Implementation Of Neural Network

by Aditya Galih Sulaksono

Submission date: 13-Jun-2023 09:34AM (UTC+0700)
Submission ID: 2114916077
File name: Implementation_Of_Neural_Network.pdf (221.59K)
Word count: 3913
Character count: 20313

Implementation Of Neural Network And Canny Edge Detection To Recognize the Crime Through Surveillance Cameras

Erwien Tjipta Wijaya[1] and Aditya Galih Sulaksono[2]
{erwien@unmer.ac.id[1], adit@unmer.ac.id[2]}

[1,2]Departement of Information System, University of Merdeka, Terusan Dieng No 57-59, Malang, Indonesia

Abstract: The development of the smart city era in various countries is still offset by the magnitude of crime rates in every corner and downtown. Delivery of slow information to find out early crime will affect the precautions that should be taken, so as to require intelligent system-based tools. This research object is based in Jakarta Smart City (JSC) as an integrated command center with 5000 more surveillance cameras that have been installed at all strategic points in Jakarta. Surveillance cameras that have been fitted with intelligent system tools will send information to the JSC command center so that the category of criminal activity is immediately detected and immediate prevention. Among the analytical methods used are computational intelligence using Neural-network System, through video surveillance camera performed image processing using facial recognition techniques and body motion with canny edge detection algorithm segmentation. Measurement testing using a method of absolute percentage error (MAPE). So, the result level to the accuracy is 65%. Difficulty in the process of recognition because it is difficult to get the position of taking the corner of the suspect image and the victim. Implementation of intelligent system-based technology installed in surveillance cameras is expected to reduce crime rates.

Keywords: smart cities, surveillance cameras, neural-network system, image processing, canny edge detection, facial emotion, body gestures, recognition, MAPE.

1. Introduction

The most prominent requirement that is in big cities, for example, in Jakarta as the center of the capital of the country Indonesia has the highest crime rate. According to data from each local Police from 2014 to 2016, Jakarta was ranked number 1 with a total crime rate of 43.842 (Badan Pusat Statistik, 2017). Therefore it is necessary to suppress and reduce the crime, in addition to how to conduct community development should also be supported by the use of technology, where the task is to supervise the community in every corner and even in the middle of the city.

As proof of the city is said to be called Smart city is to apply the use of Information technology and the latest communications which are all equipped with the Internet of Things technology (Ijaz et al, 2016). Surveillance cameras can be enhanced in function and performance by implanting the intelligence system to recognize early crime acts.

The identification of criminal acts is done by reading the human behavior of the combination of facial and body movement of the perpetrator and the victim.

ICASI 2019, July 18-19, Banda Aceh, Indonesia

DOI 10.4108/eai.18-7-2019.2288590

To support this research is needed facial feature extraction process and body gestures feature extraction, through the stages - image processing stages will be obtained data that will be inputted into a learning process. The use of computational intelligence algorithm is needed to perform enhancement intelligence system so that human behavior can be read with clear and easy to do grouping.

The neural-network algorithm is learning system working like human brain generally. therefore, this algorithm is very suitable because it has learning process like generally human learn to understand other human body language. Henceforth the results given in the form of a probability that will be matched with the data table percentage of grouping based on crime clustering techniques.

Result approaching by threshold testing process converted to shape percentage. Henceforth percentage results are closest to the percentage table of the crime clustering techniques will be informed to the command centre which will then be carried out the type of precautionary action. The purpose of this study is to prevent and crack down on crime so that the crime action can be known early on which is expected to be able to suppress the amount of crime level especially in Jakarta.

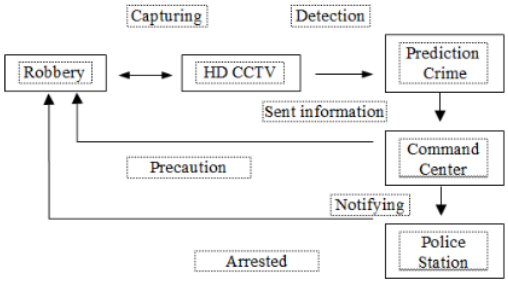

Fig. 1. Globally performance flow system.

The above description describes the flow of system performance in Jakarta Smart City, ranging from capturing the possibility of crime activities from surveillance cameras, then from the video capturing results are calculated estimates of the magnitude of crime, to further send information to the command centre with the support of internet of things to do analysis determine the right type of action, if categorized robbery without violence action will be warned through audio speakers and if categorized robbery with violence then immediately notified the nearest police office for immediate arrest.

2. Research Methodology

Data collection is done qualitatively by collecting recorded events that have been recorded by surveillance cameras. Although not all recording events can be done clustering process to be able to feature extraction group of events by giving the type of crime and give percentage of the magnitude of the risk of crime that will occur, therefore the grouping is done based on the type of crime, tools used, gender and age of the perpetrator or victim. This research is only directed to robbery incident data, which is often the case in Indonesia. The following crime clustering techniques are described in Table 1.

Table 1. Crime clustering techniques

CT	SS	SA	VS	VA	W
X	M	Yo	M	Yo	C
X	M	Yo	F	Yo	C
X	M	Mi	M	Mi	K
X	M	Mi	F	Mi	K
Y	M	Yo	M	Yo	K
Y	M	Yo	F	Yo	K
Z	M	Mi	F	Mi	B
Z	M	Ol	F	Mi	B
Z	F	Mi	M	Ol	B

CT = Crime Type, SS = Suspect Sex, SA = Suspect Age, VS = Victim Sex, VA = Victim Age, W= Weapon,
M = Male, F = Female, Yo = Young, Mi = Middle,
Ol = Old, C = Club, K = Knife, B = Bare hand,
X = Robbery With Violence, Y = Robbery Without Violence, Z = Robbery With Hipnotize

Table 1 explains that crime activities are divided into 3 sections: violent crime, non-violent crime and hypnosis. In this table, categorization is based on types of crime, sex of perpetrators of crime, age of perpetrators of crime and weapons of perpetrators of crime. Besides explaining the groups based on the victims, namely: the victim's sex and the age of the victim. In this type of crime robbery with hypnotize is very difficult to do identification of recognition through the emotions of faces and body movements of suspect and victim. Because robbery with hypnotize often occurs without any sign of violence either from suspect or victim. Therefore, this study only focuses on the type of crime robbery with violence and robbery without violence. In psychological science research that combined faces and body gestures can be a non-verbal human behaviour (Gunes, H., and Piccardi,M., 2007) for further all human activities recorded in surveillance cameras can be quickly identified the percentage threshold value of each motion and behavior, this will be explained in table 2.

Table 2. Crime threshold percentage

Type	Crime	Threshold (%)
A	Robbery without violence	61 - 79
B	Robbery with violence	80 – 100
C	Normal behavior	< 60

Table 2 illustrates the threshold limits of every facial movement and emotion that can be divided into 3 categories: violent robbery, nonviolent robbery or normal proceedings. To make it easier to categorize the types of criminal acts, it is necessary to value the percentage of crime thresholds that are useful for measuring each type of crime so that if in the identification process using intelligent system then touches the threshold that has been created, the system

will recommend the type of action to be taken warning or arrest action. In table 2 therein described various kinds of threshold boundary measurements where the threshold is used as a measurement tool of neural network learning process based on face feature extraction value and body feature extraction.

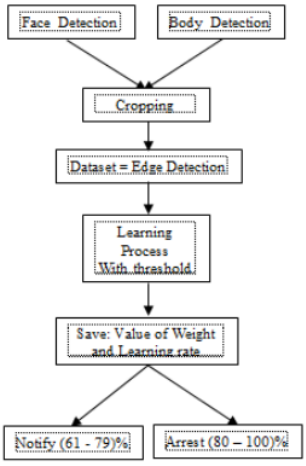

Fig. 2. Learning system.

In figure 2 described after face recognition and the body detection, the next step is to cropping the image and then done the edge detection process and then save the results of the dataset, then the dataset is done learning and the results are stored for use during the testing process.

2.1 Face and Body Detection Using OpenCV

OpenCV is a computer vision framework that is built for the purpose of detecting, especially on the face and also have developed to detect parts of the human body. This framework is written in C ++. OpenCV as an API can also be used and supports other programming languages such as Python, Java and Matlab / Octave, and not only that OpenCV can run on almost any operating system platform. Support the modules - a very complete module makes OpenCV is widely used by computer vision-based application developers.

One of the greatness of this framework is to detect faces with an accuracy of 90 to 95% on a image (Servin and Valentin, 2012). This study uses video capture media as an input that will be done emotional readings on the face and body area therefore to facilitate the process of detecting required an algorithm to solve such problems.

Viola Jones algorithm is the most widely used method for face recognition. This study emphasizes the use of Haar Cascade Classifier. The following is a script written using python language to access real-time face and eyes detection.

```
import cv2
import numpy as np

face=
cv2.CascadeClassifier('haarcascade_frontalface');
eye=cv2.CascadeClassifier('haarcascade_eye.xml');

ambil=cv2.VideoCapture(1)

While True:
        ret,gbr=ambil.read()
        abu2=cv2.cvtColor(gbr,cv2.COLOR_BGR2GRAY)
        wajah=face.detectMultiScale(abu2,1.3,5)

for (x,y,w,h) in wajah:
        cv2.rectangle(gbr,(x,y),(x+w,y+h), (255,0,0),2)
        roi_abu=abu2[y:y+h, x:x+w]
        roi_warna=gbr[y:y+h, x:x+w]
        mata=eye.detectMultiScale(roi_abu)
        for (ex,ey,ew,eh) in eyes:
                cv2.rectangle(roi_warna,(ex,ey),(ex+ew,ey+eh),(0,255,0),2)

cv2.imshow('img',gbr)
k=cv2.waitkey(30) & 0xff
if k==27:
        break

ambil.release()
cv2.destroyAllWindows()
```

The above script is used to detect faces and parts of the eye taken in real-time through the camera and display directly the results of the detection obtained. but there is still a lack of time to detect the eye, sometimes the mouth is considered part of the eye.
Meanwhile, to detect body parts can be using the scripts as follows.

```
import cv2
import numpy as np

body=cv2.CascadeClassifier('haarcascade_upperbody.xml')

ambil=cv2.VideoCapture(1)
while True:
ret,gbr=ambil.read()
        abu2=cv2.cvtColor(gbr,cv2.COLOR_BGR2GRAY)
        tubuh=body.detectMultiScale(abu2,1.3,5)
for (x,y,w,h) in tubuh:
```

```
    cv2.rectangle(gbr,(x,y),(x+w,y+h), (255,0,0),2)
cv2.imshow('img',gbr)
k=cv2.waitkey(30) & 0xff
if k==27:
        break

ambil.release()
cv2.destroyAllWindows()
```

This simple script above explains the algorithm flow to detect the upper body that can be recognized in real-time through cameras.

2.2 Facial Emotion Recognition

This research is emphasized on the emotions that most often arise when the occurrence of emotional misdeeds is anger and fear, where the moment of the suspect has anger facial emotion while victim has emotion faces fear (Ekman, P and Friesen, 2003).

Due to the many incidents of the perpetrator in committing criminal acts by using the mask it will be difficult to know the face recognition of angry emotions, for that angry emotion is not done the process of reading the face but through body gestures. These two main components will feature extraction using algorithms compiled using 3 stages: image processing stage, facial feature extraction stage or body gestures extraction stage, and emotion detection stage (Kim et al, 2015). In image processing requires effective techniques to support facial feature extraction, i.e. there are 3 areas of the face that can be done extraction consisting of - from the eye region, mouth region, and auxiliary region (Kim et al, 2015). The constraint faced in facial feature extraction is on the face capture, where the image captured produces an obscure image resulting from a low-resolution camera quality, hence to support optimum results in the image-extracting process requires a camera that has HD quality, thus the face image will be easily recognizable. After the face can be seen clearly the next step is the process of detection of the edge of the face image capture for the part to be done the pruning process is on the eye area and mouth area. Facial feature extraction is needed to identify the victim by capturing the emotion of the victim's feared face and identifying the evil faces of the perpetrator. The following is the data set and description of facial emotions described in table 3.

Table 3. Facial emotions recognized.

Type	Description
Anger	Catches are lowered and drawn together, lines appear between eyebrows, lower eyelids and may or may not be raised, eyelids over tense and may or may not be lowered due to eyebrow action, lips pressed firmly together with straight or bottom or open angles
Fear	Eyebrows are lifted and drawn together, forehead wrinkles drawn to the center, upper

	eyelid lifted and lower eyelid made, mouth open, lips slightly tense or stretch and pulled back

In table 3 describes the characteristics of face emotion recognition that is divided into 2 kinds of emotions of anger and fear. This data table still has many shortcomings, but at least this data can be representative in general to recognize human face emotions.

2.3 Body Gestures Recognition

The introduction the bodies of human emotional movement and posture to date is an unresolved problem in psychology research and non-verbal communication. Ambdy and Rosenthal have found a way to recognize human behaviour by combining face and body as a medium of communication so that it becomes a material assessment of human behaviour (Gunes and Piccardi, 2007). Generally, human behaviour also consists of 6 i.e. anxiety, anger, disgust, fear, happiness, and uncertainty (Gunes and Piccardi, 2007). In this study, not all that is used to recognize emotion bodily gestures, because just 2 of the most domains happen are anger to suspect and then fear for a victim. Before performing feature extraction, it is necessary to recognize emotion bodily gestures in order to facilitate later in recognizing human behaviour. Here is a collection of data and explanations of body movement.

Table 4. Bodies emotions recognized.

Type	Description
Anger	The body is elongated, the hands are at the waist, the hands are made into fist and are kept low sometimes holding weapons etc.
Fear	Body contracted, body support, high hands, trying to cover the body

Table 4 describes the characteristics of emotional recognition of the body, which is divided into two types of emotions, namely anger and fear, this explanation is made based on the results of psychological observations that have studied body language in every human being. in this observation still has many shortcomings because non-verbal language is very difficult to identify, but at least this data can be representative in general to recognize human face emotions.

2.3 Canny Edge Detection

This research uses canny detection algorithm. why use canny algorithm? The canny operator is the most optimal edge detection compared to other edge detection methods. This algorithm provides a low error rate, localizing the edge points. Because canny operators use the Gaussian Derivative Kernel to filter noise from the initial image soa as to get smooth edge detection results. Beside that canny algorithm has advantages (Tian-Shi Liu et al, 2016):

- Canny can perform marking of all edges corresponding to the selection of convolution parameters and is also very free to determine the desired edge thickness detection as expected.

- Canny produces the minimum distance between the edge detected with the actual edge.
- Canny easily detects edges and does not cause confusion when processing the next image.

Selection of canny edge detection parameters greatly affects the result of the ledges. Some of these parameters include :

- The standart value of gaussian deviation
- Threshold value

The following is the sequence figure of the canny edge detection described in general.

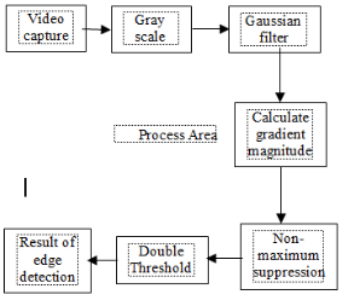

Fig. 3. Canny edge detection stage.

The figure above describes the process stages consisting of 1 input, 6 processes and 1 output. Input obtained from the video capture and produce the output of edge detection. Here is an explanation of the 6 processes :

- Grayscale :
  Techniques to do the gray process is required to be processed by computer, hence from that required gray scale image.

  $$Grayscale = \frac{R+G+B}{3}$$

- Gaussian Filter :
  Gaussian filter is used to minimize noise in the image.
  The gaussian 1D distribution equation is as follows:

  $$G(x) = \frac{1}{\sqrt{2\sigma}} e^{\frac{x^2}{2\sigma^2}}$$

  The gaussian 1D distribution equation is as follows:

  $$G(x,y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{x^2+y^2}{2\sigma^2}}$$

- Calculate gradient magnitude:
  The gradient magnitude operator is performed to calculate the amplitude edge detection where the pixel changes the gray level.

  $$|G| = \sqrt{G_x^2 + G_y^2}$$

$$G = \arctan(\frac{G_x}{G_y})$$

- Non-Maximum Suppression
  Omission of non-maximum values searched area found in stages gaussian filter and calculate gradient magnitude besides also eliminating values that are not maximum.
  $r = \alpha b + (1 - \alpha)a$

- Double Threshold
  The result of non-maximum is still not perfect, because the edge detection is still not good and there is some noise in the image. Therefore need to takes double threshold to overcome it. The way is to set two thresholds that is high and lowlimit.

2.4 Neural Network System

Artificial neural networks are methods that can duplicate the performance of the human brain. a network of small processing units modeled on the basis of the human nervous system. ANN is an adaptive system that can change its structure to solve problems based on external and internal information that flows through the network. Because of its adaptive nature, ANN is often called the adaptive network.

2.5 Preprocessing

The first step that needs to be done in the preprocessing stage is to collect the sample data in the form of recording of robbery events which is the result of the video camera surveillance videotape which is then analyzed by experts in the field of psychology. These experts identify the habits of facial emotion and gestures from both the offender and the victim. In the process of facial emotions and bodily gestures recognize these experts use the reference data described in Table 3. facial emotions recognized, while for list of the body emotions recognized are described in Table 4. the results of the analysis is clustering based on the type of crime, especially robbery. finally got 3 grouping, namely: robbery with violence, robbery without violence, and robbery with hypnotize. Because of the difficulty of identifying robbery with hypnotize, this study emphasizes robbery with violence and robbery without violence. The explanation of the results of crime clustering techniques is given in Table 1. And then weighted by the percentage of crime thresholds that aims to facilitate the identification of the type of crime, making it easier to determine the type of action or prevention described in Table 2.
After the qualitative data is obtained, the next step is to perform image processing. there are 2 stages:

- Training Process Stage
  Video capture after the segmentation process, next is to store attribute information into the database for training. Training process is to change the attribute of pixel data into the numbers - input numbers (x) which will be multiplied by weighting and then sent to learning algorithm to calculate the activation function. The learning process will stop when the epoch has been reached or the error threshold value does not exceed the desired value. Then the result of the activation function then saves the latest weight to the database which will be used as summing with the weights for the image input from video capture during testing process. The job diagram description is described in general in Figure 4.

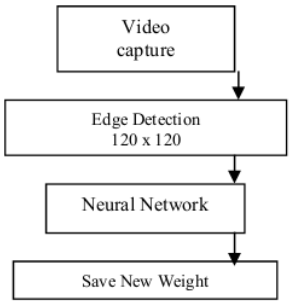

Fig.4. Preprocessing workflow.

- Testing Process Stage
  Image is obtained in real time from video capture which then done segmentation process. The result will be in pixel data form as input (x) then multiplied by the weights stored by using the following derivative.
  $y\_in_k = w_{oj} + \sum_{k=1}^{p} z_j v_{jk}$

  $y_j = f(y_{-in_k}) = \frac{1}{1+e^{-y\_in_k}}$

Once the data is available then the preprocessing step is as follows:

- Backpropagation Neural network is a multi-layers network; backpropagation algorithm is generalization of delta rule (Windrow-Hoff) that is applying gradient descent method to minimize total squared error from output calculated by network. In this study using 14,400 units of input $(x_1, x_2, x_3, \ldots, x_n)$, 50 unit of hidden layers $(z_1, z_2, z_3, \ldots, z_n)$and for output of 1 unit of output (eye region, mouth region, hand, torso, shoulder, and head). the neural network architecture can be seen in Figure 5.

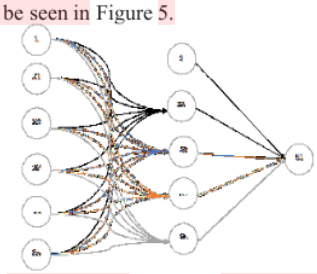

Fig. 5. Neural-networks architecture

2.5 Processing

At this stage the whole process is merged, here are the steps:

- Image Acquisition, is the stage of shooting through video.
- Preprocessing, is doing the Grays calling process to change the image that has the color into a picture that has the level to gray.

- Segmentation divides the image into its respective intensity area so that it can distinguish between object and background or called edge detection. The Canny Algorithm Detection explanation has been described above.
- After done edge detection, the data in the form of pixel value continued to neural-network. This research uses backpropagation algorithm for identification process. Why using backpropagation? because this algorithm improves weaknesses in the previous ANN, with features to minimize errors by adjusting weights based on the desired output and target differences. There are three main steps: input retrieval, error correction, and weight adjustment.
- The next step calculates the magnitude to the accuracy using the Mean Absolute Percentage Error (MAPE) approach.
  $MAPE = \frac{1}{n}\sum|A - F|$

  Description :
  A : Actual
  F : Forecast

3. Conclusions

In this paper there are still many shortcomings in the delivery from the process to the results. But this paper represents the research that has been done. In addition to the above-mentioned deficiencies. The results of this study still have a 65% accuracy rate in the introduction of crime acts. This weakness occurs because of the difficulty to get capturing image angle position a suspect and a victim. So, it still needs a lot of refinement, especially on the quality of surveillance cameras and image processing technic.

Acknowledgements

The authors express their gratitude to the Jakarta Smart City especially as a place of research, as well as to the college institution of Merdeka University of Malang as a place to work and gain knowledge thanks to its support so far and related parties that are not mentioned by name one by one.

References

[1] Badan Pusat Statistik (2017) 'Criminal Statistic'. Available at: https://www.bps.go.id/publication/2017/12/22/197562b7ad0ced87c08fada5/statistik-kriminal-2017.html. 2017.

[2] Ekman, P and Friesen, W. V. (2003) *Unmasking The Face, A guide to recognizing emotions from facial expressions,*. Cambridge MA: Malor Books.

[3] Gunes, H., and Piccardi,M. (2007) 'Bi-Modal Emotion Recognition From Expressive Face And Body Gestures', *Journal of Network and Computer Applications*, 30(4).

[4] Kim, M.W and Hoon, Y.,Joo, and Park, J.B (2015) '"Emotion Detection Algorithm Using Frontal Face Image", ICCAS2005, June 2-5, KINTEX, Gyeonggi-Do, Korea,'. Available at: http://www.researchgate.net/publication/228870902.

[5] Servin, E and Valentin, P.S. (2012) 'Facial Recognition Using OpenCV', *Journal of Mobile, Embedded and Distributed Systems*, IV(1).

[6] Sidra ijaz, Munam Ali Shah, Abid khan, and M. A. (2016) 'Smart Cities: A Survey on Security Concerns', *International Journal of Advanced Computer Science and*

Applications, 7(2).

[7] Tian-Shi Liu, Rui-XiangLiu, Ping-Zeng, and Shao-Wei Pan (2016) 'Improved Canny Algorithm for Edge Detection of Core Image.', *The Open Automation and Control Systems Journal*, 6.

Implementation Of Neural Network

ORIGINALITY REPORT

11%	6%	5%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to SIM Global Education Student Paper	1%
2	Submitted to Kingston University Student Paper	1%
3	repository.usd.ac.id Internet Source	1%
4	Submitted to Colorado State University, Global Campus Student Paper	1%
5	"Emotion Recognition", Wiley, 2015 Publication	1%
6	Ulfi Andayani, Ema Budhiarti Nababan, Baihaqi Siregar, Muhammad Anggia Muchtar, Tigor Hamonangan Nasution, Ikhsan Siregar. "Optimization backpropagation algorithm based on Nguyen-Widrom adaptive weight and adaptive learning rate", 2017 4th International Conference on Industrial Engineering and Applications (ICIEA), 2017 Publication	1%

7	Submitted to University of Nottingham Student Paper	1 %
8	blog.csdn.net Internet Source	1 %
9	Submitted to Universitas Merdeka Malang Student Paper	<1 %
10	is.muni.cz Internet Source	<1 %
11	Submitted to Universitas Gunadarma Student Paper	<1 %
12	M. Ridho Vernanda Padantyo, I. Gede Pasek Suta Wijaya, Fitri Bimantoro, Rina Lestari. "COVID-19 prediction based on GLCM features of radiography image using SVM, KNN, and backpropagation ANN classifiers", AIP Publishing, 2023 Publication	<1 %
13	Submitted to Universitas Brawijaya Student Paper	<1 %
14	Hui Li, Wenyu Wang, Xintong Wang, Qingkai Han, Jinguo Liu, Zhaoye Qin, Jian Xiong, Zhongwei Guan. "A nonlinear analytical model of composite plate structure with an MRE function layer considering internal magnetic and temperature fields", Composites Science and Technology, 2020	<1 %

Publication

15	Khem Puthea, Rudy Hartanto, Risanuri Hidayat. "The Attendance Marking System based on Eigenface Recognition using OpenCV and Python", Journal of Physics: Conference Series, 2020 Publication	<1 %
16	H. Gunes, M. Piccardi. "A Bimodal Face and Body Gesture Database for Automatic Analysis of Human Nonverbal Affective Behavior", 18th International Conference on Pattern Recognition (ICPR'06), 2006 Publication	<1 %
17	Nila Armelia Windasari, Halim Budi Santoso, Jyun-Cheng Wang. "Memorable Digital Tourism Experience: Utilization of Emotions and Sensory Stimuli With Service-Dominant Logic", Emerald, 2022 Publication	<1 %
18	edoc.pub Internet Source	<1 %
19	benthamopen.com Internet Source	<1 %
20	prosiding.unirow.ac.id Internet Source	<1 %
21	www.slideshare.net Internet Source	<1 %

22	Juan Jorge Schäffer. "Linear differential equations and functional analysis. VI", Mathematische Annalen, 08/1962 Publication	<1%
23	Priyadarshini D. Kalwad, Suvarna G. Kanakaraddi, Ashok K. Chikaraddi, T. Preeti, Karuna C. Gull. "Chapter 26 XCEPTION: Facial Expression Detection Using Deep Learning Techniques", Springer Science and Business Media LLC, 2022 Publication	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

Implementation Of Neural Network

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12