International Journal of Social Science, Educational, Economics, Agriculture Research and Technology (IJSET) | ISSN (e) : 2827-766X
Volumes 4 No. 12 (2025)

IMPLEMENTATION OF DEEP LEARNING ALGORITHM FOR PT GROWTH SUMATERA'S FACE DETECTION ATTENDANCE SYSTEM

Norita Tampubolon[1]*, Penggabean Siahaan[2], Lewika Tampubolon[3],Zailani Sinabariba[4]. Muhammad Syahputra Novelan[5].
Universitas Panca Budi, Medan
E-mail: queen.norita07@gmail.com, gabesiahaan@gmail.com, lewikatampubolon87@gmail.com, queen.norita07@gmail.com, putranovelan@dosen.pancabudi.ac.id.

Received	: 15 October2025	Published	:20 January 2026
Revised	: 01 November 2025	DOI	:https://doi.org/10.54443/ijset.v4i12.1552
Accepted	: 25 November 2025	Publish Link	:https://www.ijset.org/index.php/ijset/index

Abstract

Attendance is a data collection activity to determine the number of employees present, arrival times, and departure times in a company. Attendance is divided into two types: manual and automatic. Manual attendance is an attendance process carried out using a handwritten note or signature form. Automatic attendance is an attendance process that involves technology. With facial recognition technology, an attendance system can be developed. Facial recognition technology is a computer technology that functions to determine facial location, facial size, feature detection, background image ignoring, and facial image identification. Facial recognition involves several variables, such as source images, processed images, extracted images, and a person's identity data. Deep learning with Convolutional Neural Networks is one method used to predict and classify different human facial images. This facial detection attendance system application is designed and built on a desktop platform, using the Python programming language. The application of deep learning algorithms with convolutional neural networks (CNN) in this facial detection attendance system can streamline the existing attendance system.

Keywords: Face Detection, Deep Learning, Convolutional neural network, CNN

INTRODUCTION

Technological and scientific advancements have continuously developed and impacted human life. These advances have had a positive impact that is universally felt, such as increased information accessibility and improved communication facilities. Technological and information advancements have significantly changed various aspects of life, including the industrial sector, particularly employee attendance systems. Attendance systems are essential in various sectors, including companies, government agencies, private institutions, and others, to monitor and track the attendance of employees working in those sectors. Currently, attendance systems can be created using human body parts such as fingers and the face. The face is a unique part of the human body. Every person in the world has a unique facial contour. Therefore, everyone uses their face to identify themselves so that others can recognize them. The face is used as a part that can be recognized by computers. Face detection and facial recognition are techniques used to perform facial recognition processes on computers. PT Growth Sumatera still uses a manual system, namely only using a form that is checked by the employees who are present. This form is used to control employee attendance. With such an attendance system, it will not be good because employees might check the attendance list yesterday or tomorrow, then HRD will have to process it twice to input the results of the employee attendance form.

FORMULATION OF THE PROBLEM

From the background above, the following problem formulation can be taken:

1. How is the attendance system and application designed for PT Growth Sumatera employees?
2. How is the deep learning algorithm implemented at PT Growth Sumatera?

METHOD

This scientific article uses a combination of qualitative and quantitative methods. The descriptive qualitative method utilizes a literature review, which is research based on literature reviews and analysis of the development of problems in society. This qualitative method is used to obtain statistical data, through participant observation, and in-depth interviews. The research analysis conducted in this study involves a literature review, reading, and further observation of companies to draw conclusions based on the research findings.

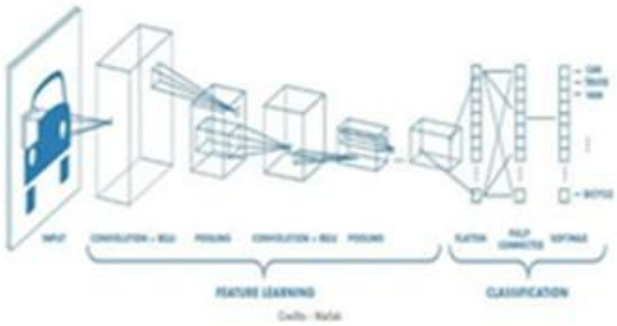
RESULTS AND DISCUSSION

Deep Learning

Deep learning is a branch of machine learning that specifically focuses on the use of multi-layered neural networks to mimic the way the human brain processes information. The term "deep" refers to the depth of the neural network, where the model consists of multiple layers that gradually become able to understand increasingly complex and abstract patterns in data. Deep learning is sometimes called deep structured learning or hierarchical learning.

Convolutional Neural Network

Convolutional Neural Network (CNN) is a development of multilayer perceptron (MLP) designed to process two-dimensional data in the form of images. CNN is included in the type of Deep Neural Network due to its high network depth and is widely applied to image data. Basically, image classification can be done with MLP, however, the MLP method is less suitable for use because it does not retain spatial information from the image data and considers each pixel as an independent feature, resulting in poor results. The initial research that underlies the discovery of CNN was first conducted by Hubel and Wiesel (Hubel & Wiesel, T, 1968) on the visual cortex in the cat's sense of sight. Technically, CNN is a trainable architecture and consists of several stages. The input and output of each stage consist of several arrays commonly called feature maps. Each stage consists of three layers: convolution, activation function layer, and pooling layer.
The following is the architecture of the Convolutional Neural Network:



Based on the image above, the first stage in the CNN architecture is the convolution stage. This stage is carried out using a kernel of a certain size. The calculation of the number of kernels used depends on the number of features generated. Then proceed to the activation function, usually using the ReLU (Rectifier Linear Unit) activation function. Next, after exiting the activation function process, it goes through the pooling process. This process is repeated several times until sufficient feature maps are obtained to proceed to the fully connected neural network, and from the fully connected network is the output class.

1. Convolution layer

   This is part of the CNN architecture stage. This stage performs a convolution operation on the output of the previous layer. This layer is the main process underlying the CNN architecture network. Convolution is a mathematical term that describes the repeated application of one function to the output of another function. The convolution operation operates on two real-valued argument functions. This operation applies the output function as a Feature Map to the input image. These inputs and outputs can be viewed as two real-valued arguments.

   The hyperparameters used in the equation below are used to calculate the number of activated neurons in a single output. Consider the following equation.

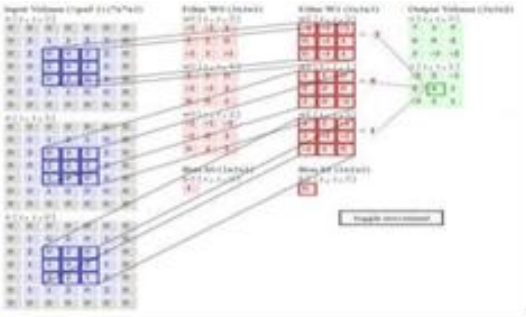
   $(WF + 2P)/(S + 1)$

   Information :

   W = Image volume size

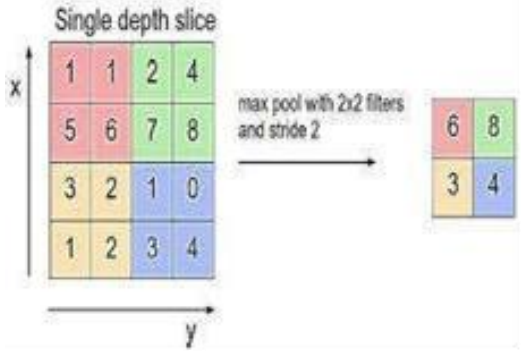
P = Padding value used
F = Filter Size
S = Stride Size

Based on the equation above, the spatial size of the output volume can be calculated where the hyperparameters used are the volume size (W), filter (F), applied Stride (S) and the amount of zero padding used (P). Stride is the value used to slide the filter through the input image and Zero Padding is the value to get zeros around the image border. A convolutional layer consists of neurons arranged in such a way that they form a filter with a length and height (pixels). For example, the first layer in a feature extraction layer is usually a conv. layer with a size of 5x5x3. The length is 5 pixels, the height is 5 pixels, and the thickness/number of filters is 3, corresponding to the image's channels. These three filters will be shifted throughout the image. Each shift will perform a "dot" operation between the input and the filter value, resulting in an output, commonly referred to as an activation map or feature map. See the following illustration.



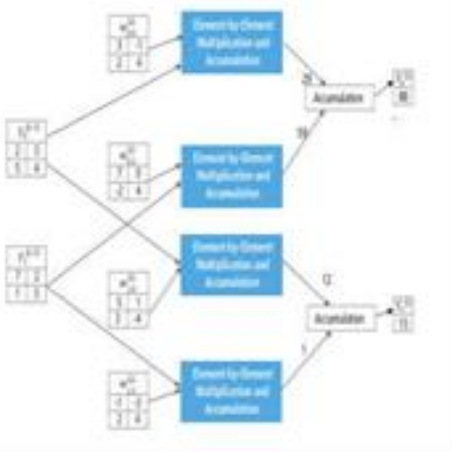
2. Pooling Operation

Pooling is a reduction in matrix size using pooling operations. The Pooling Layer is usually located after the conv. Basically, the pooling layer consists of a filter with a certain size and stride that will alternately shift across the entire feature map area. In the pooling layer, there are two types of pooling commonly used, namely average pooling and max-pooling. The value taken in average pooling is the average value, while in max-pooling it is the maximum value. Pooling layers inserted between successive convolutional layers in the CNN model architecture can progressively reduce the size of the output volume on the Feature Map, thereby reducing the number of parameters and calculations in the network, to control Overfitting. The pooling layer works on each feature map stack and reduces its size. The form of the pooling layer generally uses a 2x2 filter that is applied with two steps and operates on each slice of its input. The following is an example of a max-pooling operation image.



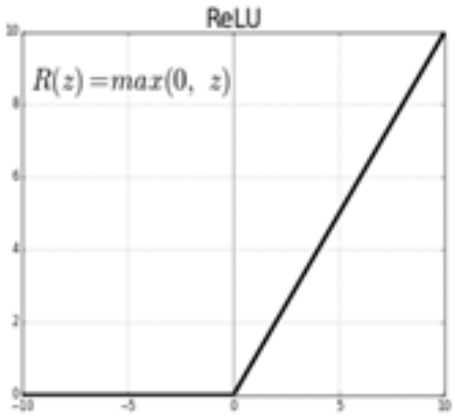
3. Fully-Connected Layer

A Fully-Connected Layer is a layer where all activated neurons from the previous layer are connected to neurons in the next layer, just like a regular neural network. This layer is typically used in Multi-Layer Perceptrons (MLPs), which aim to transform data dimensions so that data can be classified linearly. The difference between a fully connected layer and a regular convolutional layer is that the neurons in a convolutional layer are connected only to specific regions of the input, while a fully connected layer has neurons that are fully connected. However, both layers still operate on dot products, so their functionality is not significantly different. Here's the fully connected layer process:



4. ReLU Activation Function

Essentially, the ReLU (Rectified Linear Unit) function performs a threshold from 0 to infinity. This function is currently one of the most popular functions. The following is a graph of the ReLu activation function.



In this function, if the input from the neurons is a negative number, the function will translate that value to 0. If the input is positive, the output of the neuron is the activation value itself. This activation function has the advantage of speeding up the configuration process performed with Stochastic Gradient Descent (SGD) compared to the sigmoid and tanh functions. However, this activation also has the disadvantage of being fragile during the training process and can cause the unit to die.

Absence

Attendance is the process of creating data for an attendance register, commonly used by institutions or agencies that urgently require such a system. Attendance represents a system that must be used as an attendance system concept. When the system requires data, it will be used as an application capable of running and creating the attendance data.

Image

An image is a representation (picture), similarity, or imitation of an object. The output of a data recording system can be optical in the form of a photograph, analog in the form of video signals such as images on a television monitor, or digital which can be directly stored on a storage medium. Literally, an image is a picture in a two-dimensional plane. Mathematically, a continuous function of light intensity in a two-dimensional plane constitutes an image. A light source illuminates an object, and the object then reflects back part of the light beam. This reflected light is then captured by optical instruments, such as the human eye, cameras, scanners, and so on. Thus, the recorded image is a shadow of the object.

Image Processing

Digital image processing is a science that studies image processing techniques. An image in question is a still image (photo) or a moving image that comes from a webcam/camera. What is meant by digital here is image processing or images that are done digitally using a computer. Basically, an image processing system consists of objects to be processed, which will be used to represent objects in the form of digital images, will produce a new digital image to be analyzed, so that the image analysis process to get the desired results. Digital image processing has several objectives, including improving image quality seen from radiometric aspects (contrast enhancement, color transformation, image restoration) and from geometric aspects (rotation, translation, scale, geometric transformation), carrying out the process of extracting information and object descriptions or recognizing objects contained in the image, performing compression or data reduction for the purpose of data storage, data transmission, and processing time.

Face Detection and Recognition

1. Feature Extraction

   A feature is a distinctive sign that distinguishes one image from another. It is no different from any other image. (Indra, 2016). Each image feature is obtained from the feature extraction process. Processes that can be used to determine shape features are edge detection, threshold, segmentation, and moment calculations such as mean, median, and standard deviation of each local image.

2 Face Detection

   Face detection is a computer technology used in several systems and applications to detect faces. This technology is built using specific algorithms focused on human detection. In facial recognition technology, face detection is the initial stage of processing to recognize a person's face. Face detection determines where facial features appear in the input image. The success of this face detection process has a significant impact on the performance and usability of a facial recognition system.

3. Face Recognition (Face Recognition)

   The face is the focal point of human interaction, a key component of human identity and emotion. We can recognize thousands of faces due to frequent, fleeting interactions, even over long periods of time. We can also recognize someone even if they've changed in age, wearing glasses, or a change in hairstyle. Therefore, many face recognition systems utilize the face as a means of recognition.

Python Programming Language

The Python programming language was first released in 1991 by Guido van Rossum at the Scitchting Mathematisch Centrum in the Netherlands. Guido chose the name Python because he was a fan of the British comedy group Monty Python. Python is open source, with most versions licensed under the GFL. Like most applications, each version has its own unique characteristics. Python 2 was developed from Python 1 with various additional features. Compared to version 1, Python 2 is more transparent and inclusive for application development. A new feature of Python 2 is the Python Enhancement Proposal (PEP), which can provide informative guidance for its users. Python 2 has improved support for Unicode and is equipped with programmatic features, including memory management. Python version 2 is very well developed and is quite widely used by programmers. Python 3 is an extension of Python 2 with a focus on tidying up the codebase and eliminating redundancy. The biggest change in Python 3 includes making print statements in built-in functions. One of Python 3's weaknesses is its lack of compatibility with Python 2, so programmers are very cautious about migrating to this version. However, with the increasing development of Python 3 and the decreasing support for Python 2, this version has gradually increased in use.

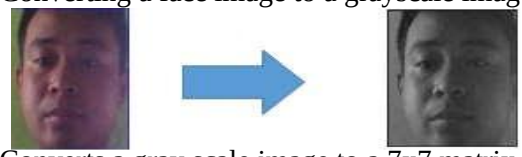
Application System Design

To do attendance with faces on this system, employees must first register by filling in the NPK name and inputting a facial image that is supervised by HRD/admin staff, then do facial image training with CNN and supervised by HRD/admin staff, after doing image training, then employees do attendance when entering or leaving, and the recording of the results of the facial detection attendance will be summarized every month by HRD, then the results of the report are submitted to the head of HRD.

1. Formation of methods/algorithms

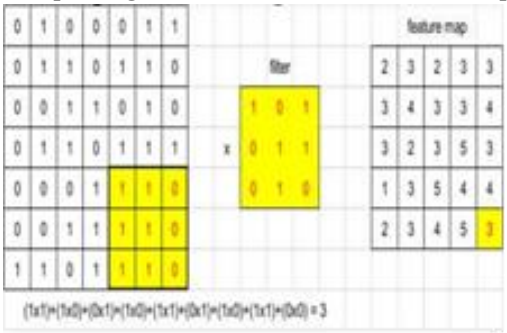
   Facial recognition in this system uses a deep learning algorithm, namely a convolutional neural network. A convolutional neural network has several stages, namely:

a. Converting a face image to a grayscale image



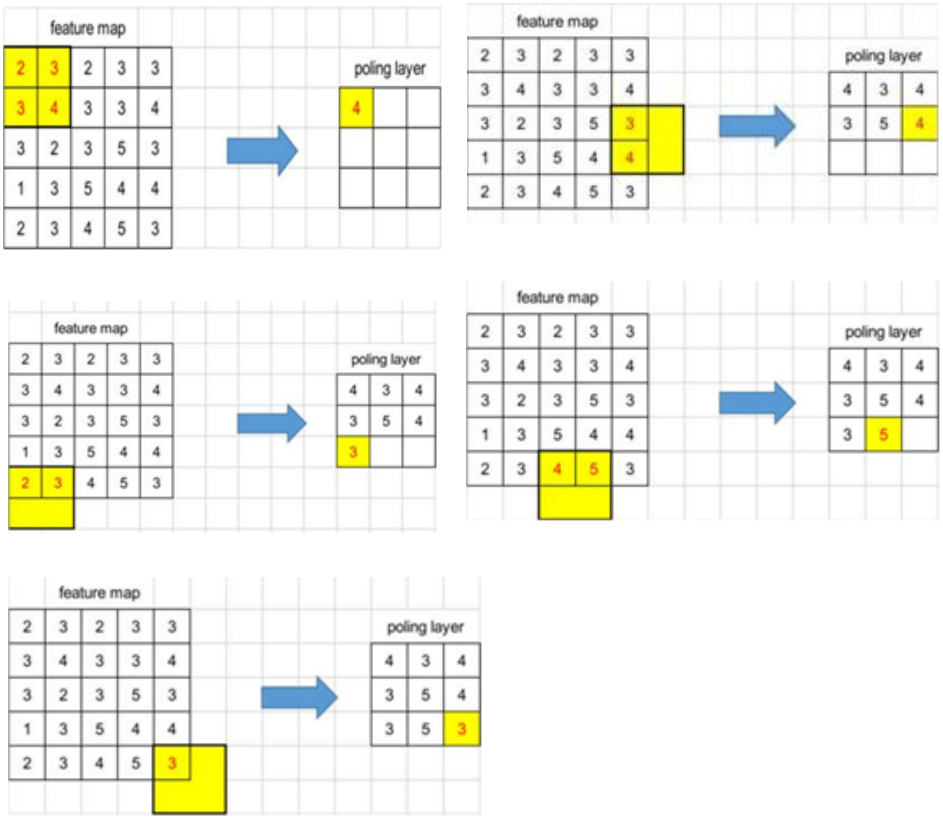
b. Converts a gray scale image to a 7x7 matrix arrangement

c. convolution, which is multiplying the 7x7 matrix of the resulting changes with the 3x3 kernel/filter matrix and placing the results in a 5x5 feature map



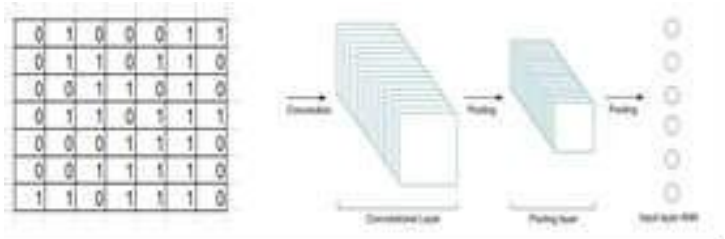
d. After calculating the convolution, the next step is the activation function. The activation function used is ReLu, which changes all negative numbers to 0. Because the results of the convolution are all positive, the activation function is still used, but there is no calculation in this case.

e. Max pooling, in this step we determine the shift of the 2x2 matrix from right to left, which is taken from the convolution result matrix, namely the features map or convolution layer, then the largest value of the 2x2 matrix is placed in the 3x3 polling layer,



f. Flattening: The flattening stage converts the matrix in the pooling layer into a single column (a single vector). This vector will then become part of the input layer in an artificial neural network (ANN).

From the pooling layer (on the left) from the previous process, we simply take each row and combine them into one column. The first row (4, 3, 4) is combined with the second row (3, 5, 4) and then the third row (3, 5, 3). After we perform flattening, we input the results into the input layer of an ANN. The display is as follows:



All pooling layers will be a single vector. So, if a pooling layer is a 3×3 matrix, it will be a single vector with 9 rows. If there are 10 pooling layers of the same size (3×3), then there will be a single vector with 90 rows as input to the ANN.

g. Fully connected layer, after the flattening process is carried out, then the fully connected layer process is carried out, namely in it there is an artificial neural network (ANN)

Formation of a Rule Base

The rule base of the convolutional network algorithm in this system includes training datasets of facial images and making predictions.

Rules or Rules for Training Facial Image Dataset

1. Importing Keras and TensorFlow libraries. Preparing the image dataset to be trained and accessed with the glob module in Python.
2. Imported the Image Data Generator module from Keras to rescale the RGB image to gray scale
3. Define our training set object with the name train datagen. All parameters and commands used are the same as those in the Keras documentation. then specify the folder directory where we prepare the image data, specify the target_size (image size) of 200x200, specify the Batch_size of 32, meaning that for 8000 images in the directory, it will be divided into several sizes with each containing 32 images (ada700/32 batches). Then specify the class_mode which is categorical.
4. Creating a sequential cnn model.
5. Adding the first convolution layer with a filter parameter of 16 feature detectors measuring 3×3 and adding relu for its activation function.
6. Defines the input shape (200, 200, 3) which resizes all images to 200 x 200 pixels and 3 arrays.
7. Determine max pooling with pool size (2)
8. Add a second convolution layer with filter parameters of 32 feature detectors measuring 3×3 with the same activation function, input shape, and max pooling as the first layer.
9. Adding 3 more convolution layers with filter parameters of 64 feature detectors measuring 3×3 with the same activation function, input shape, and max pooling as the first and second layers.
10. Added the flatten command to resize a matrix into a single vector.
11. Added danse command to perform fully connected process.
12. Added compile command to run existing ANN process in fully connected

13. The final rule for applying CNN training. In this step, we apply the created CNN model to the training_set. All parameters can be seen in the Keras documentation. The steps_per_epoch parameter is filled with the number of images in the training_set (700) divided by the number of files per epoch (32 files), so it is filled with 700/32. The epochs parameter is filled according to our wishes. Then save the training results in a file format.

Rules for Predicting Face Images

1. Loading facial images and loading the h5 model with the keras library to predict facial images with the results of the h5 model training
2. If the prediction results match the image, it will display the name of the image owner. If the prediction results do not match, it will display non-matching text.
3. Added eye blink detection feature to prevent cheating attendance using photos.

System Testing

System testing is the process of executing a program with the goal of finding errors. Software testing is considered effective if it can uncover errors that would otherwise remain undiscovered. The primary goal of system testing is to identify errors and software functions that do not meet development objectives by systematically uncovering these types of errors with minimal effort and time, also known as black box testing.

System Performance Verification

This is a reminder within the system that informs the user if there is an error in inputting or updating data, as well as confirmation if the user wishes to delete data. It also informs the user that actions such as saving, changing, or deleting data have been successfully performed. The system validation and verification displays include the following:

1. Incorrect NPK input information
   This message box displays information stating that the input in the NPK column must be numeric. The following shows an incorrect input in the NPK column.
2. Information input incorrect name
   This message box indicates that the name field must contain letters. The following shows an incorrectly entered name field.
3. Information has been saved data
   This message box displays information stating that the data, including the NPK, name, and facial image, has been saved and is ready for training. The following displays the saved data information.
4. Information has been done Absent
   This information is in the form of a message box stating that the user or employee has checked in. The following is a display of the training information image.

CLOSING

Conclusion

Based on the results of the analysis and design of the application program for implementing the deep learning method for the face detection attendance system at PT Karya Komponen Precise, the following conclusions can be drawn:

1. The application of deep learning algorithms with Convolutional networks (CNN) can be implemented in face detection attendance systems.
2. Deep learning algorithms with Convolutional networks (CNN) can recognize different facial images by means of prediction and classification.
3. Basic deep learning algorithms with Convolutional Networks (CNNs) cannot yet perform real-time facial recognition.

REFERENCES

Al Fatta, H. 2007. Information System Analysis and Design. Andi, Yogyakarta.

Alexander, LW, Sentinumo, SR, & Sambul, AM (2017). Implementation of a Facial Recognition Algorithm to Detect Visual Hacking. E-Journal of Informatics Engineering, Vol. 11, No. 1

Kristanto, A. (2018). Information System Design and Applications, Gava Media, Yogyakarta,

Munawar. (2018). Object-oriented system design analysis with UML (unified modeling language). Bandung: Informatika Bandung.

Putra, D. (2018). Digital Image Processing, Andi, Yogyakarta

Pawito. (2007). Qualitative Research. (A. Rahim (ed)). LkiS Yogyakarta

Primartha, R. (2018). Learning Machine Learning: Theory and Practice, Bandung: Informatics, Bandung

Santoso, Aditya. and Ariyanto, Gunawan, Implementation of Keras-Based Deep Learning for Face Recognition, Jurnal Emitter, ISSN 1411- 8890 Vol.18 No. 01

Santoso and Yulianto, Nugroho,. Design of an Attendance System Application Based on SMS Gateway, Jurnal Emitter, ISSN 1486-9590 Vol.18 No. 01

Sukamto, RA, & Shalahudin, M. ,(2013). Structured and Object-Oriented Software Engineering, Informatics, Bandung.

Wiryadinata, R., Sagita, R., Wardoyo, S., & Priswanto. (2016). Face Recognition On Attendance System Using Method of Dynamic Times Wrapping, Principal Component Analysis and Gabor Wavelet. Engineering Dynamics.

Wahyono, T. 2018 fundamentals of python for machine learning, Gava Media, Yogyakarta.

Novelan, MS, & Syahrizal. (2020). Expert System for Diagnosing Water Pump Machine Damage Using Case Based Reasoning Method. Informatics Journal: IT Development Journal, 5(1).

Novelan, MS (2020). Design of an Expert System for Rubber Plant Disease Diagnosis Using the Certainty Factor Method. Informatics Journal: IT Development Journal, 5(2).

Surianda, D., & Novelan, MS (2021). Implementation of the Naïve Bayes Method for Sentiment Analysis of Grab Indonesia Services. Digital Zone: Journal of Information and Communication Technology, 12(1), 1-10.

Novelan, MS, & Sari, RP (2021). Application of the Weighted Product (WP) Method to Determine Road Repair Priorities Based on Damage Classification. Kaputama Informatics Journal (JIK), 5(1), 128-135.

Sinaga, ASR, Novelan, MS, & Amin, M. (2020). Sentiment Analysis of Tokopedia App Reviews on Google Play Store Using the Naïve Bayes Algorithm. Journal of Information Technology, 4(2).

Novelan, MS (2021). Implementation of the Sugeno Fuzzy Method in a Decision Support System for Determining Bread Production Quantities. Informatics Journal: IT Development Journal, 6(1).

Putri, AR, & Novelan, MS (2021). Decision Support System for Motorcycle Spare Part Supplier Selection Using the Technique for Others Reference by Similarity to Ideal Solution (TOPSIS) Method. Kaputama Informatics Journal (JIK), 5(2), 241-248.

Novelan, MS, & Fachri, B. (2020). Implementation of the SAW (Simple Additive Weighting) Method in Determining Scholarships for Outstanding Students. Kaputama Informatics Journal (JIK), 4(1), 11-18.

Simanjuntak, MR, & Novelan, MS (2021). Implementation of the Weighted Product Method for a Decision Support System for Selecting the Best Employees. Kaputama Informatics Journal (JIK), 5(1), 108-115.

Novelan, MS, & Ramadhan, M. (2021). Expert System for Diagnosing Android Smartphone Damage Using the Forward Chaining Method. Kaputama Informatics Journal (JIK), 5(1), 84-91.