

rabayu_2019_IOP_Conf._Ser._
Mater._Sci._Eng._619_012021.p
df
by

FILE	RABAYU_2019_IOP_CONF._SER._MATER._SCI._ENG._619_012021.PDF (1.03M)		
TIME SUBMITTED	20-JAN-2020 09:04PM (UTC+0700)	WORD COUNT	4130
SUBMISSION ID	1244041474	CHARACTER COUNT	20839

PAPER · OPEN ACCESS

Car Driver Drowsiness Recognition Android-Based System

1
To cite this article: Indrabayu *et al* 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **619** 012021

View the [article online](#) for updates and enhancements.

Car Driver Drowsiness Recognition Android-Based System

Indrabayu¹, Siti Khumaerah Mufti², Intan Sari Areni³

^{1,2,4} Department of Informatic Engineering, Faculty of Engineering, Hasanuddin University, Indonesia

³ Department of Electrical Engineering Faculty of Engineering, Hasanuddin University, Indonesia

Email: indrabayu@unhas.ac.id, khumufti@gmail.com, intan@unhas.ac.id

Abstract. — The growth of the number of vehicles in Indonesia especially in terrestrial mode is also offset by the increasing number of accidents. One of the main causes of traffic accidents is due to drivers who driving in drowsiness. This condition leads to loss of concentration and even awareness of the driver. The drowsiness recognition system built into android smartphone platforms which owned by almost all walks of life, can help the community and government in providing direct warning especially to motorists while driving in a drowsy state. The method used is the Haar-Cascade Classifier to detect the driver's face and utilize the detectable face size ratio to create ROI in the mouth area. Then the using of the FindContour algorithm to find contours in dark areas when the mouth is open because it evaporates. The size of the contour will be compared to the Mouth ROI size for the drowsy driver's decision. System creation using Android Studio with C ++ and Java programming languages, as well as OpenCV libraries for Haar-Cascade and FindContour. Based on the test to 10 respondents, the accuracy of the system recognizes sleepiness on the driver reached 88.7% during the day.

1. Introduction

Current transportation means has become a major need for human beings especially in Indonesia. According to data from the Central Bureau of Statistics from 2003 to 2011, there has been a surge in the number of motor vehicles, both passenger cars, buses and trucks reaching 12-15% per year, and is forecast to continue increasing [1]. However, the number of accidents occurring in land transportation in Indonesia is quite high. According to data from the Police Force of the Republic of Indonesia, tiredness and drowsiness are the main causes of traffic accidents in 2016 which until July reached 41% [2]. Specifically, accidents caused by drivers are not ready in response to traffic and sleepy situations are closely related to the driver's vision. This condition not only endangers drivers but also other road users. Therefore, it is advisable to rest for a while when getting sleepy. However, there are still many drivers who ignore the safety protocol, one of them because of the lack of direct warning [3]. Several studies have provided solutions by creating a detection system that can recognize drowsiness by image processing and machine learning using a camera mounted on the steering wheel, as well as a car dashboard. Therefore, it takes a more practical and affordable system, which can recognize drowsiness and can alert the driver before it actually loses concentration. One of them is, a system built in the smartphone android platform.

Smartphones are almost owned by all walks of life. Utilizing a smartphone camera to detect and recognize sleepiness in the driver can be an alternative to help provide immediate alert when the driver is sleepy. Therefore, in this study, made an android-based system that can recognize sleepiness in car drivers. Using smartphone's camera to record the driver's face in real-time. The system will then process image data using Haar-Cascade Classifier to detect faces. Then the FindContour algorithm takes advantage of contour sizes when the mouth opens because it evaporates to recognize sleepiness in the driver.

2. Parameters to detect drowsiness on the driver

Drowsiness is one of the symptoms caused by fatigue that reduces strength, reaction speed, coordination ability, balance and also the ability of a person in making decisions. Therefore, when drowsy, it is not recommended to perform activities that require high concentrations with dangerous consequences such as driving [4]. Please follow these instructions as carefully as possible so all articles within a conference have the same style to the title page. This paragraph follows a section title so it should not be indented. According to Narender Kumar, the driver's sleepiness condition can be seen from the psychological condition, behaviour, and visual appearance of the driver. Furthermore, to make the driver drowsiness system can be designed with techniques to analyse visual appearance such as eyes and mouth [6]. One of them can be done based on the condition of the mouth when yawning.

The yawning mouth can be used as a parameter to detect sleepiness in the driver. One of the studies conducted by Kumar and Barwar in 2014 that decides someone is yawning or not by checking the size of the mouth. If the height of the mouth is greater than a certain threshold, it means that someone is yawning, a mistake then appears when the system fails to distinguish the yawning condition by laughing, talking and singing [7]. According to the National Sleep Foundation basically humans yawning under the following conditions [8]:

- Drowsiness
- Seeing others yawn
- Bored
- Disease
- The brains working hardly
- Body is having less movement

However, under the normal conditions humans will yawn for about six seconds [9].

3. System development

The aim of this work is to recognize sleepiness or drowsiness in driver's yawn to give the driver a direct alert.

3.1. Capturing driver image

First step is to record driver's face using smartphone front camera. The shooting angle is set between -30 to 30 degrees, with 30 frame per second, mp4 format for video, and RGB format for image frame. This will be the input for the system development.

3.2. Face detection

Face on the recording data is detected using Haar Cascade Classifier. A complex algorithm to detect faces on image using a square feature that gives a specific indication on a picture or image, called Haar-Like Feature. The image that is processed is the image in the grayscale form. So, the RGB format will be converted to grayscale first. The grayscale frame will be detected using the haar-like feature that traces the image using a box-shaped sliding window. After the sliding window ends, the box size will be minimized and sliding back through the image. And so on until box size cannot be minimized anymore. So it will be a lot of repetitive detection results.



Figure 1. Illustration of the grayscale frame with red-box sliding windows with haar-like feature inside.



Figure 2. Illustration comparison of detection results with modified minNeighbours: (a) value 0; (b) value 3; (c) value 8.

System creation uses modified value of minNeighbours to 8, to minimize the possibility of the system detecting facial-like features on non-face objects. By modifying minNeighbours to 8, then only features (or surroundings) with at least eight times detection will be classified as a strong classifier as shown on Figure 2.

3.3. Creating MouthROI

The next stage is to determine the ROI (Range of Interest) on the mouth of the face. By determining the ROI, it can be localized to certain parts that are desired to be processed, so it can be more targeted and accurate. ROI is set at the mouth position, the center-bottom of the face. ROI size is also set as not too big and not too fit on the size of the mouth to anticipate the size of the mouth when it opens. MouthROI determined MouthROI using these following stages:

3.3.1. Creating FaceROI. A previously detected face, converted to ROI under the name FaceROI which will be the place of the process to determine MouthROI by utilizing coordinate value x and y as well as the height (h) and width (w) from FaceROI.

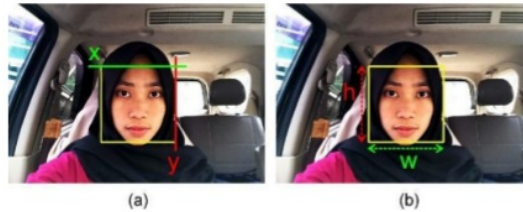


Figure 3. (a) Illustration of FaceROI coordinate; (b) Illustration of FaceROI size

3.3.2. Specifying the coordinate and determining the size of MouthROI. Coordinates will determine the position of MouthROI in FaceROI. The desired coordinates are shown in illustration figure 3. Coordinate of MouthROI area are marked with red dots and named (x', y') . To find the value of (x', y') using following equation:

$$x' = \left(r_x + \frac{1}{3} w_r \right) \tag{1}$$

$$y' = \left(r_y + 1.3 h_r \right) \tag{2}$$

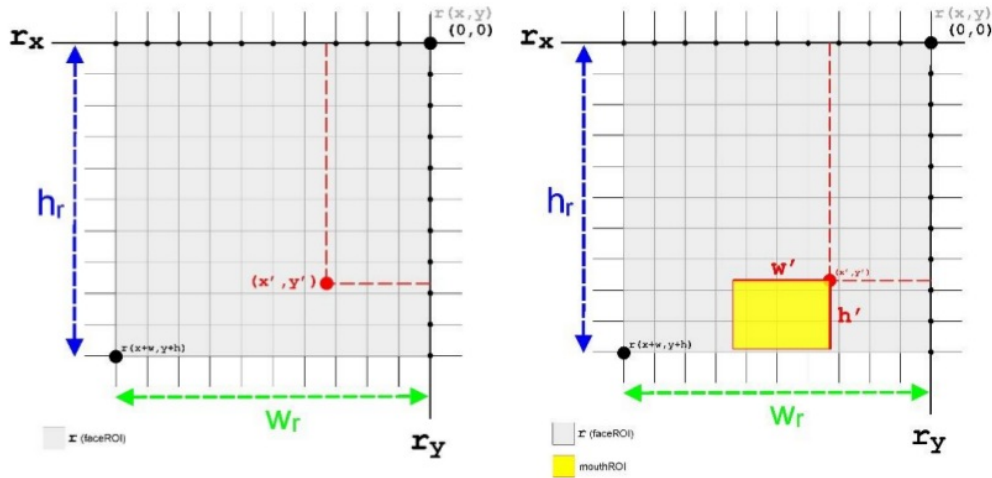


Figure 4. Illustration of desired coordinates of MouthROI within the FaceROI.

Figure 5. Illustration of desired size of MouthROI

Next, the desired size of MouthROI is a rectangle consisting of a height of (h') and width (w'). The desired size is displayed using a yellow rectangle in figure 5. To find the size of MouthROI using the following equation:

$$h' = \frac{1}{4.5} h_r \tag{3}$$

$$w' = \frac{1}{3.18} w_r \tag{4}$$

And then the MouthROI we had made, will displayed using red bounding box.

3.4 Finding contour

The next stage is the contour finding for the dark areas contained in the area of MouthROI. The contour will be used for decision making process. The flow of the contour search process is shown in Figure 8 below:

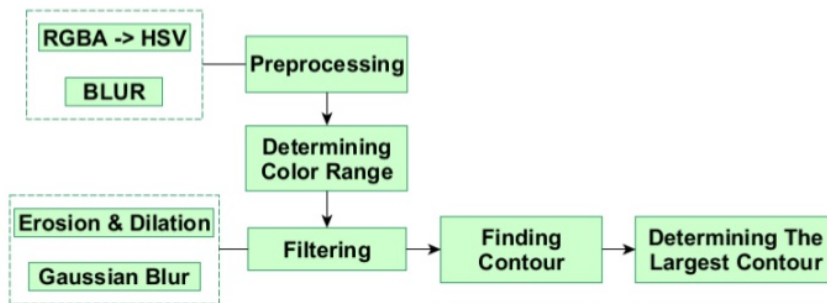





Figure 6. Contours finding process flow.

The process flow described as follows:

- (1) *Pre-processing*. Stage to prepare the image for further process. The following steps in pre-processing are shown in table 1.

Table 1. Pre-processing stages

Proses	Illustration	Description
Image input from MouthROI		7 Image in RGB format
1 Converting RGB image to HSV		Image is converted into HSV format, because HSV is better at distinguishing colours (better than RGB).
2 Blur (size kernel 4x4)		Blur used to smooth out the border and reduce the noise on the image

(2) *Determining colour range.* Colour range is five-collected colours that are considered to represent a variety of colours in dark areas when the mouth is open. The colour value of HSV in OpenCV is [0-180, 0,255, 0,255]. To convert the initial value of HSV to be compatible with OpenCV, use the following equations:

$$H_{OpenCV} = Hue \times 0,5 \quad (5)$$

$$S_{OpenCV} = Saturation \times 2,55 \quad (6)$$

$$V_{OpenCV} = Value \times 2,55 \quad (7)$$

Table 2. Colour conversion table

Colour (in RGB)	Hex Code	RGB	HSV	HSV-OpenCV
A	#000000	[0, 0, 0]	[0, 0, 0]	[0, 0, 0]
B	#374A2B	[55, 74, 43]	[96, 42, 29]	[48, 207, 74]
C	#290001	[41, 0, 1]	[360, 100, 16]	[180, 255, 40]
D	#754447	[117, 68, 71]	[356, 42, 46]	[178, 107, 117]
E	#BF7575	[191, 117, 117]	[0, 39, 75]	[0, 99, 191]

Furthermore the color will be defined using inRange function of OpenCV. One colour consists of one edge of the upper threshold and one edge of lower threshold.

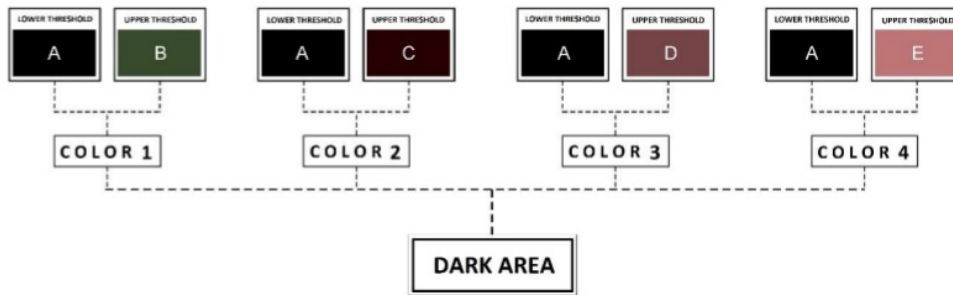


Figure 7. Illustration of colour definition for colour range of dark area.

(3) *Filtering.* This stage aims to improve the image quality. There are two types of process shown in Table 3.

Table 3. Filtering process illustration

Proses	Illustration	Description
Input for filter process is the image that contains colours that has been defined		Filtered image in binary form
1. Dilation (5 iterations) Erosion (10 iterations)		Also knows as closing operation that useful to smooth the image and remove small holes (noise)
2. <i>Gaussian Blur</i> (kernel 9x9)		Gaussian blur operation is used to smooth the image. Compared to other blur techniques, Gaussian blur provides the most tolerant output for colour differences.

(4) *Contour finding.* Contour finding based on pre-defined colours that has been filtered before, so our objects have less noise and smoother edges. Using the FindContour function from OpenCV, using CV_RETR_External mode which only looks for the outer contours of the object and CV_CHAIN_APPROX_SIMPLE method, the process will be using minimal computing and limit energy sources.

(5) *Determining the largest contour.* Inconsistencies of light and conditions around *MouthROI* can cause a color discoloration and cause a shadow on a particular part of the image. The shadow can add dark areas to the mouth, so that miscalculation can affect decision making. For that of all the contour areas that have been detected, will be selected one of the largest.

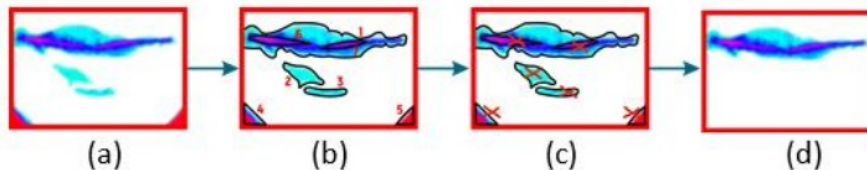


Figure 8. Illustration of determining the largest contours process. (a) the contours detected; (b) there are 7 contour detected; (c) the entire contours are compared to each other and the small ones are eliminated; (d) one of the largest contours saved as the output

The output contour will be given a red bounding box. Both the contour and the bounding box will be displayed on the system view.



Figure 9. Screenshot of system view. The green line is the contour and the red-rectangle around the contour is bounding box (bb) contour.

4. Decision making

The decision-making process is the process of determining the condition of the driver based on predefined parameters. Basically decision making is divided into 2 namely: open mouth decision; and drowsy driver. In decision-making using the comparison between MouthROI and BB contours is illustrated in figure 10.



Figure 10. Illustration of comparison instruments for decision making. (The initial shown on the picture, will be used for the rest of the paper).

4.1. Pre-experiment

Before setting the parameter for decision making, there are several pre-experiments to perform the basis form for determining the way to placing smartphone camera and decision-making parameter. Through the system that can processing 10 frames per second, there are three pre-experiment described as follows:

- (1) *Limit Angle of Shooting Tolerance*: The first test is to determine the tolerance of the proper shooting angle, the camera position from the driver's face. Tests were performed on two video data taken at different angles. To get the average value (Nt), the equation is:

$$Nt = \frac{Na}{Nb} \times 100 \quad (8)$$

where Na is the total number of frames with the face detected, and Nb is the total number of frames in the video. So the results of the test are shown in Table 4 as follows:

Table 4. Result of tolerance angle of shooting for the best picture search testing

Vid	Angle	Duration (Seconds)	Number of Frames with detectable face (a)	Average
1	<30°	15	131	87%
2	> 30°	15	88	58%

Based on Table 4, with a shooting angle of more than 30 degrees, the system tends to fail to detect the driver's face. So the tolerance of shooting angle used is -30 to 30 degrees.

(2) *Threshold Mortality Size Limit.* The test is performed to obtain a size comparison to differentiate the yawning condition with other open mouth conditions. Research conducted by Kumar and Barwar in 2014 decided the parameters of a person to yawn can be seen from the size of the mouth height is greater than a certain threshold. However, errors occur when the system fails to distinguish the yawning condition with the condition when the respondent laughs, speaks and sings [7]. Therefore, for pre-experimental testing used the comparison of some conditions:

- **A** is the condition when the driver speaks and the contour is detected
- **B** is the condition when the driver evaporates and the contour is detected
- **C** is the condition when the driver laughs and the contour is detected
- **D** is the condition when the driver sings and contours Detected

Tests were performed on three video data with the primary goal of obtaining the threshold with the minimum possible size, but with a high tolerance level to distinguish the condition of yawning with others. Therefore that compared not only high-size, but also volume. The estimated threshold used is shown in figure 11, the orange shading indicates the range of sizes that are the parameters of whether a threshold is grouped as **X**, **Y**, or **Z**.

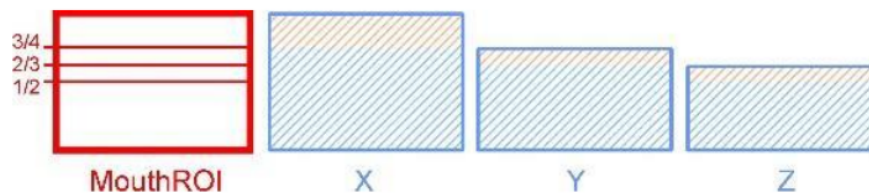


Figure 11. Illustration of size **m** (red box) and **b** (blue box)

- **X threshold** occurs if and only if the frame detected size **b** is greater than or equal to 3/4 size **m**
- **Y threshold** occurs if and only if the frame detected size **b** greater than or equal to 2/3 the size of **m**, but smaller than 3/4 size **b**
- **The Z threshold** occurs if and only if the frame detected size **b** is greater than or equal to 1/2 size **m**, but less than 2 / 3 sizes of **b**

the test results of three video data testing shown in table 5 below:

Table 5. Best size of threshold testing result

Vid	Condition	Seconds	Frame Detected					
			Volume Threshold			High Threshold		
			X	Y	Z	X	Y	Z
1	A	3	0	0	2	0	0	2
	C	5	0	1	4	0	0	2
	B	8	1	3	32	12	23	31
2	B	8	0	3	6	22	22	12
	C	10	0	0	3	0	0	13
3	C	5	0	0	2	1	0	5
	D	6	1	0	3	0	0	1

Based on Table 5 it can be concluded that the evaporative condition (**b**) is most detected at the **Z** threshold.

- (3) *Best Frame Limit Threshold*: The best frame number threshold search is done to distinguish yawning from sleepiness with other conditions. After testing of 3 video data then got the result as follows:

Table 6. Best frame threshold searching test result.

Vid	Condition	Sec	Total Frame		Average (%)
			Volume	High	
1	B	7	29	58	43
2	B	7	9	53	30.5
3	B	7	21	27	30
AVERAGE (%)					34.5

So the best value as a parameter for decision making is the detection of open mouth 35 times on the frame for 7 seconds (35 of 70 frames).

4.2. Decision taking

Using three pre-experiment results, decision taking was created and described as follows:

- (1) *Open mouth detection*. Open mouth detection will detect condition that approach the condition when mouth is opened because of yawning. The size of **m** and **b** will be compared. If the volume **b** or height **b** is greater than half of the volume or height of **m**, it will be defined as a mouth open. Every single mouth open is defined as one counter and starting the counter.
- (2) *Drowsiness recognition*. Basically humans will yawning for about 6 seconds [9], that's why the timer set for 7 second that will be used to count the duration of the mouth open during yawning. System could process 10 frames per second, it means the 7-seconds timer is equal to 70 frames. In this section, system will recognize sleepiness or drowsiness in driver's yawn – as National Sleep Foundation has claimed that sleepy is not the only reason of human's yawning, it could means bored, disease, imitating, hot brain, and body less movement [8] –. When counter (mouth open detected) in 35 frames or more as long as the 7-seconds timer starts hitting 5 until 7, the alert will ring.

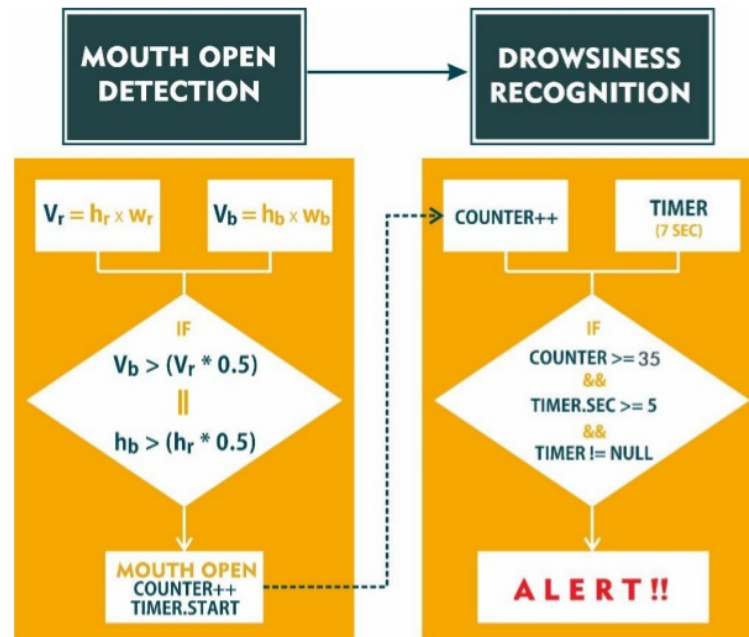


Figure 12. The flow of decision-taking system



Figure 13. Screenshot of system detection result: (a) normal condition; (b) mouth open detected; (c) drowsiness recognised.

5. Experimental results

The performance of drowsiness recognition of this system has been measured using 10 respondents in 3al-time in the middle of the day. The smartphone was installed on car dashboard under 30 degrees from driver's face. To determine the accuracy of the system, data measured using confusion matrix variables.

Table 7. Confusion matrix

	Result of Detection (Drowsiness)	Result of Detection (Not Drowsy)
Actual Condition (Drowsy)	True Positive (TP)	False Negative (FN)
Actual Condition (Not Drowsy)	False Positive (FP)	True Negative (TN)

Table 7 is described as follows:

- TP = The actual condition is drowsy and there is a warning by the system.
- FN = The actual condition is sleepy and there is no warning by the system.
- FP = The actual condition is not sleepy and there is a warning by the system.
- TN = The actual condition is not sleepy and there is no warning by the system.

Therefore, to determine the level of accuracy (A) of drowsy detection, the equation follow is used:

$$A = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (9)$$

Screenshots of system measured using confusion matrix through 10 respondents are presented in pictures below:



Figure 14. True positive (TP); respondents is yawning and alerts appear



Figure 15. True negative (TN); condition when driver is not yawning and no alert shown. Mouth open conditions like laughing, singing, and speaking are grouping in this true negative situation.



Figure 16. False positive (FP) system error and fails to alerts driver while yawning.



Figure 17. False negative (FN) condition when respondents yawn but no alert shown. Error detection could cause by Haar-Cascade fails to detect face because there are things on driver's face, or respondents yawn under 5 seconds.

6. Conclusions

Utilizing *Range of Interest* (ROI) after Haar-Cascade algorithm had detected face, could elevate the system for further feature-face detection, as well as the mouth detection. The dark feature that made

when mouth is open could be taking into account using FindContour algorithm to track the movement of the dark areas in mouth. If the contour size was bigger than half of ROI size, it will be considered as an *open* mouth. To recognize the sleepiness in it, the number of frame is measured that if in 5-7 seconds, system detected at least 35 frames contained *open* mouth, it will be decided as yawning and notification will appear. The result reaches 88.7% accuracy.

References

- [1] (2011) Central Bureau of Statistics. [On line]. Available: <https://www.bps.go.id/>
- [2] (2016) Victim of Accidents [Online]. Available: www.menyusurijalan.com/
- [3] Kusyono, "Several Ergonomic Factors Associated with Low Back Pain Complaints at Angkot Drivers Department Gunungsari-Celancang (PP) Cirebon," thesis, UNDIP, Bandung, 2012.
- [4] Anriani, Rina, "Recognition of Human Face Features for Drowsy Driver Detection," Hasanuddin University, Makassar, 2017.
- [5] (2013) Kompas Cyber Media [Online]. Available: <https://lifestyle.kompas.com/>
- [6] Kaplan, S., Guvensan, MA, Yavuz, AG, Karalurt, Y., "Driver Behavior Analysis for Safe Driving: A Survey," IEEE Trans. Intell. Transp. Syst. 16, 3017-3032. doi: 10.1109 / TITS.2015.2462084, 2015.
- [7] Kumar, Narendra, Dr. NC Barwar, "Detection of Eye Blinking and Yawning for Monitoring Driver's Drowsiness in Real Time," International Journal of Application or Innovation in Engineering and Management (IJAIEM), Vol 3, Issue 11, Nov. 2014.
- [8] (2016) National Sleep Foundation [Online]. Available: www.sleep.org/
- [9] Underwood, Emily. (2016) Science Magazine: The Bigger Your Brain, The Longer You Yawn. [On line]. Available: www.sciencemag.org/
- [10] Turiyanto MD., Djoko P., Dikairono R., "Implementation of Facial Recognition Technique Based on LBP Feature on Food Introduction Robot," thesis, Surabaya Institute of Technology, Surabaya.

Acknowledgments

I would like to express our deepest appreciation to all those who provided us the possibility and assistance. Also, for all of our respondents who had spent their precious time to help us conducted this research.

ORIGINALITY REPORT

%**5**

SIMILARITY INDEX

%**4**

INTERNET SOURCES

%**4**

PUBLICATIONS

%**3**

STUDENT PAPERS

PRIMARY SOURCES

1

elar.usfeu.ru

Internet Source

%**2**

2

irep.iium.edu.my

Internet Source

%**1**

3

Submitted to University Tun Hussein Onn
Malaysia

Student Paper

%**1**

4

M P Hatta, S S Yuni. "Changes in rainfall and discharge in Saddang Watershed, South Sulawesi, Indonesia", IOP Conference Series: Earth and Environmental Science, 2019

Publication

<%**1**

5

Gulraiz Khan, Saira Jabeen, Muhammad Zeeshan Khan, Muhammad Usman Ghani Khan, Razi Iqbal. "Blockchain-enabled deep semantic video-to-video summarization for IoT devices", Computers & Electrical Engineering, 2020

Publication

<%**1**

6

Internet Source

<% 1

7

"Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2018 (ISMAC-CVB)", Springer Science and Business Media LLC, 2019

Publication

<% 1

EXCLUDE QUOTES ON

EXCLUDE BIBLIOGRAPHY ON

EXCLUDE MATCHES < 5 WORDS