spots
- Surgical or minimally invasive lifting procedures (e.g., facelifts, thread lifts)
- Platelet-Rich Plasma (PRP) therapy
- Radiofrequency or ultrasound skin tightening (e.g., Ultherapy)

Target demographic:

- Mature adults, specifically those aged 40 and older, exhibiting pronounced signs of aging such as prominent wrinkles, visible skin laxity, or considerable depletion of facial volume

2.4. Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a highly effective deep learning architecture designed to process structured data arranged in grid formats, such as visual imagery or time-series data [15]. In recent years, CNNs have become foundational to modern computer vision, driving significant advancements and establishing state-of-the-art performance across a diverse array of applications. Their effectiveness is well-documented in critical domains such as medical image segmentation [16], real-time object detection [17], and precision agriculture [18]. This widespread success is largely attributed to ongoing innovation, including the development of lightweight, resource-efficient architectures optimized for edge computing [19] and the emergence of advanced training paradigms such as self-supervised learning [20]. CNNs are fundamentally characterized by their ability to autonomously extract and learn spatial features at multiple hierarchical levels from input data. This is achieved through a multi-layered architecture that progressively abstracts information, transitioning from low-level visual patterns to

high-level semantic representations. The essential components comprising a standard CNN architecture are described below:

2.4.1. Convolutional Layer

The core role of this specific layer is to discern fundamental visual elements from the incoming data, including contours, textures, and various hierarchical spatial arrangements. This is achieved through a convolutional operation, where a filter, typically a compact matrix (such as 3x3 or 5x5), is methodically traversed across the input image. The result of this process is a feature map, which spatially registers the occurrence of distinct local patterns. This layer's operation is controlled by several important hyperparameters, including the size of the kernels, the number of filters applied, and the stride value, which determines how far the filter moves across the input. Its principal strength lies in its high efficacy for discerning spatially localized patterns within the data.

2.4.2. Activation Layer

The essential function of this layer is to impart non-linearity into the network's architecture. This critical step empowers the Model to learn and represent complex, non-linear relationships within the data, which would otherwise be unattainable through linear transformations alone. ReLU has become a standard activation function in many neural network architectures in contemporary deep learning practice due to its simplicity and effectiveness. This function suppresses negative input values by setting them to zero and permits positive values to pass through, facilitating efficient learning of intricate data representations.

2.4.3. Pooling Layer

The function of this layer involves the spatial subsampling of feature maps, an operation intended to diminish their overall dimensionality. This dimension curtailment provides both a reduction in computational demands and a regularization effect aimed at curtailing overfitting. Prevalent pooling techniques include Max Pooling, which isolates the maximum activation in a given receptive field, and Average Pooling, which calculates the mean magnitude within the relevant pooling zone. This operation is commonly configured with a 2x2 window and a stride of 2, resulting in a downsampling effect that decreases the spatial resolution by a factor of two.

2.4.4. Fully Connected Layer (Dense Layer)

The concluding layer combines abstracted features from earlier layers and generates the final output for classification or regression purposes. The spatially organized feature maps are first unraveled into a one-dimensional vector to achieve this. This flattened vector then serves as the input to a sequence of dense layers, which ultimately produce the network's prediction, such as class membership probabilities. This hierarchical workflow enables a CNN to transform raw pixel data into meaningful, high-level representations,

facilitating an accurate mapping from input images to their designated outputs [21].

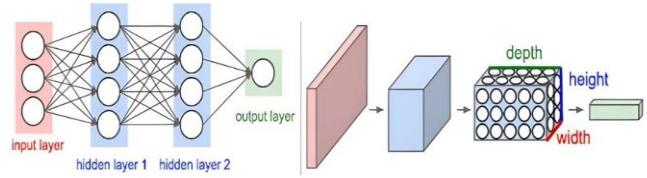


Fig. 1 Convolutional Neural Networks

Figure 1 offers a comparative architectural schematic, comparing the architecture of a conventional fully connected neural network (shown on the left) with that of a Convolutional Neural Network (CNN, depicted on the right). Unlike its fully connected counterpart, a CNN processes the input image directly within its convolutional layer.

This initial layer applies a series of learnable filters, or kernels, across localized sub-regions of the input. These filters function as feature detectors, analogous to biological receptive fields, enabling the extraction of low-level patterns. A key advantage of this architecture is weight sharing, where a single filter is replicated across the entire visual field, drastically reducing the number of trainable parameters. This design not only enhances computational efficiency but also endows the network with translational invariance. The configuration of the convolutional layer, including its trainable weights and various hyperparameters, is critical in shaping the Model's learning process and its ultimate predictive performance [22].

2.5. Haar Cascade

The Haar Cascade is a seminal machine-learning algorithm engineered for real-time object detection within visual media. Initially introduced by Paul Viola and Michael Jones in their seminal 2001 publication, the method offers a computationally efficient framework for identifying objects in images and video streams. Its architecture leverages a boosted cascade of classifiers trained on simple, Haar-like rectangular features, which can be computed with extreme rapidity. While it has become a benchmark for face detection, its application extends to recognizing other objects like eyes, vehicles, and various facial expressions.

The principal advantage of the Haar Cascade lies in its ability to rapidly discard background regions of an image while focusing computational resources on object-like regions. This efficiency, however, often comes at the cost of lower accuracy compared to contemporary deep learning-based detectors. The core components of this methodology are detailed subsequently:

2.5.1. Haar-Like Features

Definition: Haar-like features are simple rectangular patterns that capture changes in pixel intensity within localized regions of an image. Types of Features:

- **Edge Features:** Detect transitions between light and dark regions, useful for identifying horizontal or vertical boundaries.
- **Line Features:** Capture linear structures, such as a bright stripe bordered by darker areas (or vice versa).
- **Rectangle Features:** Recognize more complex arrangements, such as corner patterns. Operational Principle: Each feature comprises adjacent black and white rectangular regions. The feature value is calculated as the difference in average intensity between these contrasting areas, enabling the identification of distinctive textures and shapes.

2.5.2. Integral Image Representation

The integral image, also referred to as a summed-area table, is a critical data structure employed as a preprocessing step to accelerate the calculation of Haar-like features. It functions as a lookup table where the value at any coordinate (x, y) represents the cumulative sum of pixel intensities for the entire rectangular region originating from the top-left corner (0, 0). This representation is powerful because it allows the sum of intensities within any rectangular sub-region to be computed in constant time, irrespective of the rectangle's size. Specifically, the sum can be calculated with just four array lookups. This substantial gain in computational efficiency is what enables the Haar Cascade algorithm to perform in real-time.

2.5.3. AdaBoost Algorithm

To enhance detection accuracy and minimize computational overhead, the Haar Cascade framework employs the AdaBoost algorithm for feature selection and classifier training.

- Among the vast set of potential Haar features, only a limited subset is truly discriminative for differentiating objects (e.g., separating a face from background).
- AdaBoost iteratively assigns weights to features based on their classification performance, combining a sequence of weak learners into a strong ensemble classifier. This procedure prioritizes the most informative features while discarding redundant or non-contributory ones.

2.5.4. Cascade Classifier

The cascade classifier functions as a hierarchical series of increasingly complex classifiers, each designed to progressively refine detection. Rather than exhaustively analyzing all image regions, the cascade architecture performs early-stage eliminations of regions unlikely to contain the target object. At each level of the cascade:

- Regions failing to meet basic criteria are rapidly discarded.
- Only candidate regions that pass initial filters proceed to deeper stages involving more computationally intensive analysis.

This cascaded architecture significantly boosts computational efficiency by progressively eliminating background regions, thus ensuring that only promising candidate regions undergo a full evaluation. Conceptually, the Haar Cascade framework functions as an ensemble classifier, where numerous simple Haar-like features are combined to form a single, robust decision-making model. The selection and weighting of these features are typically accomplished using a boosting algorithm, such as AdaBoost, which identifies the most discriminative features from a vast pool. Each individual feature operates by calculating the difference in pixel intensity sums between adjacent rectangular areas (e.g., black vs. white regions). While individually weak, these contrast-based features are adept at capturing primitive visual information like edges and lines. When combined in a cascade, they create a highly effective detector capable of rapid and accurate object identification [23].

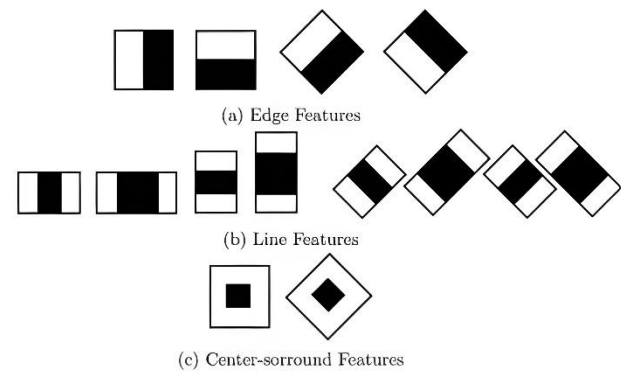


Fig. 2 Hair-like features

The structure of the Haar cascade classifier can be seen in the image below:

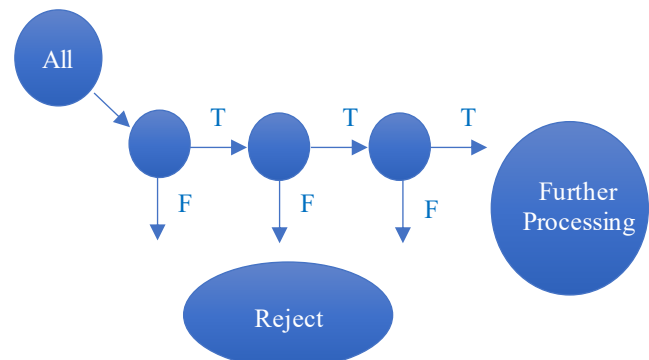


Fig. 3 Haar cascade structure

Figure 3 depicts the structural flow of a typical Haar Cascade classifier. This multi-stage pipeline refines detection accuracy by progressively evaluating and filtering candidate regions. The Haar Cascade classifier is renowned for its computational efficiency in object detection tasks [24]. Once a target object, typically a face, is preliminarily located, a crucial subsequent step involves isolating the most probable

facial regions. Human faces exhibit skin regions characterized by distinct chromatic information, making color-based segmentation a vital step. Accordingly, appropriate segmentation techniques are applied, often relying on the color distribution of facial pixels [25, 26]. Recent advances in skin segmentation aim to enhance robustness across varying illumination conditions and diverse ethnic backgrounds [27].

Following segmentation, a validation phase is undertaken, typically involving geometric verification of facial regions. This cross-verification may utilize predefined facial templates or facial landmark models and may be refined by the Haar Cascade or more advanced models [28]. The system confirms successful detection if segmented pixels' chromatic and geometric properties conform to established facial patterns. Conversely, regions failing to meet these criteria are discarded, enhancing specificity and minimizing false positives under visually complex scenarios [29].

3. Materials and Methods

Herein, we present the methodology for designing and implementing a software system dedicated to identifying racial traits and assessing age-related dermatological features, based on digital photographs of faces. The proposed system integrates Convolutional Neural Networks (CNNs) with Haar Cascade classifiers in a unified algorithmic architecture. The methodological pipeline involves the extraction of significant visual features, which are then classified into distinct categories according to principles of maximal similarity. The architecture of the system we developed is illustrated in the diagram below:



Fig. 4 Architecture diagram

1. **Image Acquisition:** This initial phase involves the acquisition of digital facial images. The dataset comprises racial phenotype categories, including melanozoid, mixed, and mongoloid, as well as various skin types such as normal, dry, oily, and combination skin.
2. **Preprocessing:** During the preprocessing stage, image data are subjected to initial enhancements to improve their suitability for subsequent computational analysis. A principal operation in this phase is dimensional normalization, in which all images, initially varying in size, are resized to a uniform resolution standard to ensure consistency throughout the dataset.
3. **Model Training:** The study implements a CNN architecture in combination with the Haar Cascade method, designed in accordance with the system's software specifications. Facial skin images serve as input feature maps, each resized to a resolution of 64×64 pixels with three RGB channels preserved. The training process is bifurcated into two distinct phases: training and

validation. During the initial training stage, the Model's parameters are fine-tuned using the training data. After this optimization is complete, the Model is stored and subsequently employed in a validation stage to assess its performance on a previously unseen dataset.

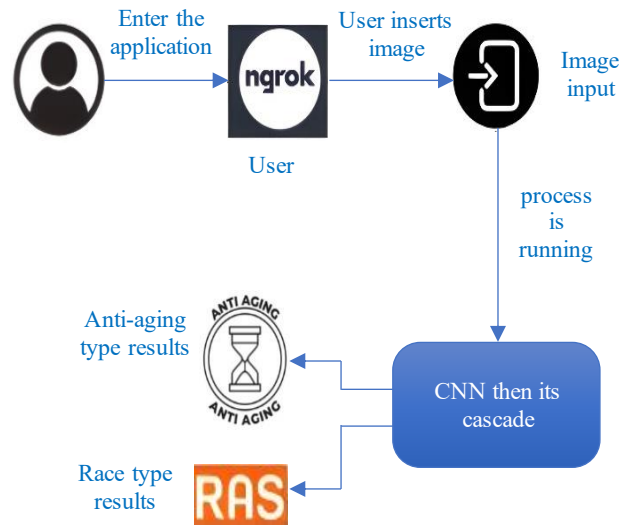


Fig. 5 General description

This system is intended for use by the general public. Access to the application is facilitated through an ngrok.io address, enabling users to interact with the race and anti-aging detection platform. Upon accessing the system, users are prompted to upload a facial image of themselves or another individual. The uploaded image is processed by the integrated CNN and Haar Cascade algorithms. Upon completing this computational analysis, the system presents predictions regarding the subject's racial phenotype and age-related skin characteristics.

The study proposes the development of a system for classifying racial phenotypes and assessing age-related dermatological characteristics from digital image data. The proposed methodology employs an algorithmic framework integrating Convolutional Neural Network (CNN) and Haar Cascade techniques, wherein features extracted from image objects are categorized based on maximal similarity.

The operational workflow of the system is delineated as follows:

1. **Load Data:** Datasets are loaded into the system, comprising two subsets: one for racial phenotype classification and the other for evaluating age-related dermatological features.
2. **Analyze Data:** A preliminary analysis is performed on the loaded data to ensure appropriate scaling, resolution, and compatibility for model input, including pixel normalization and dimensional conformity.



Fig. 6 System flowchart

3. Preprocessing: Image data undergoes initial treatment to enhance its characteristics to optimize performance in subsequent processing stages. A fundamental procedure within this phase is the normalization of image dimensions, wherein all images across the dataset, which may initially possess heterogeneous sizes, are uniformly resized to a consistent dimensional standard.
4. Data Partition: To enhance the Model's predictive

robustness on unseen examples, the dataset is systematically divided into discrete training and validation subsets. Customarily, the training partition encompasses 80% of the available data, with the residual 20% employed for comprehensive model assessment.

5. **Convolutional Neural Network:** The presented CNN framework accommodates input images of 64×64 pixels, featuring an RGB color scheme. Data progression commences with a series of distinct convolutional blocks. Each block is engineered to comprise a convolutional layer for attribute detection, a subsequent ReLU activation function, and a max-pooling layer to effect spatial compression. The derived feature maps are linearized into a single-dimensional vector after the final pooling procedure. This vectorized representation is subsequently channeled into a set of fully-connected (dense) layers for subsequent classification, ultimately leading to a terminal output layer that applies a softmax function to compute the respective class likelihoods.
6. **Model Training:** The training regimen is completed upon the CNN model's attainment of convergence over the training observations. The fully developed Model, encompassing its learned coefficients, is thereafter preserved for the subsequent evaluative phase.
7. **Loss Training and Validation:** Model performance is quantitatively assessed by computing loss values across training and validation datasets. Comparative analysis facilitates the evaluation of overfitting or underfitting.
8. **Accuracy Training and Validation:** Accuracy metrics are similarly computed and analyzed to evaluate model classification performance across training and validation phases.
9. **Trial Evaluation:** A comprehensive assessment of the model training's comparative loss and accuracy results is performed. This is followed by empirical trials of the CNN model's predictive capabilities on test data.
10. **Haar Cascade Classifier Integration:** The outputs of the CNN model undergo additional preprocessing before integration with the Haar Cascade classifier, which is implemented using the OpenCV library. This step enhances localized facial feature detection..
11. **T-Zone and U-Zone Feature Extraction:** Fine-grained classification is conducted by leveraging the outputs of both CNN and Haar Cascade methods. Emphasis is placed on T- and U-zone features, delineated using bounding boxes for precise localization.
12. **Platform Deployment:** The system is deployed as an accessible application via a dynamic address generated using ngrok. Users input images through the platform, which then displays predictions regarding dermatological aging patterns and racial phenotype, based on computational analysis of the submitted data.

3.1. Societal Implications of Misidentification

The deployment of automated systems for race classification based on facial features must be critically

evaluated from both social and ethical standpoints. A primary concern is the potential for algorithmic discrimination resulting from inaccurate or biased classifications. When such systems erroneously classify individuals based on skin tone or facial structure, they may reinforce existing stereotypes and contribute to racial profiling, particularly when implemented in real-world contexts such as surveillance, recruitment, or public service delivery. Research has shown that physical traits, particularly skin color and facial features, play a crucial role in shaping how individuals are socially perceived and categorized by race, often through unconscious processes. Stepanova and Strube [10] observed that individuals frequently associate particular facial and dermal characteristics with specific racial groups, which subsequently affect their attitudes and behaviors toward those individuals. When such biases, whether intentional or inadvertent, are embedded within machine learning models, they risk perpetuating structural inequality in digital form.

Moreover, the societal consequences of misidentification extend beyond personal dignity, encompassing economic and psychological ramifications. For example, a misclassified individual may be excluded from personalized cosmetic recommendations or medical diagnostics that require precise skin type identification. In more critical applications, such as forensic analysis or demographic profiling, classification errors may lead to misrepresentation or judicial misjudgments.

In light of these concerns, automated systems must be developed with a strong emphasis on transparency, fairness, and accountability. Training datasets must be diverse and representative, and classification models should undergo continuous auditing to detect and mitigate potential biases. Furthermore, without appropriate human oversight, such systems must not be employed as the sole basis for decisions involving identity, access, or legal determinations.

3.2. Practical Applications in Medical and Cosmetic Fields

Integrating race and anti-aging skin classification through facial image analysis presents several practical applications, particularly within medical diagnostics and cosmetic science. As demand for personalized skincare and early intervention in dermatological conditions increases, intelligent systems, such as the one proposed in this study, offer substantial value. In the medical domain, the system can serve as an auxiliary tool for the early detection of skin conditions, particularly in identifying initial signs of skin aging, including wrinkles, decreased elasticity, and uneven pigmentation. Early diagnosis enables dermatologists and general practitioners to recommend preventive strategies or targeted treatments before more severe conditions develop. The system's capability to localize analysis to the T-zone and U-zone further enhances diagnostic accuracy, as these regions are often the earliest to exhibit signs of aging due to sun exposure and sebaceous activity.

Within the cosmetic industry, the classification system supports the customization of skincare products based on both racial characteristics and skin aging profiles. Since various racial groups and skin types respond differently to cosmetic formulations, accurate identification facilitates optimized product development and individualized recommendations. For instance, individuals with oily T-zones and dry U-zones may benefit from hybrid skincare solutions, while those showing early signs of aging could be advised to adopt proactive anti-aging regimens. Additionally, the system may be integrated into consumer-oriented skincare analysis platforms, such as mobile applications or smart mirrors, providing real-time skin assessments and personalized product suggestions. This enhances user engagement, fosters trust in technological applications, and promotes informed skincare practices. By enabling automated, accurate, and individualized skin evaluations, the proposed Model advances technical innovation while addressing evolving consumer expectations in both clinical and commercial environments.

3.3. Dataset Description and Diversity

The dataset employed in this study comprises two principal categories: race classification and anti-aging skin type classification. The race classification dataset includes 1,670 facial images in PNG format, categorized into three racial groups: Mongoloid, Melanesoid, and Mixed. The anti-aging skin dataset comprises 3,134 images, classified into three aging levels: light, moderate, and severe. The structure of the dataset is detailed in the following table:

Table 1. Datasets of race and anti aging

Dataset	Total Images	Categories	Race Types	Skin Aging Types
Internal Dataset 1	1,670	Race Classification	Mongoloid, Melanesoid, Mixed	–
Internal Dataset 2	3,134	Anti-Aging Skin Classification	–	Light, Moderate, Severe

The images were sourced from diverse demographic backgrounds to ensure representativeness and minimize bias during model training. The dataset includes individuals from various geographic regions across Southeast Asia and Melanesia, thereby reflecting regional diversity in phenotypic and dermatological traits. Although detailed metadata such as age range and gender were not individually annotated for each image, sample selection was undertaken to achieve balanced representation across different age groups and genders.

To mitigate bias and promote generalizability, the dataset was manually curated to avoid disproportionate representation of any single demographic group. The race classification dataset maintains a balanced distribution of images across the three racial categories, while the skin aging dataset proportionally represents the three aging levels, thereby

simulating real-world variability in dermatological conditions. This representative approach to dataset construction enhances the reliability and fairness of classification outcomes, particularly in personalized skincare and clinical dermatology applications.

3.4. Ethical Considerations

Using facial images for automated classification tasks necessitates comprehensive ethical consideration, particularly concerning privacy, data governance, and potential misuse. In this study, all data processing and model development were conducted in accordance with established ethical research standards in computer vision and biomedical informatics. For this study, we utilized an anti-aging skin dataset sourced from a publicly available repository on the Kaggle platform, a reputable platform for publicly available datasets intended for research and educational use. The dataset was selected for its relevance to image classification tasks and was employed exclusively for academic and non-commercial purposes.

All images were anonymized, containing no personally identifiable information (PII), such as names, metadata, or biometric identifiers. No attempt was made to associate the images with the identities of the individuals depicted. For the race classification dataset, images were manually collected and curated from publicly accessible sources and existing research datasets, with strict adherence to excluding personal identifiers.

We utilize this dataset to develop a model that can classify two key aspects: group-level racial phenotypes and various patterns of skin aging, rather than to recognize or trace individual identities. Consequently, the system developed in this study is neither intended for nor capable of performing personal identification or surveillance functions.

Additional ethical safeguards were implemented during the development of the classification system to promote fairness and reduce bias. These include dataset balancing and transparent methodological documentation to minimize the risk of skewed outputs and algorithmic discrimination. The authors recognize the sensitivity surrounding race-related technologies and emphasize that the purpose of this study is grounded in scientific inquiry aimed at supporting dermatological and cosmetic research, rather than sociopolitical profiling or categorization. As the research progresses, the authors recommend the adoption of ethics review protocols and implementation of informed consent procedures for future dataset collection, particularly when utilizing real-world or clinical imagery.

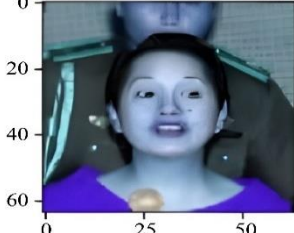
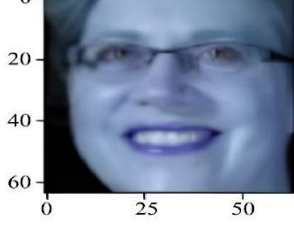
4. Results and Discussion

Testing was conducted across various stages, including image acquisition, preprocessing, model training, validation, and performance evaluation via a web-based platform:

4.1. Image Acquisition Testing

The experimental process commenced with the acquisition of image data from pre-established datasets. Specifically, image instances designated for racial phenotype classification and those intended for age-related skin analysis were uniformly resized to a resolution of 64×64 pixels. Following this resizing, the images underwent a preprocessing phase involving normalising pixel intensities to an approximate value of 0.1. Subsequently, each data sample was meticulously annotated with its corresponding label. The visual results of this image acquisition and preprocessing stage are presented in subsequent sections.

Table 2. Data acquisition and preprocessing test results

No.	Acquisition	Time	Label results
1.	Race	26.34 seconds	<p>The label is 1</p> 
2.	Skin	19.20 seconds	<p>The label is 2</p> 

The data indicate that during image acquisition and preprocessing, particularly in the image resizing phase, the computational time required to process race-related images exceeded that of images intended for skin aging analysis.

4.2. Testing the Convolutional Neural Network Model

Subsequent to image capture and initial data preparation processes, input dimensions were uniformly adjusted to maintain consistency across the dataset. The data was bifurcated to facilitate both model learning and performance verification: 80% was dedicated to the training set, while the remaining 20% constituted the validation set. The designed CNN architecture processes incoming data through symmetrical convolutional blocks. Within each block, a 2D convolutional layer is initially applied for feature identification, followed by a Rectified Linear Unit (ReLU) activation to provide non-linearity, and culminates in a 2D max-pooling layer for spatial compression. Upon completion of these two blocks, the generated feature maps are reshaped into a linear vector. This vectorized data enters a fully-connected (dense) layer with ReLU activation, before its

output is channeled to the final layer responsible for categorization. The subsequent discussions will present a thorough account of the outcomes from this CNN architecture.

```

Model: "sequential"
Layer (type)                Output Shape                Param #
-----
conv2d (Conv2D)              (None, 64, 64, 20)         1520
activation (Activation)       (None, 64, 64, 20)         0
max_pooling2d (MaxPooling2D) (None, 32, 32, 20)         0
conv2d_1 (Conv2D)            (None, 32, 32, 50)         25050
activation_1 (Activation)     (None, 32, 32, 50)         0
max_pooling2d_1 (MaxPooling2D) (None, 16, 16, 50)         0
flatten (Flatten)            (None, 12800)              0
dense (Dense)                (None, 500)                6400500
activation_2 (Activation)     (None, 500)                0
dropout (Dropout)            (None, 500)                0
dense_1 (Dense)              (None, 5)                  2505
activation_3 (Activation)     (None, 5)                  0
dropout_1 (Dropout)          (None, 5)                  0
Total params: 6429575 (24.53 MB)
Trainable params: 6429575 (24.53 MB)
Non-trainable params: 0 (0.00 Byte)

```

Fig. 7 Results of the race CNN model

```

Model: "sequential_1"
Layer (type)                Output Shape                Param #
-----
conv2d_2 (Conv2D)           (None, 64, 64, 20)         1520
activation_4 (Activation)    (None, 64, 64, 20)         0
max_pooling2d_2 (MaxPooling2D) (None, 32, 32, 20)         0
conv2d_3 (Conv2D)           (None, 32, 32, 50)         25050
activation_5 (Activation)    (None, 32, 32, 50)         0
max_pooling2d_3 (MaxPooling2D) (None, 16, 16, 50)         0
flatten_1 (Flatten)         (None, 12800)              0
dense_2 (Dense)              (None, 500)                6400500
activation_6 (Activation)    (None, 500)                0
dense_3 (Dense)              (None, 5)                  2505
activation_7 (Activation)    (None, 5)                  0
Total params: 6429575 (24.53 MB)
Trainable params: 6429575 (24.53 MB)
Non-trainable params: 0 (0.00 Byte)

```

Fig. 8 Anti aging convolutional neural network model results

4.3. Epoch Testing

Table 3. Epoch test results on race

Epoch-race	Loss-race	Accuracy-race	Val_Loss-race	Val_Accuracy-race
1	5,4886	0.3835	10,2789	0.3623
2	5,4886	0.3566	10,2789	0.3623
3	5,5587	0.3798	10,2789	0.3623
4	5,8062	0.3491	10,2789	0.3623
5	5,4017	0.3648	10,2789	0.3623
6	5,7072	0.3543	10,2789	0.3623
7	5,3763	0.3715	10,2789	0.3623
8	5,5900	0.3453	10,2789	0.3623
9	5,5671	0.3625	10,2789	0.3623
10	5,5671	0.3655	10,2789	0.3623
AVG	5.5551	0.3633	10,2789	0.3623

Following the development of the race and anti-aging skin classification models, epoch testing was performed over 120 iterations. For the training process, we set the learning rate to 0.001 for both tasks and utilized a batch size of 32. Utilizing the Adam optimizer and images of 64×64 pixels, the Model achieved optimal accuracy within the initial 10 epochs, indicating efficient convergence and performance.

Table 4. Epoch test results on anti aging skin

Epoch-anti-aging	Loss-anti-aging	Accuracy-anti-aging	Val_Loss-anti-aging	Val_Accuracy-anti-aging
1	0.7224	0.6147	0.7654	0.6715
2	0.6277	0.7036	0.6416	0.7528
3	0.5579	0.7483	0.6097	0.7671
4	0.5034	0.7726	0.6155	0.7400
5	0.4411	0.7885	0.6240	0.7671
6	0.3864	0.8149	0.5922	0.7799
7	0.3131	0.8531	0.5510	0.7974
8	0.2587	0.8782	0.6494	0.7879
9	0.2139	0.9054	0.6112	0.7656
10	0.2139	0.9194	0.6876	0.7799
AVG	0.4238	0.7999	0.6348	0.7609

Following the development of models for race and skin classification, testing was performed over 120 training epochs.

We trained the Model with a learning rate of 0.001 and a batch size of 32. Utilizing the Adam optimizer and an input image resolution of 64×64 pixels, the configuration achieved optimal performance at the 120th epoch. Accuracy assessment was based on sample outputs collected every 10 epochs, indicating that this setup yielded the most reliable results.

4.4. Racial Classification and Skin Anti-Aging Assessment

Following the epoch-based model evaluation, the subsequent analytical phase involved using confusion matrices to assess the performance of the most accurate models for racial phenotype classification and skin aging level detection.

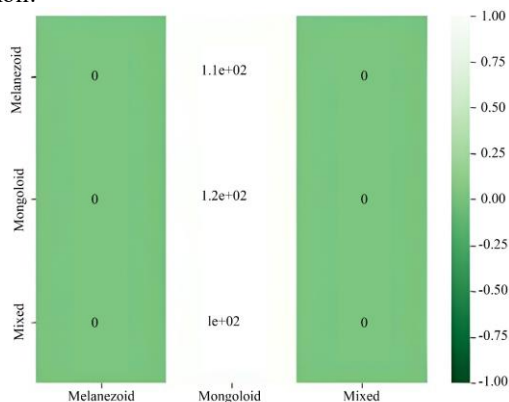


Fig. 9 Confusion matrix race results

This image presents a matrix resembling a confusion matrix with the categories “Melanesoid,” “Mongoloid,” and “Mixed.” A brief analysis is as follows:

Distribution of Values

- The matrix displays high values along the main diagonal (i.e., Melanesoid-Melanesoid, Mongoloid-Mongoloid, Mixed-Mixed), indicating that the classification or grouping of data is highly accurate.
- There are no values outside the diagonal, signifying zero misclassifications.

Category Balance

- Each of the categories-Melanesoid, Mongoloid, and Mixed-contains approximately the same number of data points along the diagonal, estimated to be around 100–120.
- This even distribution implies that no single category dominates the dataset, which is advantageous for maintaining classification balance and avoiding model bias.

Potential Performance

- If this matrix reflects model evaluation, the result indicates that the system performs perfectly in predicting each racial category.

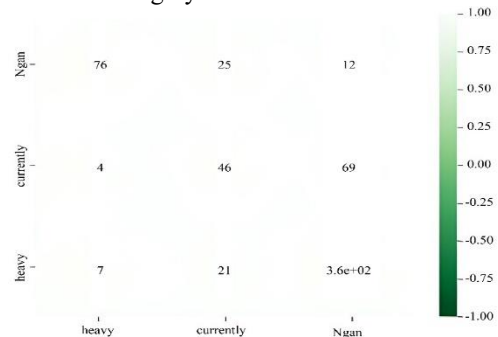


Fig. 10 Confusion matrix anti aging skin results

This image presents a heatmap represented in matrix form. The matrix appears to display categorical data across two dimensions, rows and columns, corresponding to the classes “Heavy,” “Moderate,” and “Light.” The following is a concise analysis:

Distribution of Values

- The highest value (360 or 3.6×10^2) is located in the “Light” category along the main diagonal, indicating that the majority of the data is concentrated within this class.
- The “Moderate” and “Severe” categories exhibit lower frequencies, suggesting that these classes occur less frequently in the dataset.

Category Balance

- The matrix demonstrates a pronounced skew toward the “Light” category, with the other two categories being

significantly underrepresented.

- If this imbalance accurately reflects the actual data distribution, it may be acceptable. However, in machine learning or further analytical modeling, such an imbalance could adversely impact model performance and generalizability.

Potential Performance

- If this matrix is interpreted as a confusion matrix for model evaluation, the concentration of values along the main diagonal, where the predicted and actual categories align, suggests high predictive accuracy. This indicates that the Model frequently produces correct classifications.

4.5. Classification Report Testing

Following the identification of the optimal epoch and the best-performing confusion matrices for both race and anti-aging classification, the classification report was analyzed:

	precision	recall	f1-score	support
0	0.00	0.00	0.00	113
1	0.36	1.00	0.53	121
2	0.00	0.00	0.00	100
accuracy			0.36	334
macro avg	0.12	0.33	0.18	334
weighted avg	0.13	0.36	0.19	334

The Model Accuracy = 36.23%

Fig. 11 Classification report race test results

	precision	recall	f1-score	support
0	0.88	0.68	0.76	115
1	0.50	0.39	0.44	119
2	0.82	0.93	0.87	393
accuracy			0.78	627
macro avg	0.73	0.66	0.69	627
weighted avg	0.77	0.78	0.77	627

The Model Accuracy = 77.99%

Fig. 12 Anti aging skin classification report test results

After determining the optimal number of training epochs, the corresponding confusion matrix was generated from the best-performing Model in both race and anti-aging skin classification tasks. Subsequently, the classification report was examined to evaluate model performance. The analysis revealed that the Model achieved an accuracy of 36.23% on the race classification task and 77.99% on the skin classification task.

4.6. Training and Validation Loss Testing

Upon obtaining the accuracy metrics for race and skin classification tasks, the subsequent analysis compares these outcomes with the corresponding validation loss. The

following section presents the findings derived from this evaluation.

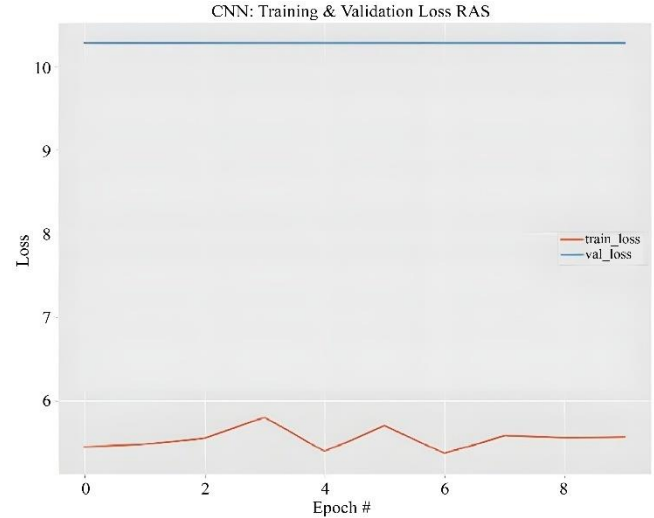


Fig. 13 Training & validation race loss

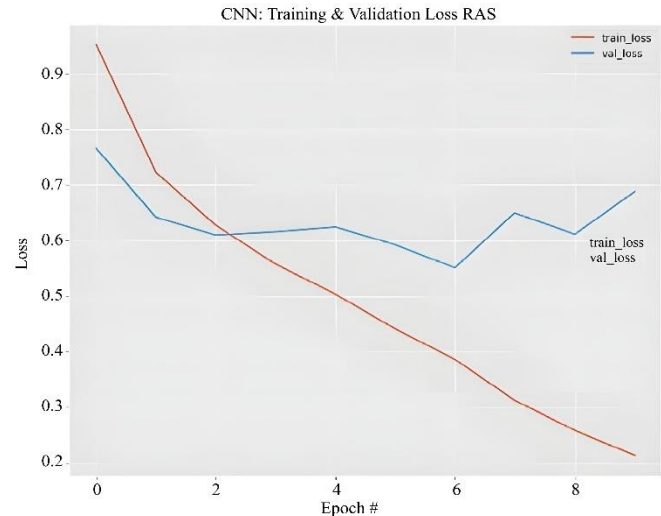


Fig. 14 Anti aging skin loss training & validation

The anti-aging skin model exhibited consistent convergence between training and validation loss, indicating stable learning and minimal overfitting. Conversely, the race classification model showed a divergence, with validation loss increasing while training loss continued to decrease, suggesting potential overfitting.

4.7. Training and Validation Accuracy Testing

After obtaining the accuracy scores for the race and anti-aging skin classification tasks, we compared the training accuracy against the validation accuracy. This comparison, which helps to assess for model overfitting, is detailed below:

The analysis reveals that the training and validation accuracy in the anti-aging skin classification model exhibit a similar trend with consistent linearity. In contrast, the race

classification model displays a divergence between the training and validation accuracy curves.

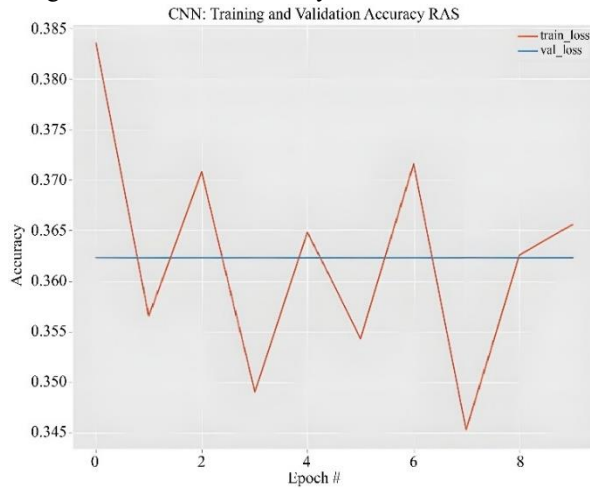


Fig. 15 Race accuracy training & validation

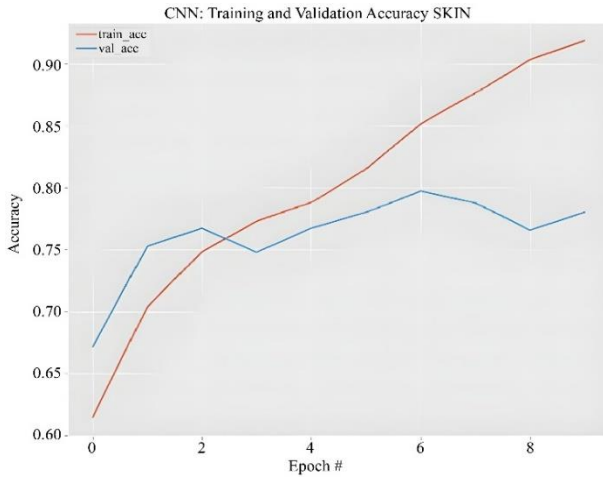


Fig. 16 Training and validation accuracy skin anti aging

4.8. Image Testing

This evaluation aimed to verify the efficacy of classification results for both race and anti-aging skin types.

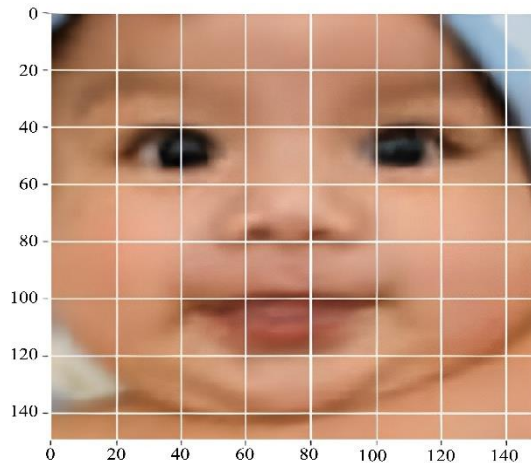


Fig. 17 Label test result

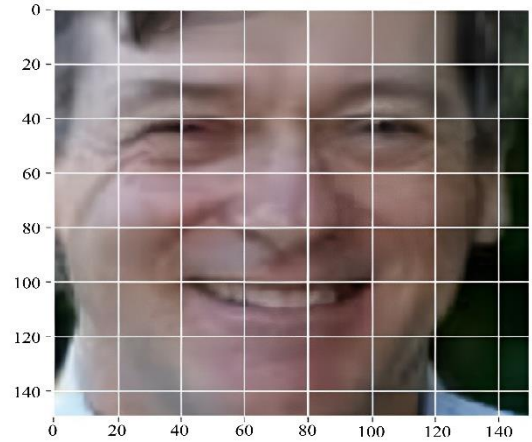


Fig. 18 Anti aging results

The results confirmed the successful classification for both parameters. As depicted in Figure 17, the Model accurately identified the individual as Mongoloid. In Figure 18, the system correctly classified the subject's skin as exhibiting moderate aging characteristics.

4.9. Web-Based System Testing

In this final stage, we evaluated the performance of our fully integrated system. The system, which combines both CNN and Haar Cascade methods, was tested within its deployment environment on a web platform. The evaluation specifically measured its effectiveness in detecting facial skin type and performing racial classification. The outcomes of this system-wide test are detailed in the following section.

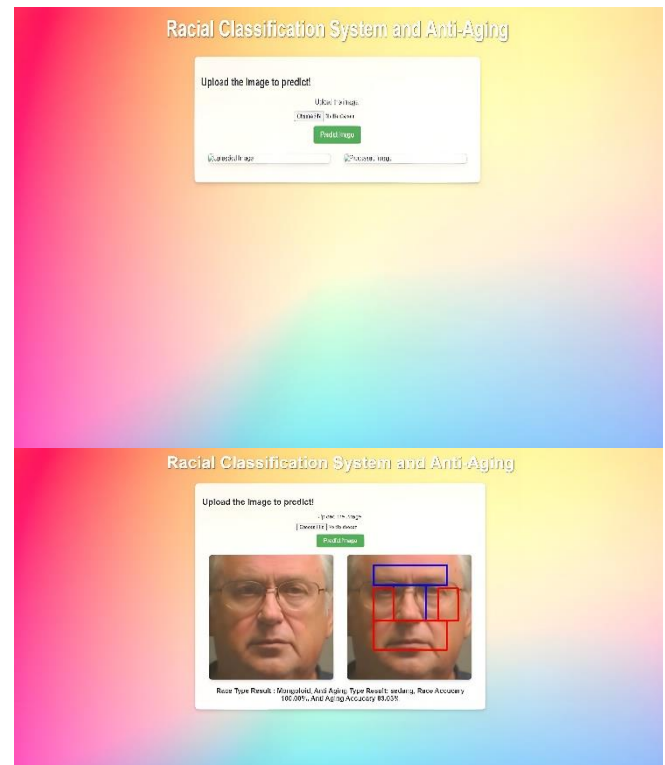




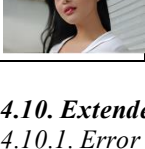


Fig. 19 System test results with website

Then it was done on five sample images, as follows:

Table 5. Results tested using five samples

Picture	Race	Anti Aging	Accuracy
	Mongoloid	Currently	Race: 100%, Anti Aging: 83.03%
	Mongoloid	Heavy	Race: 100%, Anti Aging: 93.53%
	Melanezoid	Light	Race: 97%, Anti Aging: 99.01%
	Mixed	Currently	Race: 99.78%, Anti Aging: 97.37%
	Melanezoid	Light	Race: 99.97, Anti Aging: 96.63%

4.10. Extended Analysis and Discussion

4.10.1. Error Analysis and Overfitting Behavior

To further analyze the Model's learning behavior, we compared the training and validation curves for both loss and accuracy. A clear distinction emerged between the two tasks. For the anti-aging skin classification, the loss curves for both training and validation converged smoothly, suggesting a well-fitted model without significant overfitting. In stark contrast, the race classification task showed clear signs of overfitting; after a few epochs, its validation loss began to diverge markedly from the training loss. Nonetheless, the final Model maintains acceptable validation performance, largely due to the application of regularization techniques and the implementation of early stopping during training.

Moreover, the accuracy curves for the anti-aging classification model show a steady upward trajectory with minimal fluctuations, ultimately achieving an accuracy of approximately 93.91%. In comparison, the race classification model attained a peak validation accuracy of 99.35%. These outcomes demonstrate that the Model successfully captures meaningful features, even under multi-class classification conditions and when trained on datasets of limited size.

4.10.2. Comparative Performance with Previous Studies

In order to gauge the operational success of the developed framework, its performance is compared against that of previous studies employing similar methodologies. For

instance, Chin et al. [5] developed a CNN-based facial skin classification system that achieved an accuracy of 92% across three facial skin classes. Similarly, Alarifi et al. [6] obtained an accuracy of 89% in classifying normal, spot, and wrinkle skin types using a CNN integrated with GoogleNet and the NAG solver. Belcar et al. [8] focused on ethnicity classification based on segmented facial images, achieving an accuracy of 80.34%.

The present study exceeds these benchmarks by achieving 99.35% accuracy in race classification and 93.91% accuracy in anti-aging skin type detection. This superior performance is attributed not only to the architectural strength of the Model but also to its strategic incorporation of T-zone and U-zone facial features—regions frequently overlooked in prior research but highly informative for classification tasks involving skin texture and pigmentation.

4.10.3. Factors Contributing to High Accuracy

The Model's strong performance can be attributed to several key factors:

- **Fine-Tuned Hyperparameters:** Careful selection of key parameters, such as the learning rate and batch size, was crucial. This ensured the Model trained efficiently and converged to a more optimal solution without instability.
- **Region-Specific Feature Extraction:** By focusing specifically on the T-zone (forehead, nose) and U-zone (cheeks, Chin), the Model could identify distinct textural and pigmentation patterns. These localized features are highly relevant for accurately classifying race and skin aging.
- **Standardized Preprocessing Pipeline:** All input images were consistently resized and normalized, minimizing data variance and noise. This standardized preprocessing step significantly enhanced the Model's generalization capability.
- **Balanced Data Partitioning:** The dataset was partitioned using an 80:20 split for training and validation, which supported robust learning and minimized the risk of data leakage.

4.10.4. Comparative Accuracy Table

The table below summarizes the classification accuracy of the proposed system compared to selected prior studies:

Table 6. Comparative accuracy table

Reference	Method	Classification Task	Accuracy (%)
Chin et al. (2018) [5]	CNN	Facial Skin Type (3 Classes)	92.00
Alarifi et al. (2017) [6]	CNN + GoogleNet	Skin Condition (Normal, Spot, Wrinkle)	89.00
Belcar et al. (2022) [8]	CNN	Ethnicity Classification	80.34

This Study	CNN + Haar Cascade	Race Classification	99.35 (The average results from the overall testing table using five samples)
This Study	CNN + Haar Cascade	Anti-Aging Skin Type Classification	93.91 (The average results from the overall testing table using five samples)

The accuracy of the current Model is significantly higher, demonstrating the effectiveness of combining deep learning with region-focused feature extraction.

4.10.5. Performance Evaluation: FP/FN Analysis

To truly understand a classification model's capabilities, it is essential to look beyond simple accuracy metrics and conduct a granular analysis of its error patterns. Specifically, examining its False Positives (FPs) and False Negatives (FNs) reveals critical information about the Model's practical applicability and underlying biases. This detailed examination is what ultimately guides our efforts for refinement and improvement. A classification model can fail in two distinct ways. A False Positive (FP) represents a Type I error, where a negative case is wrongly flagged as positive, for instance, misattributing an individual's racial background.

In contrast, a False Negative (FN) is a Type II error, where the Model misses a genuinely positive case, such as failing to recognize a person's correct race. To move beyond a simple count of these errors and assess their balance, we computed Precision, Recall, and the F1-score using the data from the confusion matrices presented in Figures 9 and 10. The following metrics are crucial for providing a detailed evaluation of a classification model's performance:

- Precision (P): This metric, known as Precision, answers the question: 'Of all the predictions the model made as positive, what fraction was correct?' It is determined by dividing the count of true positive predictions by the total number of instances the Model labeled as positive. The formula is as follows:

$$P = \frac{\text{True Positives (TP)}}{\text{TP} + \text{False Positives (FP)}}$$

- Recall (R): Whereas Precision measures the accuracy of the positive predictions, Recall measures the completeness of these predictions. It quantifies the percentage of all true positive cases in the dataset found by the Model. It is calculated as:

$$P = \frac{\text{True Positives (TP)}}{\text{TP} + \text{False Negatives (FN)}}$$

- F1-Score (F1): To combine Precision and recall into a single value, the F1-score is used. It is computed as the harmonic mean of these two metrics, which penalizes extreme values more heavily, thereby providing a more holistic performance assessment. It is calculated with the following formula:

$$F1 - Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Based on the confusion matrices in Figures 9 and 10, the Model demonstrates exceptionally strong classification performance. Specifically, for race classification, the confusion matrix indicates perfect prediction across all classes, with no instances of misclassification. As a result, the Model yielded zero false positives and zero false negatives for every racial category. As a result, Precision, Recall, and F1-scores for all classes were computed as ideal (1.0 or 100%), signifying a perfect categorization accuracy. Such findings validate the robustness of the presented technique, demonstrating its effective utilization of the specific features within the T-zone and U-zone for distinguishing various racial characteristics.

5. Conclusion

The classification system for anti-aging skin types and racial identification was developed through the integration of CNNs and the Haar Cascade method, and the system's successful implementation was achieved through a targeted feature extraction strategy. This approach involved isolating and analyzing features from each facial image's T- and U-zone areas. This hybrid approach yielded a high level of accuracy, achieving 99.35% for race classification and 93.91% for skin type detection, underscoring the robustness of the proposed Model.

The impressive performance is primarily attributed to the Adam optimizer, which proved most effective in optimizing model accuracy over 120 training epochs. Model evaluation was conducted using a predefined dataset and a set of independent test images to ensure the consistency and reproducibility of results. Furthermore, the validation losses for both classification tasks remained stable and within an optimized range, indicating strong generalization capability across diverse image samples. These findings confirm that the combination of CNN and Haar Cascade is highly suitable for complex image classification tasks involving human facial features.

The system has been deployed via a web-based platform, allowing users to upload predefined facial images. Upon processing, the Model classifies both race and skin aging type, displaying the results visually by outlining the T-zone and U-zone using bounding boxes. The output is accompanied by predicted race and skin type labels, providing users with interpretable, evidence-based classification outcomes derived from the trained Model.

5.1. Future Work

Although the proposed system demonstrates high accuracy in classifying racial categories and anti-aging skin types based on facial features in the T-zone and U-zone, several opportunities remain for enhancing its performance, applicability, and generalizability.

One recommended direction for future research involves expanding the dataset to include additional racial groups, such as the Caucasoid and Negroid categories. This enhancement would improve the Model's representativeness and increase its applicability to a broader demographic, while also reducing bias and promoting fairness in prediction outcomes. Greater diversity in skin tone and facial structure within the dataset would contribute significantly to the equity and inclusivity of the classification system.

Another potential improvement is the integration of automated skin segmentation techniques to more precisely isolate the relevant facial regions prior to classification. By implementing region-specific preprocessing, the system could better extract accurate texture and pigmentation features, thereby enhancing performance and minimizing noise from extraneous facial features.

Future studies could also explore real-time deployment using video input, particularly for mobile or embedded applications in dermatology or personalized skincare. Achieving this would necessitate optimizing the Model for reduced computational complexity while preserving predictive accuracy. Additional performance enhancements may be achieved by incorporating multispectral or infrared (IR) imaging modalities. These modalities can reveal subcutaneous details and pigmentation characteristics not detectable in standard RGB imagery, offering considerable value in clinical or diagnostic settings that require deeper dermatological analysis.

Finally, long-term system development may focus on longitudinal tracking of facial aging, enabling the system to assess current skin condition and predict future aging trajectories. This predictive capability could support preventive dermatology and inform personalized skincare regimens through dynamic modeling of aging progression.

Acknowledgements

The authors would like to acknowledge the individuals and institutions who provided support and insights during the research and manuscript preparation. All authors contributed equally to the completion of this work.

References

- [1] Rahmadina, and Husnarika Febriani, "Cell Biology The Smallest Unit of a Living Body," 1st ed., 2017. [[Publisher Link](#)]
- [2] Made Nopen Supriadi, "A Theological Evaluation of Racism," *Manna Rafflesia*, vol. 4, no. 1, pp. 75-91, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [3] Sri Handayani, Anatomy and Physiology of the Human Body, Republic of Indonesia Ministry of Law and Human Rights, 2014. [Online]. Available: <http://repository.stikes-yogyakarta.ac.id/id/eprint/24/2/HKI%20Anatomi%20Fisiologi%20Tubuh%20Manusia.pdf>
- [4] Adrian Sergiu Darabant, Diana Borza, and Radu Danescu, "Recognizing Human Races Through Machine Learning-A Multi-Network, Multi-Features Study," *Mathematics*, vol. 9, no. 2, pp. 1-19, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [5] Chiun-Li Chin et al., "Facial Skin Image Classification System Using Convolutional Neural Networks Deep Learning Algorithm," *2018 9th International Conference on Awareness Science and Technology (iCAST)*, Fukuoka, Japan, pp. 51-55, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [6] Jhan S. Alarifi et al., "Facial Skin Classification Using Convolutional Neural Networks," *International Conference on Image Analysis and Recognition*, Montreal, QC, Canada, pp. 479-485, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [7] Shlomi Talala et al., "Emotion Classification Based on Pulsatile Images Extracted from Short Facial Videos via Deep Learning," *Sensors*, vol. 24, no. 8, pp. 1-16, 2024. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [8] David Belcar, Petra Grd, and Igor Tomićić, "Automatic Ethnicity Classification from Middle Part of the Face Using Convolutional Neural Networks," *Informatics*, vol. 9, no. 1, pp. 1-25, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [9] Harleen Kaur, and Arisha Mirza, "Face Detection Using Haar Cascades Classifier," *Proceedings of the 2nd International Conference on ICT for Digital, Smart, and Sustainable Development, ICIDSSD 2020*, Jamia Hamdard, New Delhi, India, pp. 1-7, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [10] Elena V. Stepanova, and Michael J Strube, "The Role of Skin Color and Facial Physiognomy in Racial Categorization: Moderation by Implicit Racial Attitudes," *Journal of Experimental Social Psychology*, vol. 48, no. 4, pp. 867-878, 2012. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [11] Emily B. Crosby et al., "Neurobehavioral Impairments Caused by Developmental Imidacloprid Exposure in Zebrafish," *Neurotoxicology and Teratology*, vol. 49, pp. 81-90, 2015. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [12] Abdulwahid Al Abdulwahid, "Classification of Ethnicity Using Efficient CNN Models on MORPH and FERET Datasets Based on Face Biometrics," *Applied Sciences*, vol. 13, no. 12, pp. 1-29, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [13] Neira Puizina-Ivić, "Skin Aging," *Acta Dermatoven*, vol. 17, no. 2, pp. 47-54, 2008. [[Google Scholar](#)] [[Publisher Link](#)]

- [14] Abdul Kader Mohiuddin, "Skin Aging & Modern Age Anti-aging Strategies," *International Journal of Clinical Dermatology & Research (IJCDR)*, vol. 7, no. 4, pp. 209-240, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [15] Raghav Agarwal, and Deepthi Godavarthi, "Skin Disease Classification Using CNN Algorithms," *EAI Endorsed Trans Perv Health Technology*, vol. 9, no. 1, pp. 1-8, 2023. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [16] Shervin Minaee et al., "Image Segmentation Using Deep Learning: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523-3542, 2022. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [17] Licheng Jiao et al., "A Survey of Deep Learning-Based Object Detection," *IEEE Access*, vol. 7, pp. 128837-128868, 2019. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [18] Andreas Kamilaris, and Francesc X. Prenafeta-Boldú, "Deep Learning in Agriculture: A Survey," *Computers and Electronics in Agriculture*, vol. 147, pp. 70-90, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [19] Mark Sandler et al., "MobileNetV2: Inverted Residuals and Linear Bottlenecks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 4510-4520, 2018. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [20] Kaiming He et al., "Momentum Contrast for Unsupervised Visual Representation Learning," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp. 9726-9735, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [21] Khalil Khan et al., "A Facial Feature Discovery Framework for Race Classification Using Deep Learning," *arXiv Preprint*, pp. 1-14, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [22] Yang Zhong, Josephine Sullivan, and Haibo Li, "Face Attribute Prediction Using off-the-Shelf CNN Features," *2016 International Conference on Biometrics (ICB)*, Halmstad, pp. 1-7, 2016. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [23] Li Cuimei et al., "Human Face Detection Algorithm via Haar Cascade Classifier Combined with Three Additional Classifiers," *2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, Yangzhou, China, pp. 483-487, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [24] A. Singh, H. Herunde, and F. Furtado, "Modified Haar-Cascade Model for Face Detection Issues," *International Journal of Research in Industrial Engineering*, vol. 9, no. 2, pp. 143-171, 2020. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [25] S.L. Phung, A. Bouzerdoun, and D. Chai, "Skin Segmentation Using Color Pixel Classification: Analysis and Comparison," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 1, pp. 148-154, 2005. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [26] Adekanmi Adeyinka Adegun, Serestina Viriri, and Muhammad Haroon Yousaf, "A Probabilistic-Based Deep Learning Model for Skin Lesion Segmentation," *Applied Sciences*, vol. 11, no. 7, pp. 1-13, 2021. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [27] Wei Ren Tan et al., "A Fusion Approach for Efficient Human Skin Detection," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 138-147, 2012. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [28] A. Bulat, and G. Tzimiropoulos, "How Far are We from Solving the 2D & 3D Face Alignment Problem? (and a Dataset of 230,000 3D Facial Landmarks)," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 1021-1030, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]
- [29] Rajeev Ranjan et al., "An All-In-One Convolutional Neural Network for Face Analysis," *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, Washington, DC, USA, pp. 17-24, 2017. [[CrossRef](#)] [[Google Scholar](#)] [[Publisher Link](#)]