Facial Skin Type Detection for Race Classification using Convolutional Neural Network and Haar Cascade Method

I Indriyani[1*], Ida Ayu Dwi Giriantari[2], Made Sudarma[3] and I Made Oka Widyantara[4]

[1*]Doctor of Engineering Program, Universitas Udayana, Bali, Indonesia. Institute of Technology and Business (ITB) STIKOM Bali, Bali, Indonesia. Indry.joice@gmail.com, Orcid: https://orcid.org/0000-0002-8371-1240

[2]Department of Electrical Engineering, Universitas Udayana, Bali, Indonesia. Dayu.giriantari@unud.ac.id, Orcid: https://orcid.org/0000-0002-2326-2594

[3]Department of Electrical and Computer System Engineering, Faculty of Engineering, Universitas Udayana, Bali, Indonesia. msudarma@unud.ac.id, Orcid: https://orcid.org/0000-0002-8331-0519

[4]Electrical Engineering Department, Faculty of Engineering, Udayana University, Bali, Indonesia. oka.widyantara@unud.ac.id, Orcid: https://orcid.org/0000-0002-0979-5116

Received: March 07, 2023; Accepted: May 12, 2023; Published: June 30, 2023

Abstract

Every organism possesses a unique structural makeup, extending from molecules to organ systems. One such organ, the skin, is our body's largest and the most complex. Its diversity in color corresponds with various human races, although the facial skin type, an often-overlooked factor, also plays a significant role in race identification. In this research, a system was developed to recognize racial classifications from facial skin types by focusing on features in the facial T and U areas, using the Convolutional Neural Network (CNN) and Haar cascade methodologies. CNN was employed due to its ability to use the convolution process, moving a kernel across an image, multiplying it with the applied filter, and thereby generating new representative information. It is especially effective in image recognition and processing. The Haar cascade method, on the other hand, was used to outline the T and U areas on the face for the skin type detection system. The T area, known for oil detection, identifies skin types, while the U area identifies race types by forming facial patterns. This system, trained on 1670 race and 60 skin type datasets and optimized using the Adam optimizer, exhibited high accuracy levels. Upon testing with five new samples, it demonstrated an average accuracy of 98% in race detection and 97% in skin type detection.

Keywords: Convolutional Neural Network, CNNs, Haar Cascade, T-Region, U-Region.

1 Introduction

Every living being is created with its own structure, such as having molecules, cells, organs, tissues, organisms, and organ systems (G. Bruce, 2019). However, among other living things, humans were created perfectly, because humans are God's creations who were endowed with the ability of reasoning to think and work on this earth. However, just like other living things, humans too have cells, tissues, organs, organ systems, and organisms. An example is the organ called the skin, which is the largest

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 14, number: 2 (June), pp. 41-58. DOI: *10.58346/JOWUA.2023.I2.004*

*Corresponding author: Institute of Technology and Business (ITB) STIKOM Bali, Bali, Indonesia. Doctor of Engineering Program, Universitas Udayana, Bali, Indonesia.

external organ of the human body, one of the most complex and extensive, as well as being flexible and delicate (Bragazzi, N.L., 2019). This skin has many significant roles and tasks in maintaining human life and health.

Speaking of the human skin, there is compelling research that posits that racial conflict and disputes about the meaning of race, as well as disagreements across different racial groups, were forces instrumental in the evolution and transformation of American evangelicalism in the postwar decade (Ballard, T.P., 2019). Essentially, human skin exhibits a broad array of colors commonly associated with various racial groups. This reality, till now, is frequently a source of conflict stemming from racism, which can significantly disrupt the freedom and peace of life of an individual or a group. Concurrently, humans must understand the significance of race in human life. Such comprehension can foster a society that thrives on mutual respect amidst differences.

Race is generally defined as a distinct human population differentiated from others by the frequency of certain genes. The classification of race is often based on physical features or behavioral traits used to categorize humans into different population groups. This classification can describe populations from a wide range of perspectives, including but not limited to phenotype, geography, physical characteristics, and ethnic ancestry inherited (Hochman, A., 2019).

However, many people still resort to exclusively using skin color to distinguish between races. This approach often lacks accuracy, given the multifaceted nature of race and the multitude of types it can be classified into. Fundamental characteristics of a race can typically be identified by skin color, height, head shape, facial features, hair type, eye shape, nose shape, body structure, etc. Facial skin type, although often overlooked, can also play a key role in determining human race.

In light of this, the author plans to develop a system that can detect racial types based on the skin types in two specific facial areas, namely, T and U. The T area typically comprises the forehead and the bridge of the nose, resembling a 'T' shape, and is often the oilier part of the face. The U area, on the other hand, typically includes the cheeks and chin, outlining a 'U' shape on the face. These regions can provide information pivotal for the accurate identification of a person's racial type.

The author intends to leverage the capabilities of Convolutional Neural Networks (CNN) (Guo, G., 2020) in this study. CNNs are a class of deep neural networks, applied mostly to analyzing visual imagery. They are designed to learn spatial hierarchies of features automatically and adaptively through a process known as convolution. Moving a convolution kernel of a certain size over an image enables the system to extract valuable and representative information (Rosda, R., 2023) (Zalmi, W.F., 2023) from the product of the image part and the filter applied (Vo, T., 2018).

To augment the performance of the system, the author also plans to integrate the haar cascade method, which forms classifiers based on the value results derived from haar features. This integration is expected to yield more accurate results (Kharat, D., 2019). As such, the use of feature detection in the T and U areas of the face should lead to more precise results in identifying patterns of facial type and, consequently, racial type. This approach contributes significantly to the ongoing dialogue and research concerning racial classification and identification.

2 Methodology

The following is a block diagram of research that has been carried out on racial and facial skin type detection systems:

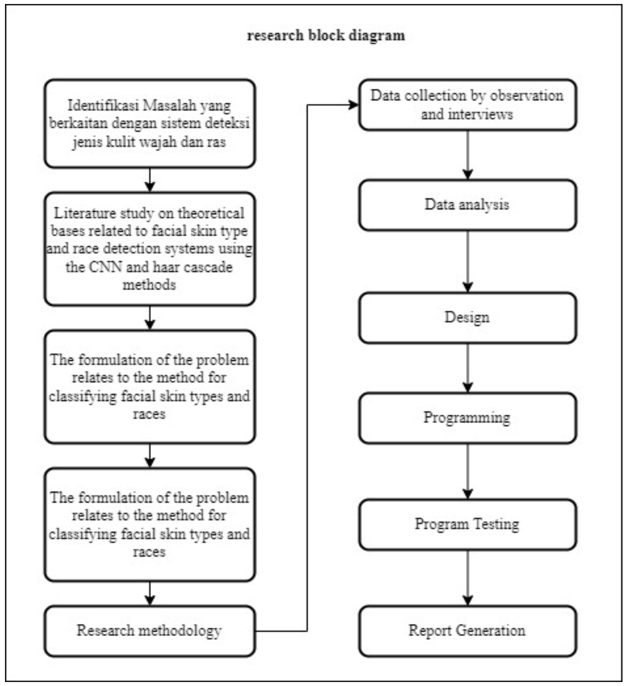


Figure 1: Research Block Diagram

Literature Review

A review of the literature concerning the method used by the author is described below:

The first study (Chan, H.T., 2021) comprised the results of a computer vision technology system that can analyze faces by classifying skin types and detecting acne, which is then used as a recommendation for skin care products (Chan, H.T., 2021). The author succeeded in using computer vision technology to classify facial skin types but did not use the basic features of the T and U areas to detect them. The second item of literature (Vo, T., 2018) constituted experimental results showing that for the VN Faces dataset, the RR-VGG model with augmented input images produced the best accuracy of 88.87%, while the RR-CNN achieved 88.64% (Vo, T., 2018). The author succeeded in using the convolutional neural network method to classify a racial type from a set of datasets but did not use the basic features of the T and U areas for detecting it. The third item of research (Zhang, K., 2017) developed two new strategies to improve the performance of the CNN cascade for face detection (Zhang, K., 2017), succeeding in using the CNN method to classify and detect faces, but did not use the basic features of the T and U areas in such detection. The fourth study (Kharat, D., 2019) concluded that the application of the Haar Cascade method can detect faces based on the image area, i.e., with various facial expressions, under different environmental conditions (Kharat, D., 2019). The author, while using the Haar Cascade method to classify faces in different environments, did not use the basic features of the T and U areas in detecting them. The fifth item of literature (Lin, H.Y., 2021) involves CNN results with Taguchi parametric optimization for detecting facial skin conditions. The original CNN model achieved 79.71% accuracy and average accuracy, and the CNN with Taguchi parametric optimization was 86.95% (Lin, H.Y., 2021). The study managed to show that the proposed CNN optimization method can effectively improve the detection accuracy of facial skin conditions but did not use the basic features of the T and U areas in detecting it. The sixth study (Li, X., 2020) successfully built a new face detection model-the RFE-MTCNN. According to the unique cascade character characteristics of MTCNN, two different receptive field enhancement modules are used to optimize the network structure, and the AM-Softmax loss function is introduced to improve network discrimination (Li, X., 2020). The study showed that the use of the CNN method can improve the accuracy of face detection, but did not use the basic features of the T and U areas in detecting it. The seventh item of literature (Koshy, R., 2019) revealed that CNN can provide 100% accurate results, indicating that it can efficiently classify two-dimensional diffusion images as real or fake (Koshy, R., 2019). This study succeeded in using CNN to outperform advanced approaches used for facial activity detection but did not use the basic features of the T and U areas in detecting it. The eighth study (Yu, B., 2018) concluded that anchor-based face detectors and CNN-based cascade face detectors can improve the accuracy of face detection (Yu, B., 2018) but did not use the basic features of the T and U areas in detecting it. The ninth item of research (Guo, G., 2020), by proposing a fast face detection method based on the DCF extracted by an elaborately designed CNN (Guo, G., 2020). Though the study confirmed that the CNN method for face detection can achieve promising performance on several face detection datasets, it did not use the basic features of the T and U areas in detecting it. the tenth and final study (Zheng, Q., 2022) proposed a deep learning-based approach to the skin feature detection task (Zheng, Q., 2022), using computer vision to produce promising results for the detection of acne, pigmentation, and wrinkles on the face, but did not use the basic features of the T and U areas in detecting them.

Skin

As is known, the skin is the largest and outermost organ and covers the entire surface of the body (Juwanda, F.S., 2021). The main function of the skin is to protect the body from injury and pathogens. It also regulates body temperature, and even controls fluid loss that is not felt, so that it can store vitamin D, water, and fat. The skin is considered a vital organ because it is the center for protecting the health of living beings. This skin has very complex, elastic, and sensitive properties, varying with climate, age, gender, and race. Physically, the skin has a variety of colors, ranging from light, and black, to pink on the soles of the feet and hands. Facial skin is one of the organs of the human body that has the ability to protect the bones and muscles in the face (Hemasree, V., 2022). The following are the types of facial skin (Zhang, K., 2017):

a. Normal skin
This is a skin condition with a balance between the amount of water and oil. Therefore, the face is neither too dry nor too oily.
b. Oily skin
This is generally the easiest to detect because most people with oily facial skin tend to have a smoother facial appearance, large facial pores, and blackheads.
c. Dry skin
This type usually occurs due to low moisture levels in the outermost layer of the skin. This skin type will look like normal but dry skin, and cracked skin, which looks pale, flaky, dull, rough, and lacks elasticity.
d. Combination skin
This is a combination of oily skin in the T-zone (chin, nose, and forehead area) and dry skin in the cheek area.

Race

Race is a classification of humans with identical biological characteristics and is not based on socially structured characteristics. Race is defined as grouping people based on certain characteristics. Usually, the characteristics of this race are recognized from physical signs, such as body shape, head shape, facial expression and lower jawbone, protruding cheekbones, tooth shape, nose shape, eye color, skin color, and hair color. These are divided into 3 types, namely the Mongoloid Race, Melanoid Race, and Mixed Race (Malhotra, K.C., 2019):

a. Melanezoid race is a group of people who look black.
b. Mongoloid race has inherent characteristics like yellow skin, such as American Mongoloid, Malayan Mongoloid, and Asiatic Mongoloid.
c. Mixed Race is a race that has organisms with parents or parents who come from different varieties or populations.

Computer Vision

Computer vision is a cutting-edge technology rooted in the realm of artificial intelligence (AI) that enables computers to gain a comprehensive understanding of the visual world. Essentially, it trains computers to interpret, process, and draw insights from digital images, videos, or other forms of visual data. Based on these insights, the system can then generate recommendations for subsequent actions (Lin, T.Y., 2022).

In this context, computer vision propels the field of deep learning, a subfield of AI that imitates the workings of the human brain in processing data for decision-making. For this study, we employ a specific method in deep learning known as Convolutional Neural Networks (CNN). Computer vision serves the following multiple essential functions:

a. Enhanced clarity of information: Computer vision algorithms are capable of extracting intricate details from images that might go unnoticed by the human eye. This ability substantially reduces the likelihood of errors caused by incomplete or incorrect information.
b. Image quality improvement: Advanced image processing techniques in computer vision, such as image enhancement, restoration, and super-resolution, can significantly enhance the quality of images, leading to more precise and accurate analysis of the data.
c. Analysis and comprehension of images or objects: By translating pixel data into a form understandable by machines, computer vision empowers computers to understand images or objects. It enables computers to detect and recognize objects, classify them, and even identify patterns and make predictions based on past learning.

In real-world applications, computer vision is used in numerous fields ranging from security and surveillance, healthcare, agriculture, and autonomous vehicles, to industrial automation and retail. For instance, in healthcare, it can help doctors in the early diagnosis of diseases by analyzing medical images. In autonomous vehicles, it aids in object detection, lane line detection, and traffic sign recognition, *inter alia*, thereby improving road safety.

In essence, computer vision bridges the gap between digital technology and the visual world, pushing the boundaries of how machines can perceive, understand, and interact with their surroundings, much like humans. It continues to be an essential tool in developing intelligent systems that can operate efficiently in complex, real-world scenarios.

Deep Learning

Deep learning was originally introduced in the ImageNet Large Scale Visual Recognition Comparison (ILSVRC) image recognition competition held in 2012 and won with the highest level of accuracy. It is a new branch of machine learning in which the system learns to extract deep features through a hierarchical algorithm (Sezavar, A., 2019). Deep learning can be defined as deep learning with artificial neural network (ANN) capabilities, which is part of machine learning. It is the development of an Artificial Neural Network with Multi-Layer Perceptron (MLP) capabilities. MLP is different from Deep Learning, and uses the concept of a fully connected network, while Deep Learning has a simpler network scheme because it uses filters. Deep learning is usually applied to provide solutions such as classification, detection, segmentation, etc. (AlBdairi, A.J.A., 2022) (Fasihah, E., 2023).

Convolution Neural Network (CNN)

Convolutional Neural Networks (CNN) is part of deep learning which is a deep neural network, which means a kind of artificial neural network aimed at recognizing and processing an image (Bergs, T., 2020). This CNN algorithm is specifically designed to process pixel data and visual images. It is usually dedicated to high-dimensional data such as images and videos so that CNN is able to classify images or videos used to detect an object in the image and even in the area around the image (Bose, P., 2020). In this CNN, as convolution is used instead of the usual matrix multiplication, this operation is used at least in one layer. In general, there are two types of layers in CNN (Gheitasi, A., 2020):

a. The image feature extraction layer located at the beginning of the architecture, consisting of several layers, each layer consisting of neurons connected to the local area of the previous layer.
b. Classification Layer, consisting of several layers, each layer consisting of neurons fully connected to other layers.

Based on this, CNN can be considered a method that can transform the original image layer by layer from an image pixel value to a class scoring value in its classification. The following is the CNN architecture used in the research for the training model.

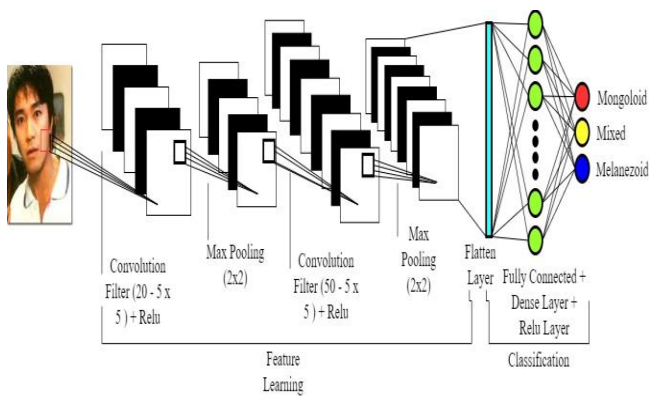


Figure 2: Race Type Architecture with CNN Method Model

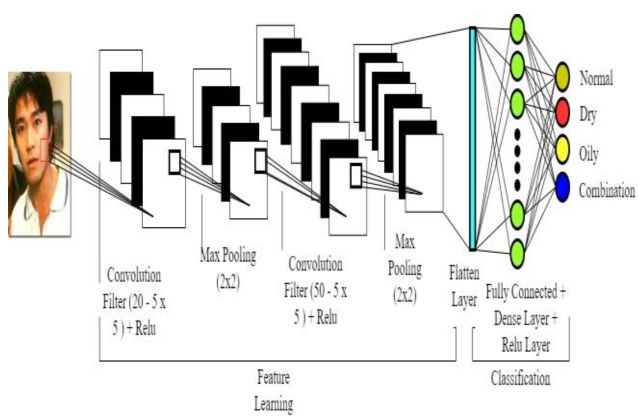


Figure 3: Model Architecture: CNN Method Facial Skin Type

In the visual comparison provided, the left-hand side illustrates the architecture of a standard Artificial Neural Network (ANN), while the right-hand side depicts the layer configuration of a Convolutional Neural Network (CNN). Each architecture has its unique arrangement of layers and processes tailored to the specific tasks they're designed to accomplish.

The convolutional layer in the CNN architecture applies a 5x5 filter, which directly receives the input from the image. This layer operates identically to the convolution operation, capable of performing linear combination filter operations on a localized region of the image. The filter represents the receptive field of the neurons connected to this local area in the input image. As such, the convolution layer has hyperparameters and parameters that control the specifics of these operations and influence the outcome (G. Bruce, 2019).

Following this, a 2x2 size max pooling operation is applied. This operation serves the purpose of dimensionality reduction, reducing the number of parameters by selectively keeping only the maximum values from the specified window, in this case, 2x2. This procedure retains the important features while drastically decreasing the spatial resolution of the feature maps. This, in turn, leads to a more manageable computational load and mitigates the risk of overfitting.

This process of convolution and pooling repeats, each time refining the feature maps to represent the input image's most significant aspects. The iterative process continues until it reaches the final stage known as the fully connected layer. This layer's purpose is to classify the input images into output labels. It is fully connected because it connects every neuron in the previous layer to every neuron in the next layer.

While developing such a system, it is important to ensure the privacy and consent of individuals whose images are used in the datasets. Data privacy principles dictate that explicit permission is obtained from individuals before their images are used. This includes providing a clear explanation of why their images are required, how they will be used and stored, and who will have access to them. Individuals should also be informed about their rights concerning their data, including their right to withdraw consent and request data deletion. This practice safeguards individual privacy and aligns with ethical guidelines for AI and machine learning research.

Haar Cascade

Haar Cascade Classifier is a library available in Open-CV, built on top of the C/C++ language with the python API (Application Programming Interface) (RamyaKalangi, R., 2021). Haar Cascade combines three basic things. The first combination is to have a complex set of features that can be calculated quickly and precisely, thereby reducing the variability within a class and increasing the variability between classes (Qin, H., 2016). The second combination is implementing an algorithm that can perform feature selection and training. The last combination is to form cascades periodically with classification results on detection and schemes that are wider, faster, and more time-efficient. The Haar Cascade function based on this cascade classifier is the object detection method of Paul Viola and Michael Jones. In 2001, they presented a paper titled "Rapid Object Detection Using Simple Enhancements". This Haar Cascade is a collection of Haar-Like functions, which are combined to form a classification. This feature is the sum of the white pixel values subtracted from the pixel values in the black areas (Ma'arif, A., 2021).

Haar Cascade can be trained to detect multiple objects. The next is to determine the area of the face that has the highest probability. The face has skin and a pixel level of color in the skin. The choice of segmentation technique is chosen for the color of the pixels on the face and validated with the haar cascade classifier. If the validated pixel matches the geometric, then the system has found the face in question; if it does not match, the system ignores it.

OpenCV

Python is a high-level programming language that is an interpretive, interactive, object-oriented, versatile illustration with the principle of forming instructions that focus on the level of reading code and can operate on all platforms. Python is a language that combines skills, expertise, and a very simple and clear code syntax and is equipped with a large and complete standard library of functions (Wakurdekar, S., 2020). In this study, the authors used open computer vision (open CV), which is an open-source library dedicated to image processing. The point is that computers have capabilities similar to visual processing in humans. OpenCV has provided many basic computer vision algorithms. and provides an object detection module using computer vision methods (Li, P., 2020).

System Design

In this study, a system was designed to detect facial skin types and races with data in the form of digital images. This system uses algorithms with the CNN and Haar Cascade methods, from which image objects will be retrieved and classified into the type with the highest similarity. System design can be seen in the following flowchart:

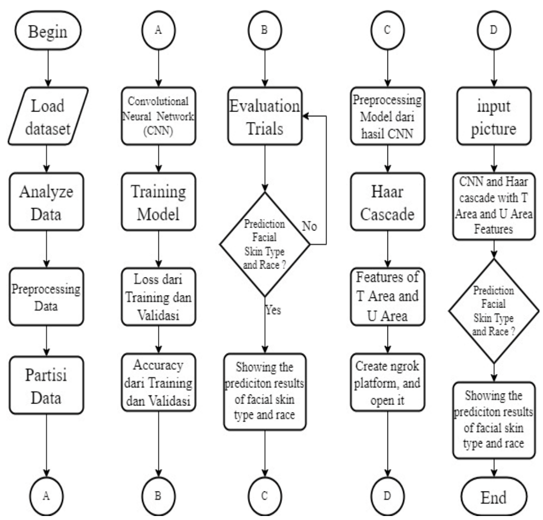


Figure 4: System Flowchart

Flowchart Description

a. Load Data
Data loading stage: In this, storage data from the dataset will be read and retrieved into the system to facilitate further processing. The dataset consists of a race-type dataset and a skin-type dataset.

b. Analyze Data
At the data analysis stage, the data seen in the dataset will be analyzed. Then the data will be analyzed programmatically so that the size of the image pixel scale is appropriate.

c. Preprocessing
At the data preprocessing stage, the image data will be processed for getting a better image when it is processed to the next stage. At the resizing stage, the image data, which originally had different sizes, will be converted into one size for the entire dataset.

d. Data Partition
   At this stage, the data will be partitioned, so that the model can be generalized properly. In the distribution of training data, it is usually divided into two parts, the first designed for training, and the second for validation. In this case, it would typically train the model on 80% of the training data and validate it on the remaining 20% .
e. Convolutional Neural Networks
   At this stage, we create a network for the CNN method, usually to resize matrix images into arrays and enter them as input to the network. Then, in determining the neural network architecture and network parameters, the use of the convolution layer is determined by the input shape layer, convolution layer, relu layer, max pooling layer, convolution layer, relu layer, max pooling layer, flatten layer, density layer, relu layer, dense layer, and the softmax layer, until the output of the output layer is obtained.
f. Training Models
   This stage will train the model that has been designed according to the network parameters that have been made and the training results will be saved.
g. Loss and Validation Training
   This stage will calculate the results of the training model and the resulting loss validation and compares them.
h. Accuracy and Validation Training
   This stage will calculate the results of the resulting accuracy training and validation models and compares them.
i. Trial Evaluation
   This stage provides an evaluation of the results of the comparison of loss and accuracy based on the results of the training model and tests the prediction results using the CNN method.
j. Haar Cascade Classifier
   After the CNN method is complete, the data from the results of the CNN model will be preprocessed, which will then be combined with the Haar Cascade method defined using OpenCV.
k. T-Area and U-Area features
   This stage continues the processing of facial skin type and race prediction results by combining the CNN method and the Haar Cascade Classifier in the T and U area features with a bounding box.
l. Platform application
   The final step is to open the application with the address generated by ngrok and test it by entering an image. The prediction results for facial skin type and race will be displayed based on the previous image input.

3 Discussion and Results

Tests are carried out on image acquisition, preprocessing, training, validation, and testing through the web platform.

Data Load Test

In the early stages of running the system, the program will download the dataset first, which will be read in the program. The following is the result of reading the facial skin type dataset and race type through indicator skin:

Table 1: Data Loading Table

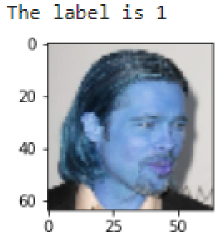
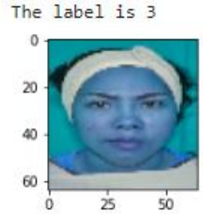
No.	Facial Skin Type	Race Type	Number of Race Type data sets	Total facial skin type data collection
1.		Mongoloid	662	
2.		Mixture	473	
3.		Melanezoid	535	
4.	Normal			17
5.	Combination			14
6.	Dry			12
7.	Greasy			17
Total			1670	60

Based on the table of data load test results, the number of race types was found to have a total of 1670 images, with the dataset taken from the source kaggle.com (https://www.kaggle.com/datasets/zuruoke/race-classification), and the facial skin type a total of 60 images, with a dataset from the author's collection through skincare clinics.

Data Processing Time

This test is carried out on image retrieval or image acquisition from the prepared dataset, after which data of each race and skin type will be resized to 64x64 and preprocessed with a pixel intensity scale of around (0.1), followed by the labeling of data. Here are the results of image acquisition:

Table 2: Data Acquisition Test Results and Preliminary Processing

No.	Acquisition	Image resizing processing time	Label results
1.	Race	6.16 seconds	
2.	Skin	1.5 sec	

Based on the data above, it can be explained that during data acquisition and data preprocessing, between data race and the time needed for the resizing process, the race process takes longer than the skin process.

Training Models

The next test is network testing using the CNN method. The following is the CNN flowchart used:

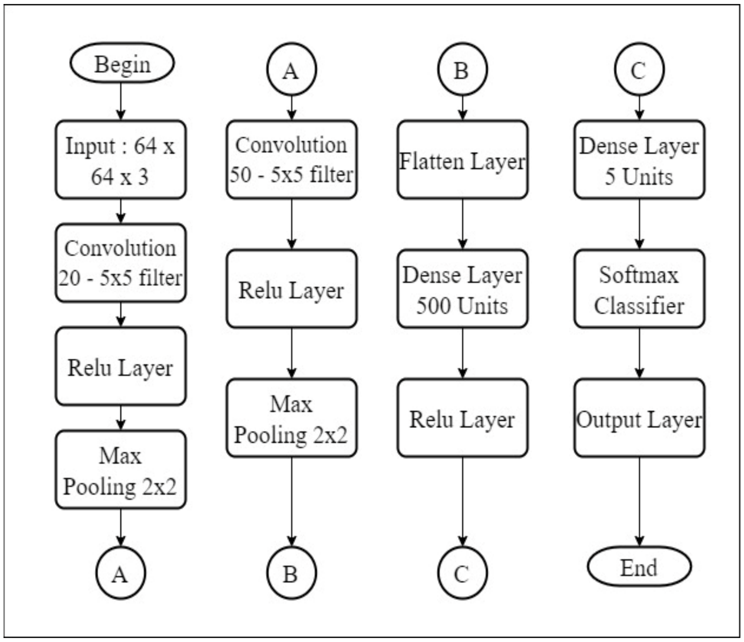


Figure 5: CNN Method Flowchart

This test is carried out after data acquisition and preprocessing so that the data is the same size. This partition will be tested with 80% training and 20% validation. The first configuration is done with a two-dimensional convulsion layer, then a reLU activation layer, and a two-dimensional maximum pooling layer. Further, the second configuration uses a two-dimensional convulsion layer, thereafter the reLU activation layer, and finally, the two-dimensional maximum pooling layer. This is followed by the fully connected and reLU activation layer configuration, and finally, the classification results.

Based on the architecture, the architecture of Figure III.1 is for the CNN method for racial types with skin indicators, while Figure III.2 is for the architecture of the CNN method for facial skin types. In the next stage, the training model is carried out with network parameters at 120 epochs. The following are the results of testing the training model.

```
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 (MaxPooling  (None, 16, 16, 50)       0
 2D)

 flatten (Flatten)           (None, 12800)             0

 dense (Dense)               (None, 500)               6400500

 activation_2 (Activation)   (None, 500)               0

 dense_1 (Dense)             (None, 5)                 2505

 activation_3 (Activation)   (None, 5)                 0

=================================================================
Total params: 6,429,575
Trainable params: 6,429,575
Non-trainable params: 0
_________________________________________________________________
```

Figure 6: CNN Race Model Results

```
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 (MaxPooling  (None, 32, 32, 20)       0
 2D)

 conv2d_3 (Conv2D)           (None, 32, 32, 50)        25050

 activation_5 (Activation)   (None, 32, 32, 50)        0

 max_pooling2d_3 (MaxPooling  (None, 16, 16, 50)       0
 2D)

 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: 6,429,575
Trainable params: 6,429,575
Non-trainable params: 0
_________________________________________________________________
```

Figure 7: CNN Skin Model Results

Table 3: Epoch Test Results on Race

Epoch	Lost	Accuracy	Val_Loss	Val_Accuracy
1	1.2277	0.4277	0.9264	0.6467
10	0.0724	0.9820	0.8097	0.7665
20	7.2529e-04	1.0000	1.2552	0.7575
30	2.0889e-04	1.0000	1.3847	0.7515
40	8.7078e-05	1.0000	1.4827	0.7665
50	4.6046e-05	1.0000	1.5603	0.7665
60	2.7920e-05	1.0000	1.6197	0.7754
70	1.7974e-05	1.0000	1.6622	0.7754
80	1.2278e-05	1.0000	1.7245	0.7754
90	8.7203e-06	1.0000	1.7657	0.7754
100	6.1038e-06	1.0000	1.8049	0.7754
120	3.5166e-06	1.0000	1.8828	0.7754
AVG	0.108437	0.950808	1.489983	0.758967

This test was carried out after modeling on race and skin; 120 epoch tests were carried out, with a race and skin LR of 0.001 and BS of 32. By using the Adam optimizer on a 64x64 image, an epoch value of 120 was obtained to produce the best accuracy in testing the race model, with an average of 0.950808 and the smallest loss with an average of 0.108437.

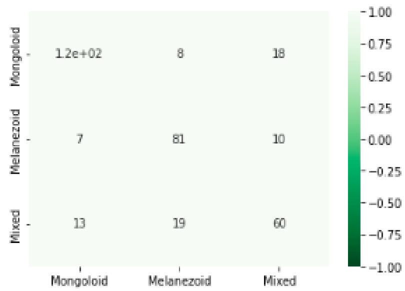
Table 4: Skin Epoch Test Results

Period	Lost	Accuracy	Val_Loss	Val_Accuracy
1	1.8792	0.1250	1.6382	0.1667
10	0.6250	0.2292	1.7315	0.2500
20	0.0146	1.0000	2.8585	0.2500
30	0.0012	1.0000	4.4604	0.2500
40	2.8025e-04	1.0000	4.7507	0.3333
50	1.8041e-04	1.0000	4.6466	0.3333
60	1.4826e-04	1.0000	4.7503	0.3333
70	1.2440e-04	1.0000	4.7846	0.3333
80	1.1213e-04	1.0000	4.8341	0.3333
90	9.9350e-05	1.0000	4.8790	0.3333
100	8.8446e-05	1.0000	4.9161	0.3333
120	7.1417e-05	1.0000	5.0077	0.3333
AVG	0.210092	0.86285	4.104808	0.295436

This test was carried out after modeling on race and skin; 120 epoch tests were carried out, with a race and skin LR of 0.001, and BS of 32. By using the Adam optimizer on a 64x64 image, an epoch value of 120 was obtained to produce the best accuracy in skin model testing, with an average of 0.86285 and the smallest loss with an average of 0.210092.

Race and Skin Matrix Tests

This test is carried out after the epoch test is obtained. The next is the confusion matrix test from the best test on race and skin.



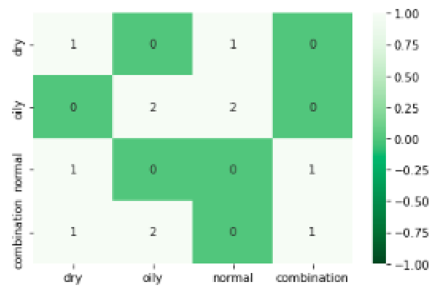


Figure 8: Confusion Matrix Contest Results Figure 9: Skin Confusion Matrix Results

Test Classification Report

This test is carried out after obtaining the epoch and confusion matrix tests. The next is to test the results of the classification report-the following are the results thereof.

	precision	recall	f1-score	support
0	0.86	0.82	0.84	144
1	0.75	0.83	0.79	98
2	0.68	0.65	0.67	92
accuracy			0.78	334
macro avg	0.76	0.77	0.76	334
weighted avg	0.78	0.78	0.78	334

The Model Accuracy = 77.54%

	precision	recall	f1-score	support
0	0.33	0.50	0.40	2
1	0.50	0.50	0.50	4
2	0.00	0.00	0.00	2
3	0.50	0.25	0.33	4
accuracy			0.33	12
macro avg	0.33	0.31	0.31	12
weighted avg	0.39	0.33	0.34	12

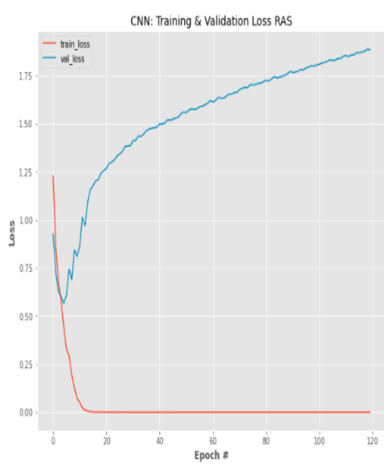
The Model Accuracy = 33.33%

Figure 10: Race Classification Report Test Results Figure 11: Skin Classification Report Test Results

This test is carried out after obtaining the epoch and confusion matrix values from the best tests on race and skin; then, the results of the classification report will produce an accuracy of 77.54% for race, and 22.50% for skin accuracy.

Loss Testing and Validation Training

This test is carried out after obtaining the accuracy results on race and skin so that training results can be compared with loss validation. Here are the test results:



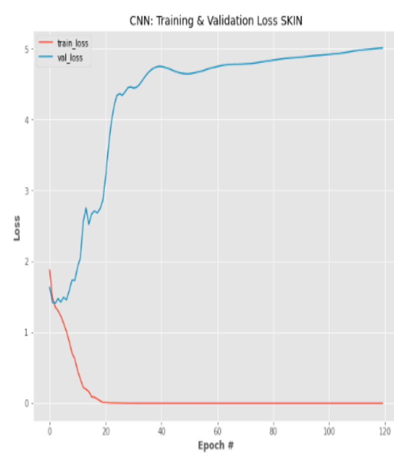


Figure 12: Race Loss Validation VS Training Figure 13: Training VS Skin Validation

This test reveals that the results of training loss with loss validation on race and skin type have similar results: the graph goes up on validation loss and down on training loss.

Validation Testing and Accuracy Training

This test is carried out after obtaining the accuracy results on race and skin so that training results can be compared with validation accuracy. Here are the test results:

Figure 14: Race Accuracy Validation VS Training

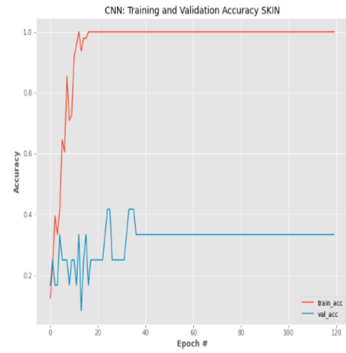


Figure 15: Skin Accuracy VS Validation Training

This test indicates that the results of training accuracy with validation accuracy on the skin have a fairly linear graph, but on the race results, the graphs of training accuracy and validation accuracy tend to be more linear, so that classification accuracy can produce the best results.

Haar Cascade Test

The next is to add the Haar Cascade Classifier method through opencv, which will be combined using the CNN method based on the T and U area features. The following is a flowchart of the Haar Cascade Classifier method:

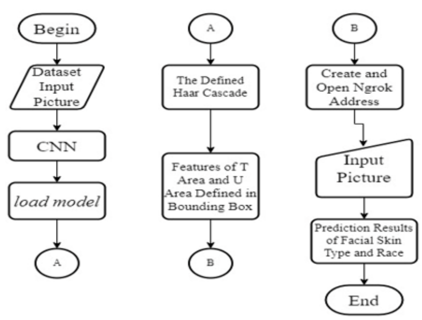


Figure 16: Haar Cascade Flowchart

Image Test

This test is done to ensure that the results of the classification test on race and skin type are successful. Here are the test results:

Figure 17: Race Label Test Results

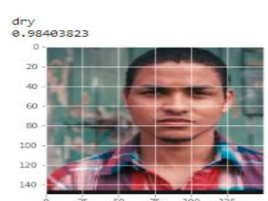


Figure 18: Skin Label Test Results

This test reveals that the test regarding the classification of race and skin type is successful, as illustrated in Figure III.15 It can detect that the image has mixed race, while Figure III.16 can detect that the image has mixed race and a normal skin type.

Testing with the Web Platform

This test is conducted to test the system as a whole in detecting facial skin type and race using the CNN and Haar Cascade methods on a website-shaped platform. The test results are as follows.

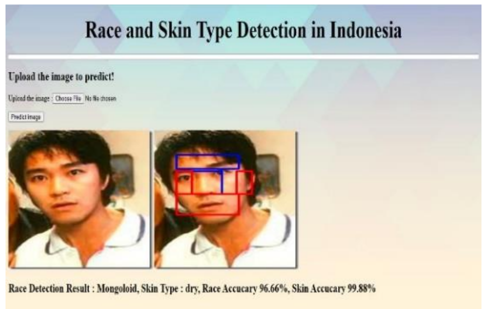
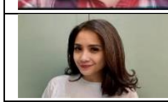
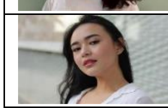


Figure 19: System Test Results with Platforms

Table 5: Overall Test Results

Picture	Race Type	Skin type	Accuracy
	Melanezoid	Normal	Race: 100.0 Skins: 99.7
	Mongoloid	Normal	Race: 96.66 Skins: 99.88
	Mixture	Normal	Race: 97.93 Skins: 98.81
	Melanezoid	Normal	Race: 99.78 Skins: 99.61
	Melanezoid	Normal	Race: 99.97 Skins: 99.98

This test provides information that the overall testing system regarding the classification of race and skin type is successful, as exemplified in Figure III.17. It can detect that the image has a melanezoid race through the website platform, while in Table III.5 it has tested 5 new image objects and produced race type and facial skin type, with accuracy on race and facial skin type.

4 Conclusion

The facial skin type and race detection system based on the features of the T and U areas using the Convolutional Neural Networks and Haar Cascade methods has been successfully implemented using 1670 race datasets and 60 facial skin type datasets. This system is very accurate because the use of Adam optimizer ensures the best level of accuracy in detecting race and skin type with 120 epochs, after testing predetermined images and datasets. The loss validation on race detection has an average of 0.108437, and that on skin detection has an average of 0.210092. Hence, the Convolutional Neural Networks method and the Haar Cascade method are highly appropriate for measuring the level of classification in the detection of the skin type and race of a person. This system works through a web-shaped platform, by inputting images that have been determined to be detected. Then CNN and Haar Cascade will work to classify images including skin type and race with categories from the results of the training samples done before. After the two methods have carried out the classification, the image will reappear with a box marked with the results of the classification in the T and U areas, and generate skin type and race type, through the accuracy obtained from the detection system. By inputting the predetermined images to be detected, CNN and Haar Cascade will work to classify these images into skin types and races with categories from the results of the training samples carried out previously. Through the Ngrok platform, the system is implemented with a simple UI display, and by entering the image to be detected and clicking the prediction button, the race type and facial skin type results will appear, accompanied by the accuracy of the detection of both types. Based on system tests with five new object samples, the system successfully detects the Melanezoid race type in the first, fourth, and fifth new objects, the Mongoloid race type in the second new object, the mixed-race type in the third new object, and normal facial skin types in the five new object samples, with an average accuracy of 98% in race detection and 97% in skin type detection. Thereafter, the type of race and type of facial skin results will appear, accompanied by the accuracy of the detection thereof. Based on system tests carried out with five new object samples, the system successfully detects the Melanezoid race type in the first, fourth, and fifth new objects, the Mongoloid race type in the second new object, the mixed-race type in the third new object and normal facial skin types in the five new object samples, with an average accuracy of 98% in race detection and 97% in skin type detection. then the type of race and type of facial skin results will appear accompanied by the accuracy of the detection of the type of race and type of facial skin. Based on system tests with five new object samples that have been carried out, the system successfully detects the melanezoid race type in the first, fourth, and fifth new objects, the mongoloid race type in the second new object, the mixed-race type in the third new object and has normal facial skin types in the five new object samples, with an average accuracy of 98% in race detection, and 97% in skin type detection.

References

[1] AlBdairi, A.J.A., Xiao, Z., Alkhayyat, A., Humaidi, A.J., Fadhel, M.A., Taher, B.H., & Al-Shamma, O. (2022). Face recognition based on Deep Learning and FPGA for ethnicity identification. *Applied Sciences*, *12*(5), 1-15.

[2] Ballard, T.P. (2019). *The Missionary Enterprise, Racial Conflicts, and the Transformation of American Evangelicalism, 1945–1980*. University of Montana.

[3] Bergs, T., Holst, C., Gupta, P., & Augspurger, T. (2020). Digital image processing with deep learning for automated cutting tool wear detection. *Procedia Manufacturing*, *48*, 947-958.

[4] Bose, P., & Bandyopadhyay, S.K. (2020). Facial spots detection using convolution neural network. *Asian Journal of Research in Computer Science*, 71-83.

[5] Bragazzi, N.L., Sellami, M., Salem, I., Conic, R., Kimak, M., Pigatto, P.D.M., & Damiani, G. (2019). Fasting and its impact on skin anatomy, physiology, and physiopathology: A comprehensive review of the literature. *Nutrients*, *11*(2), 1-15.

[6] Chan, H.T., Lin, T.Y., Deng, S.C., Hsia, C.H., & Lai, C.F. (2021). Smart Facial Skincare Products Using Computer Vision Technologies. *In IEEE Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 1674-1677.

[7] Fasihah, E., & Al Mubarokah, N.H. (2023). Validation Analysis of Learning Development on Optic Material with Core Learning Model. *SAGA: Journal of Technology and Information System*, *1*(1), 1-4.

[8] G. Bruce, *Structures and functions of prosody*. 2019.

[9] Gheitasi, A., Farsi, H., & Mohamadzadeh, S. (2020). Estimation of hand skeletal postures by

using deep convolutional neural networks. *International Journal of Engineering*, *33*(4), 552-559.

[10] Guo, G., Wang, H., Yan, Y., Zheng, J., & Li, B. (2020). A fast face detection method via convolutional neural network. *Neurocomputing*, *395*, 128-137.

[11] Hochman, A. (2019). Race and reference. *Biology & Philosophy*, *34*(2), 1-40.

[12] Hemasree, V., & Kumar, K.S. Facial Skin Texture and Distributed Dynamic Kernel Support Vector Machine (DDKSVM) Classifier for Age Estimation in Facial Wrinkles. *Journal of Internet Services and Information Security (JISIS)*, *12*(4), 84-101.

[13] Juwanda, F.S., & Zin, H.M. (2021). The development of skin analyser for skin type and skin problem detection. *Journal of ICT in Education*, *8*(3), 27-37.

[14] Kharat, D., Kajawe, P., Kopnar, S., Jinabade, N., & Wagh, J. (2019). Face detection using Haar Cascade Classifier. *International Journal of Innovative Research in Science, Engineering and Technology*, *8*(4), 3712-3715.

[15] Koshy, R., & Mahmood, A. (2019). Optimizing deep CNN architectures for face liveness detection. *Entropy*, *21*(4), 1-16.

[16] Li, P., Wang, H., Li, Y., & Liu, M. (2020). Analysis of face detection based on skin color characteristic and AdaBoost algorithm. *In Journal of Physics: Conference Series*, *1601*(5), 1-9. IOP Publishing.

[17] Li, X., Yang, Z., & Wu, H. (2020). Face detection based on receptive field enhanced multi-task cascaded convolutional neural networks. *IEEE Access*, *8*, 174922-174930.

[18] Lin, H.Y., Lin, C.J., Jeng, S.Y., & Yu, C.Y. (2021). Integrated Image Sensor and Hyperparameter Optimization of Convolutional Neural Network for Facial Skin Detection. *Sensors and Materials*, *33*(8), 2911-2924.

[19] Lin, T.Y., Chan, H.T., Hsia, C.H., & Lai, C.F. (2022). Facial Skincare Products' Recommendation with Computer Vision Technologies. *Electronics*, *11*(1), 1-13.

[20] Ma'arif, A., & Wulandari, A. (2021). Face shape-based physiognomy in LinkedIn profiles with cascade classifier and k-means clustering. *In IEEE 8th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 347-353.

[21] Malhotra, K.C., & Vasulu, T.S. (2019). Development of typological classification and its relationship to microdifferentiation in ethnic India. *Journal of Biosciences*, *44*, 1-14.

[22] Qin, H., Yan, J., Li, X., & Hu, X. (2016). Joint training of cascaded CNN for face detection. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3456-3465.

[23] RamyaKalangi, R., Lakshmiteja, N., Haritha, K., Sonika, K., & Jahnavi, T. (2021). Real Time Face Detection With Image Dataset Load On Haar Cascade Algorithm. *Turkish Journal of Computer and Mathematics Education*, *12*(7), 2366-2374.

[24] Rosda, R., Wali, M., & Imilda, I. (2023). Evaluation of the Successful Implementation of the SIMARDI Using the Technology Acceptance Model (TAM). *SAGA: Journal of Technology and Information System*, *1*(2), 56-63.

[25] Sezavar, A., Farsi, H., & Mohamadzadeh, S. (2019). A modified grasshopper optimization algorithm combined with cnn for content based image retrieval. *International Journal of Engineering*, *32*(7), 924-930.

[26] Vo, T., Nguyen, T., & Le, C.T. (2018). Race recognition using deep convolutional neural networks. *Symmetry*, *10*(11), 1-15.

[27] Wakurdekar, S., Rikhra, A., Sharma, M., & Singh, S. (2020). Real Time Facial Recognition Using Opencv Python Techniques. *International Journal of Science and Research Development (IJSDR*), *5*(8), 392–396.

[28] Yu, B., & Tao, D. (2018). Anchor cascade for efficient face detection. *IEEE Transactions on Image Processing*, *28*(5), 2490-2501.

[29] Zalmi, W.F., Sitompul, B.J., Nduru, S.W., & Lumbangaol, S.K. (2023). Welder Recruitment Decision Support System Using the SMARTER Method. *SAGA: Journal of Technology and*

Information System, *1*(2), 44-49.
[30] Zhang, K., Zhang, Z., Wang, H., Li, Z., Qiao, Y., & Liu, W. (2017). Detecting faces using inside cascaded contextual CNN. *In Proceedings of the IEEE international conference on computer vision*, 3171-3179.
[31] Zheng, Q., Purwar, A., Zhao, H., Lim, G.L., Li, L., Behera, D., & Kot, A.C. (2022). Automatic facial skin feature detection for everyone. *34*(8), 1-6.

Authors Biography


Indriyani

Indriyani is a lecturer in the Information System Department, Institute of Technology and Business STIKOM Bali, Indonesia. She finished her Magister degree in the Postgraduate Program, Informatics Engineering, Institute of Technology Sepuluh Nopember, Surabaya, Indonesia. Currently, she is enrolled in the Doctor of Engineering Program, at Udayana University, Bali, Indonesia.
E-mail: indriyani@stikom-bali.ac.id, indry.joice@gmail.com,
Orcid: https://orcid.org/0000-0002-8371-1240


Ida Ayu Giriantari

Ida Ayu Dwi Giriantari is Professor in the Electrical Engineering Department, born on 13 December 1965. She finished her undergraduate study in the Electrical Engineering Department at Udayana University in 1990. The next year, she started working as a lecturer in the department she graduated from. In 1997, she was granted a scholarship by the Indonesian Government for continuing studies at The University of New South Wales, Sydney, Australia. She obtained her Master of Engineering Science (MEngSc) from the School of Electrical Engineering and Telecommunication UNSW in 1999. Under the same scholarship and at the same University, she obtained her Doctor of Philosophy (PhD) in 2003, majoring in Electric Power. As a Senior Member of IEEE (Institute of Electronic and Electrical Engineering), she participates in the program committee at many international conferences and works as a reviewer for various journals, including Frontier of Energy. She lectures on several many subjects; in particular, Renewable Energy, Energy Management, Energy tariff, Photovoltaic, Smart Grid, High Voltage Technology, and Electric Circuit. E-mail: dayu.giriantari@unud.ac.id,
Orcid: https://orcid.org/0000-0002-2326-2594


Made Sudarma

Prof. Dr.Ir. Made Sudarma, MASc., CSEP., CPMP., IPU., ASEAN Eng. is
Professor of Information Technology in the Engineering Faculty Udayana University - Bali Indonesia. Department of Electrical and Computer Engineering
RESEARCH INTERESTS:
•Artificial Intelligence
• Neural Network and Fuzzy Logic
•Analysis Design and Algorithm
•Bioinformatics and Scientific Computing
•Cloud Computing
•Genetic Algorithm
•Image Processing
•Network Security
• Optimization
E-mail: msudarma@unud.ac.id, Orcid: https://orcid.org/0000-0002-8331-0519


I Made Oka Widyantara

I Made Oka Widyantara received a B.Eng. degree in electrical engineering from Sepuluh Nopember Institute of Technology (ITS) Surabaya, Indonesia, in 1997. He completed his M. Eng. degree in Telecommunication Information Systems from Bandung Institute of Technology (ITB), Bandung, Indonesia, in 2001 and a Dr. degree in multimedia telecommunication from Sepuluh Nopember Institute of Technology (ITS) Surabaya, Indonesia, in 2013. He is currently a lecturer in the Department of Electrical Engineering, at Udayana University, and head of the Information Technology Department, Engineering Faculty, Udayana University, Indonesia. His research interests include intelligent signal processing, image/video processing, and compression.
E-mail: oka.widyantara@unud.ac.id, Orcid: https://orcid.org/0000-0002-0979-5116