Science in Information Technology Letters
Vol. 2., No. 1, May 2021, pp. 43-53
ISSN 2722-4139
http://pubs2.ascee.org/index.php/sitech

Image processing for student emotion monitoring based on fisherface method

A H Pratomo [a,1,*], M Y Florestyanto [a], Y I Sania [a], B Ihsan [b], H H Triharminto [b]

[a] Jurusan Informatika, Fakultas Teknik Industri, Universitas Pembangunan Nasional Veteran Yogyakarta
[b] Department of Electrical Engineering, Indonesian Air Force Academy
[1] awang@upnyk.ac.id
* Corresponding Author

ARTICLE INFO

Article history
Received March 12, 2021
Revised April 20, 2021
Accepted May 15, 2021

Keywords
Image processing
Facial recognition
Fisherface

ABSTRACT

Monitoring academic emotion is an activity to provide information from students' academic emotions in the class continuously. Some research in the image processing field had done for face recognition but had not been many studies on image processing to detect student emotions. This paper aims to determine the percentage of facial recognition with fisherface and academic emotional recognition by monitoring changes in students' facial expressions using facial landmarks in various distances, camera angles, light, and attributes used on objects. The proposed method uses facial image extraction based on fisherface method for presence. Furthermore, face identification will be made with Euclidean distance by finding the smallest length of training data with test data. Emotion detection is done by facial landmarks and mathematical calculations to detect drowsiness, focus, and not focus on the face. Restful web service is used as a communication architecture to integrate data. The success rate of applications with the fisherface method obtains 96% percent accuracy of face recognition. Meanwhile, facial landmarks and mathematical calculations are used to detect emo tions, with 84 %.

1. Introduction

Academic emotion is one of the success factors for students learning in class. The emotions are recognized by looking at student's expressions in the class where change over time [1]. Emotional monitoring is done by detecting students' faces and then identifying them. After the identification is made, the process continues with emotion detection using a webcam as a sensor device for capturing the image. The image is then identified based on image extraction, and facial landmarks are searched to detect emotions. Conducted emotion detection using a face-based template based on fuzzy classifier [2]. The other method for emotion detection is based on facial landmarks to determine changes in facial expressions.

Face detection is performed to determine the face object [3]. One way to detect faces is to detect skin color. Research related to skin color detection was carried out by [4] using RGB, HSV, and YCbCr color comparisons to distinguish skin color from the background. The development of the method can be used for face detection. [5] used the skin color detection method combined with eye and mouth detection. The study was able to detect faces with a wider variety of face colors. However, the disadvantage of the skin detection method is when the background color is similar to the skin tone. In the Viola-Jones method, face detection can be applied without skin color detection. Research conducted a machine learning approach through four stages, namely Haar-like feature, integral image, adaboost learning, and cascade classifier. The results of the study reached an accuracy of 95% [6].

At the face identification stage, the features of the detected face object are extracted. In this study, face identification was carried out as a substitute for the presence of students. He (2005) uses Laplacian to extract images, then comparing them with eigenfaces and fisherfaces in three different databases [7]. In the three databases, the Laplacian faces method has the highest accuracy than the other two techniques.

However, in terms of image reduction dimensions [8], fisherfaces are better than the others by reducing the smallest image. In this research, the accuracy of the fisherface method tends to be consistent with the increasing amount of training data. Another study conducted by [9] compared the eigenface and fisherface methods. The study compared the two methods with three subsets of training data divided by lighting variations. The result showed that the lowest error rate was obtained at fisherfaces. Meanwhile, testing objects with fisherface glasses attribute also get the smallest error rate, namely 5.3%, with the smallest reduce space, which is 1.

After the faces have been identified, then the method implements emotion detection. Research related to emotion detection based on facial expressions was conducted by [10] using viola-jones and Principal Component Analysis. The method detects emotions such as happy, sad, and neutral with 99.84% accuracy in the study. The emotions seen were sleepiness and focus. In order to get the direction of the head, projection from 2D to the 3D image is employed on the camera. [11] used camera calibration based on Direct Linear Transform (DLT) and Lavenberg-Marquardt Optimization methods. These experiments indicate that the method can improve the accuracy and speed of the camera calibration with the average distance between the measured coordinates and the model's coordinates, which is 0.04548. In this research, the method is implemented to project a face image. Thus, the direction of the head can be determined. Also, pupil detection is used by masking the eye area obtained by knowing the eye area's facial landmarks.

In [12] study, sleep detection was carried out by determining whether the eyes were closed or open at specific intervals. The eye is said to be closed when the sclera and iris are not visible in the image. [13] resulted in 95% accuracy using the Euclidean distance to find the eye area . The study calculates the eye ratio by determining the eye's center of gravity and then calculates the closure degree. In the study, there was no length of the time the eyes were closed to increase reliability. [14] uses facial landmarks as a reference for the eye ratio. Facial landmark is used to detect the eye area with facial landmark points in one eye. By using mathematical calculations, the distance of the coordinates of these points has been found. After that, the eye ratio is determined to detect whether the eye is closed or open. The eye ratio calculation is used to determine the emotion of sleepiness by combining head direction detection [15]. The head direction is to overcome the condition when the eye is detected, but the head is tilted which is detected as drowsiness even though it is open. Sleep detection and focus detection will be implemented in this research to see students' academic emotions in class.

This research uses the Viola-Jones method in face detection. After the face is found, it will be identified using fisherface. Emotion detection is based on facial expressions. Emotions detected were sleepiness and focus using facial landmarks by calculating the eye ratio.

2. Method

This research consists of five stages as seen in Fig. 1.

Fig. 1. Research framework

From Fig. 1, the first stage is image acquisition. At this stage, the images are divided into two groups, i.e. training image and testing image. Training image acquisition obtains a dataset from private datasets in the form of faces at the university. Taking training dataset images is done with a webcam and then stored in the database server. The dataset is divided into ten types of data based on the lighting and attributes used on the object. The attribute category is divided into 4, namely wearing glasses (1), veiling (2), without glasses (3), and without a headscarf (4). The dataset distribution can be seen in Table 1.

Table 1. Dataset Table

Dataset	Lighting (lux)	Attribute	Number ofobject
Dataset 1	50 – 1935	1,2,3,4	26 object
Dataset 2	1935	1,2,3,4	10 object
Dataset 3	1839	1,2,3,4	10 object
Dataset 4	340	1,2,3,4	10 object
Dataset 5	25	1,2,3,4	10 object
Dataset 6	22	1,2,3,4	10 object
Dataset 7	1839	2,3,4	10 object
Dataset 8	1839	3,4	10 object
Dataset 9	1839	2,3	10 object
Dataset 10	1839	1,2,4	10 object

Data collection for each object is conducted 11 times with different poses. The poses are facing in front of eye-opening, facing in front of eye closing, looking up, looking down, facing up, facing down, facing right 25 degrees, facing right 45 degrees, facing left 25 degrees, facing left 45 degrees, and facing free front pose. The image taken is a face image with a size of 90x90 m^2 based on the results of face detection.

In the test image acquisition stage, the webcam camera captures RGB image. Then, the image is converted the greyscale image based on (1). Furthermore, the face detection step is to get the region of interest using the viola-jones method.

$$Grey = 0.299R + 0.587G + 0.114B \quad (1)$$

where *R*,*G*, and *B* are the red, green, and blue value of image. Each pixel of the images applies (1) in order to obtain grey image. After the image converts from *RGB* to grey scale, the viola-jones method is used to detect faces. The viola-jones process is carried out in four stages as in the following lines.

2.1. Features of Haar

The Haar feature [16] finds the difference between the number of black pixels and white pixels of the image with (2).

$$F_{Haar} = F_{White} - \Sigma F_{Black} \quad (2)$$

where F_{haar} is the total feature value, F_{white} is the light area feature value, and F_{black} is the dark area featurevalue.

2.2. Integral image

Integral image [17] is a technique to calculate the feature value by changing the value of each pixel into a newimage representation. The formula of integral image is

$$ABCD = D - (B + C) + A \tag{3}$$

where *ABCD* is a square with an angle *ABCD* representing pixels and *D*, *B*, *C*, *A* are the angles as in Fig. 2.

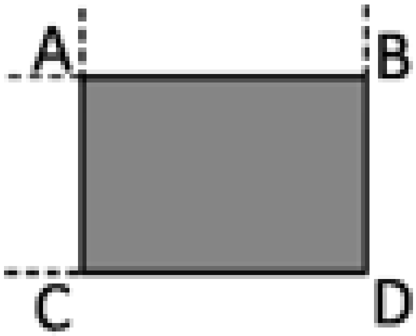

Fig. 2. Integral Image Representation

2.3. Adaboost learning

Adaboost learning combines several weak classifiers into one robust classifier where a weak classifier is acorrect answer with an inaccurate level of truth.

$$h_j(x) = \{^{1,}_{0,} \; if \; p_j f_j(x) < p_j \theta_j(x) \; others \tag{4}$$

where h_j *(x)* is a weak classification, p_j is the parity to j, θ_j is the threshold to *j* and *x* is the sub imagedimension, e.g. 24x24. Adaboost learning trains classifiers at each classification level.

2.4. Cascade Classifier

Cascade classifier is a multilevel classification method [18] to reject image areas that are not detected by faces.After obtaining the face image, the image is converted to 90x90 pixels to equalize the input image size with the training image. It should be noted that the size of the training image and the test image must be equal. After obtaining an image with a size of 90x90, then the fisherface method extracts the image to obtain the features of the image. The method is a combination of the Principal Component Analysis (PCA) and Fisher Linear Discriminant (FLD) methods [9]. The PCA process is explained in the followinglines.

- Convert the face image into a column vector x_N where N is the number of face images. The result of conversion is

$$X = [x_1 \; x_2 \dots x_n]^T \tag{5}$$

- Calculate the average image using the formula as below.

$$\bar{x} = \frac{1}{N} \Sigma_{n-1}^{N} x_n \tag{6}$$

- The transformation matrix *[W]* is obtained based on the eigenvector of the covariance matrix $[C]_x$ which is calculated based on (7)

$$[C]_x = E[x - \bar{x}][x - \bar{x}]^T \tag{7}$$

- Because the dimensions of $E[(X - \bar{X})(X - \bar{X})^T]$ are smaller than $E[(X - \bar{X})^T(X - \bar{X})]$ and A = (X-$\bar{X}$). Assuming the eigenvector of $E[(X - \bar{X})^T(X - \bar{X})]$ is v_i, then

$$A^T A_{vi} = \lambda_i v_i \tag{8}$$

$$AA^T A_{vi} = \lambda_i v_i \tag{8}$$

By multiplying A into (8), then the equation becomes,

$$[C]_X AV_i = \lambda_i v_i \tag{9}$$

- The face recognition process is carried out by utilizing (10),

$$y = [W]^T(\hat{x} - \bar{x}) \tag{10}$$

where ω is a face image that will be recognized by $\hat{X}$, which has been transformed by *[W]*. Then,it is matched with the reference image based on the minimum value of its Euclidean distance.

$$\varepsilon_K^2 = \|\omega - y_k\|^2 \tag{11}$$

- Then the FLD calculation steps are as follows:
  - Transforms the training set into a column vector (face space, Γ).
  - Forms the average face (Ψ) of the face space, and the average face value of each class (classaverage face, Ψ_i).
  - Performs calculation of the within-class scatter matrix (S_W), and the between-class scattermatrix (S_B).
  - Projects the distribution matrix (S_W and S_B) into the PCA projection matrix (W_{pca}).

$$S_{ww} = W_{pca}{}^T S_w W_{pca} \tag{12}$$

$$S_{BB} = W_{pca}{}^T S_B W_{pca} \tag{12}$$

  - Calculates the eigenvalues and eigenvectors of the distribution matrix.
  - Calculates the fisher projection matrix by sorting the Eigenvectors and selecting the non-zero Eigenvalues components. For class C, we always get C-1 eigenvectors which have non-zero Eigenvalues.
  - Calculates the optimal projection matrix (W_{OPT}).
  - Performs normalization on the optimal projection matrix

$$W_{OPT} = \frac{w_{OPT}}{\|w_{OPT}\|} \tag{13}$$

- Calculates the weight of each fisherface [19] against each face image in the training set (face key, $U_{database}$) by projecting the deviation value of the face space against the average face into the optimal projection matrix.

After extracting the image using the fisherface method, the process continues with face identification by matching the test image with the training image in the dataset using Euclidean distance. The Euclideandistance equation can be seen in (14).

$$distance = \sqrt{(x_i - x_j)^2} \tag{14}$$

where x_i is the test image and x_j is the input image.

After the face is identified, emotional identification is carried out. Emotional identification consists oftwo emotions, i.e. focus detection and sleepiness detection. Focus detection detects the direction of the head and eye direction. The head's direction can be determined by projecting a 2D face image on the camera onto a 3D face image. Therefore, the 3D coordinates of the nose tip and chin are compared. In order to find the eye's direction, the ROI of the eye is determined with facial landmarks. Then, the eye image is converted into greyscale, and the eye area masking is applied to remove the outside area of the eye. The process can be seen in Fig. 3.

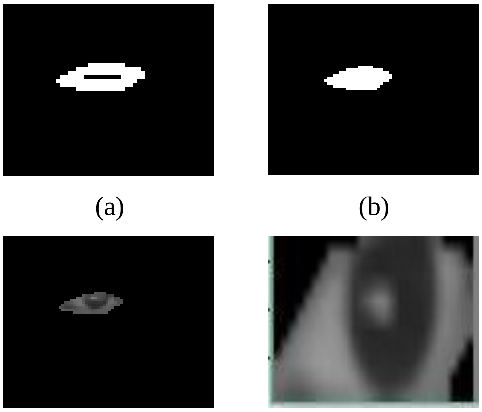

Fig. 3.(a) eye polylines, (b) fillpoly, (c) polyl masking results, (d) eye region of interest

The detection of sleepiness is done by utilizing facial landmarks on the face using the HOG features method. This method recognizes faces into 68 face points, then the face point of the eye is selected as in Fig. 4. The eye aspect ratio is calculated on the eye to determine whether the eye is closed or not based on the equation below.

$$Eye\ aspect\ ratio = \frac{(A+B)}{(2XC)} \tag{15}$$

Fig. 4. Facial landmarks of the eye

where A is the distance between points 38 and 42, B is the distance between points 39 and 41, and C isthe distance between points 37 and 40. The monitoring front page can be seen in Fig. 5.

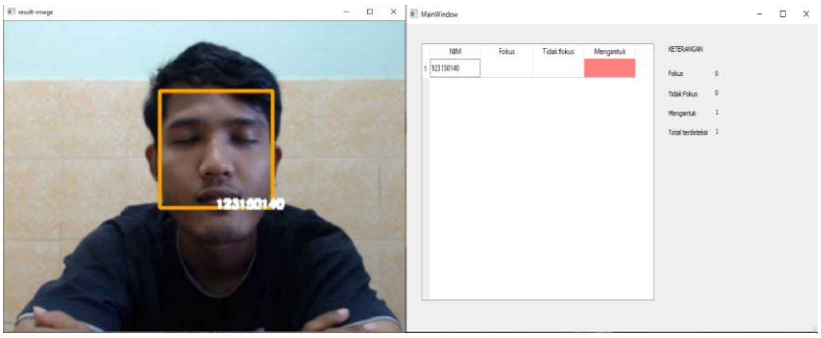

Fig. 5. Monitoring interface

Furthermore, experiments is conducted to test the proposed methods to identify faces and emotions. The experiments includes face detection, face identification and emotion detection testing. The testing canbe seen in Fig. 6.

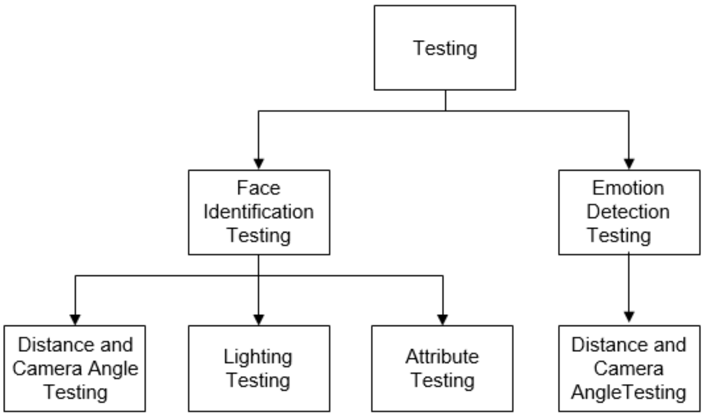

Fig. 6. Experimental Setup

The percentage of success in each type of test uses calculation with the following formula,

$$\text{Percentage} = (\text{number of correct / number of trials}) \text{ x } 100\% \tag{16}$$

3. Results and Discussion

Dataset are retrieved in several steps, i.e. image acquisition, greyscaling, face detection, and facial ROI determination. After obtaining the ROI of the image, the image is resized to 90x90. Example of a greyscaling image result can be seen in Fig.7.

Fig. 7. Example of greyscaling results

The detection of face objects to determine ROI is conducted after obtaining ROI and an image size of 90x90. Image size is changed to size 90x90 to equalize the image size of training data. Examples of the image, including face detection, ROI, and resize, can be seen in Fig. 8.

(a) (b) (c)

Fig. 8.(a) Examples of face detection results, (b) examples of ROI results, (c) examples of resize results

The dataset obtained on each object are 11 images with predetermined different poses. The exampleresults face dataset can be seen in Fig. 9.

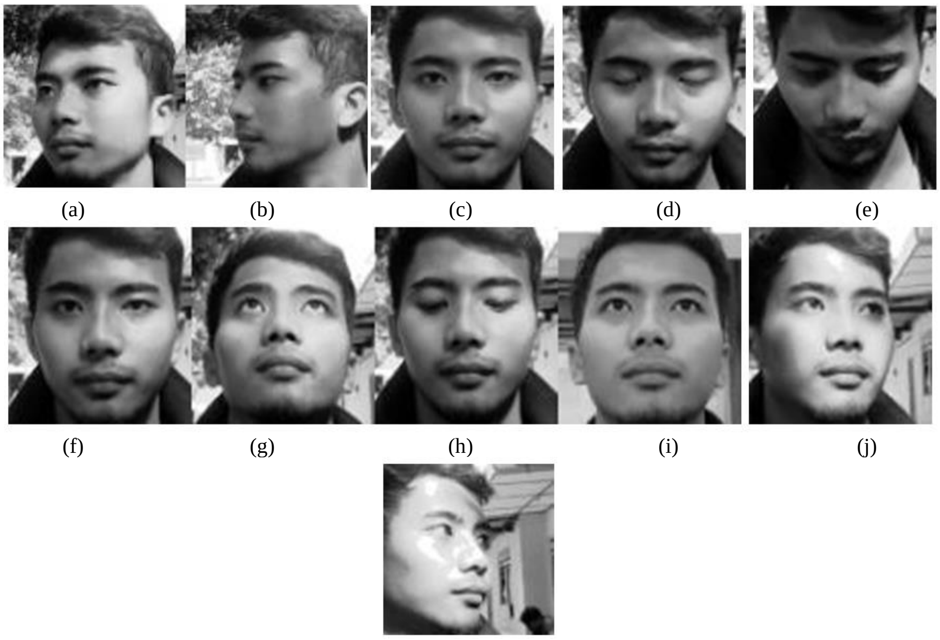

Fig. 9. Example of training data , (a) Face 25 degrees right (b) Face right 45 degrees (c) Face the front,open the eyes (d) Face the front, close the eyes (e) Face down (f) Face to the front free pose (g) Face up (h) See below (i) See above (j) Face left 25 degr ees (k) Face left 45 degrees

Fig. 9 shows an example of an object in dataset 2 with 11 different poses according to the image's description. The dataset is extracted using the fisherface method. Furthermore, face identification is obtained by comparing the training image's proximity with the test image using the Euclidean distance. After face identification, facial landmarks are determined to detect sleepiness and focus.

Several tests is conducted including the distance and camera angle test, lighting test, and the attributes test to determine the accuracy of the method.

3.1. Testing camera distances and angles

A face identification test is essential to measure the percentage of success of the fisherface method forface identification. The distance and camera angle test aim to get the optimal length and angle with the highest accuracy value. In the trial, 26 students of the Informatics Department of UPN "Veteran" Yogyakarta are used as a dataset with 11 different poses. The dataset used in testing is dataset 1. During dataset retrieval as data training dataset 1, it neglects the intensity of lighting. The Tests involved 15 individual students in determining the optimal distance and camera angle. Based on the test, the results are as shown in Table 2.

Table 2. Test data for distance and angle matching faces

Distance	Direction	Total Experiments	True Percentage	Error Percentage
50 cm	Forward	150	56%	44%
50 cm	Above 45	150	57,30%	42,70%
70 cm	Forward	150	57%	43%
70 cm	Above 45	150	34,70%	65,30%
90 cm	Forward	150	37%	63%
90 cm	Above 45	150	24%	76%

In the face matching test in Table 2, the highest results are achieved at a distance of 50 cm and 45 degrees of camera direction to the object. The percentage obtains 57.3% of accuracy. The percentage of correct face identification tends to below 60% at each camera distance and angle. The obtained results arebecause the training data retrieves at different lighting so that it affects the image's gray degree in the training data. Emotion detection testing is conducted to determine the accuracy of facial landmarks. Emotion detection tests were carried out at three different distances, i.e. 50 cm, 70 cm, and

90 cm, where there are two directions, i.e. the front and the top 45 degrees. Testing is carried out on 15 students one by one. The dataset used is dataset 1, but the dataset's use does not affect emotion detection because emotion detection is carried out without matching the test image with the training image. In each experiment, six different poses are detected: facing the blindfold, facing right, facing left, facing up, facing down, and facing forward eye-opening. The face-to-eye pose on the app will be detected as sleepy. The left-facing and right-facing poses will be detected as out of focus if the eye is not looking forward, while the left- facing and right-facing poses will be noticed in focus if the eye is looking forward or not completely looking away. Face up and face down will be detected out of focus, and face to front eye-opening will detect focus. The test results can be seen in Table 3.

Table 3. Data table testing distance and angle of emotion detection

Distance	Direction	Testing	Emotion DetectedRight
50 cm	Forward	150	84%
50 cm	Above 45	150	54%
70 cm	Forward	150	78%
70 cm	Above 45	150	57%
90 cm	Forward	150	50%
90 cm	Above 45	150	34%

Based on the test results in Table 3, it has been found that the distance with the highest percentage of true detected emotions is the distance of 50 cm with the front camera angle. The percentage is 84% accuracy. However, the fisherface method for face detection still obtains a low percentage of success because lighting affects the image extraction results. On the dataset taken with different light, the identification results are lacking. Meanwhile, in emotion detection, because the process does not match the training data, the use of datasets with different lights does not affect emotion detection results. In its development, facial landmarks can be used to detect other emotions based on changes in facial expressions.

3.2. Testing camera distances and angles

Lighting testing is conducted to determine the best lighting. In the dataset, the distance between the objectand the camera is 50 cm. The distance has the highest accuracy on the test distance and camera angle. Taking pictures for training data is done by equalizing the level of lighting; e.g. in testing at a light intensity of 22 lux, the training data is taken at a light intensity of 22 lux. The test uses 10 students. The test results can be seen in Table 4.

Table 4. The test table for lighting

Location	DataSet	Light(lux)	Testing	Detected Right	Detected False	Percentage
A	2	1935	100	71	29	71%
B	3	1839	100	81	19	81%
C	4	340	100	76	24	76%
D	5	25	100	79	21	79%
E	6	22	100	73	27	73%

3.3. Testing with attributes

After obtaining the lighting with the maximum results, testing is carried out by grouping the attributes used on the head, namely glasses and a hijab or veil. There are four groups of attributes used, namely without glasses, without a veil, with glasses, and with a veil. The training data taken were 10 studentswith 11 poses each. The lighting used to adjust the optimal lighting in the test for lighting is 1839 lux withthe attributes according to the test category. For example, in testing with glasses, the student data collection uses the glasses attribute. The test results are shown in Table 5.

Table 5. The test for with attributes used

Attribute	Data Set	Testing	Detected Right	Detected False	Percentage
Without Glasses	7	100	72	28	72%
Without Veil Without Glasses	8	100	96	4	96%
Veil Without Glasses	9	100	85	15	85%
Glasses	10	100	64	46	64%

The highest accuracy is obtained in the test results when the test object and training data are in 1839 lux lighting and do not use the attributes of glasses or hoods. The percentage in this condition is 96% because the glasses can cover facial features. Based on these tests, the fisherface method can be implemented in face recognition, and the use of facial landmarks can be used to detect emotions through changes in facial expressions.

4. Conclusion

In this study, the fisherface method for image extraction and facial landmarks combined with mathematical calculations are able to detect sleepiness and focus. Parameters used in sleep and focus detection are head direction and eye condition. Based on the results of the analysis, design, and discussionthat have been carried out, the fisherface method results on face recognition are obtained with an accuracyof up to 96%. These results were obtained in 1839 lux lighting conditions on the training data and test data while the object did not use a veil and glasses. In detecting emotions using facial landmarks, DLIB can monitor students' emotions with an accuracy of 84% at a camera distance of 50 cm in front of the object.

References

[1] A. R. Dores, F. Barbosa, C. Queirós, I. P. Carvalho, and M. D. Griffiths, "Recognizing Emotions through Facial Expressions: A Largescale Experimental Study," *Int. J. Environ. Res. Public Heal. 2020, Vol. 17, Page 7420*, vol. 17, no. 20, p. 7420, Oct. 2020, doi: 10.3390/IJERPH17207420.

[2] H. Avetisyan, O. Bruna, and J. Holub, "Overview of existing algorithms for emotion classification. Uncertainties in evaluations of accuracies.," *J. Phys. Conf. Ser.*, vol. 772, no. 1, p. 012039, Nov. 2019, doi: 10.1088/1742-6596/772/1/012039.

[3] D. Lu and L. Yan, "Face Detection and Recognition Algorithm in Digital Image Based on Computer Vision Sensor," *J. Sensors*, vol. 2021, 2021, doi: 10.1155/2021/4796768.

[4] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat, and J. Jatakia, "Human Skin Detection Using RGB, HSV and YCbCr Color Models," vol. 137, pp. 324–332, Dec. 2020, doi: 10.2991/ICCASP-16.2017.51.

[5] B. N. Manu, "Facial features monitoring for real time drowsiness detection," *Proc. 2016 12th Int. Conf. Innov. Inf. Technol. IIT 2022*, Mar. 2021, doi: 10.1109/INNOVATIONS.2016.7880030.

[6] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, 2019, doi: 10.1109/CVPR.2001.990517.

[7] X. He, S. Yan, Y. Hu, P. Niyogi, and H. J. Zhang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2019, doi: 10.1109/TPAMI.2005.55.

[8] S. C. Ng, "Principal component analysis to reduce dimension on digital image," *Procedia Comput. Sci.*, vol. 111, pp. 113–119, Jan. 2020, doi: 10.1016/J.PROCS.2017.06.017.

[9] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, 1997, doi: 10.1109/34.598228.

[10] S. Agrawal and P. Khatri, "Facial expression detection techniques: Based on Viola and Jones algorithm and principal component analysis," *Int. Conf. Adv. Comput. Commun. Technol. ACCT*, vol. 2015-April, pp. 108–112, Apr. 2022, doi: 10.1109/ACCT.2015.32.

[11] S. X. Tian, S. Lu, and Z. M. Liu, "Levenberg-Marquardt algorithm based nonlinear optimization of camera calibration for relative measurement," *Chinese Control Conf. CCC*, vol. 2015-September, pp. 4868–4872, Sep. 2019, doi: 10.1109/CHICC.2015.7260394.

[12] V. T. S. S. IEEE and S. S. IEEE, "Drowsiness detection system and method," p. 780, Oct. 2019. Available at: patents.google.

[13] W. Deng and R. Wu, "Real-Time Driver-Drowsiness Detection System Using Facial Features," *IEEE*

Access, vol. 7, pp. 118727–118738, 2019, doi: 10.1109/ACCESS.2019.2936663.

[14] J. D. irawan and E. Adriantantri, "PENDETEKSI MENGANTUK MENGGUNAKAN LIBRARY PYTHON," *Mnemon. J. Tek. Inform.*, vol. 2, no. 1, pp. 22–27, Jan. 2019, doi: 10.36040/MNEMONIC.V2I1.47.

[15] P. Tsokanaki, D. Moraitou, and G. Papantoniou, "The combined effect of sleep and time of day on emotion decoding from dynamic visual cues in older adults," *Neuropsychiatr. Dis. Treat.*, vol. 12, p. 2283, Sep. 2022, doi: 10.2147/NDT.S109959.

[16] A. Mohamed, A. Issam, B. Mohamed, and B. Abdellatif, "Real-time Detection of Vehicles Using the Haar-like Features and Artificial Neuron Networks," *Procedia Comput. Sci.*, vol. 73, pp. 24–31, Jan. 2020, doi: 10.1016/J.PROCS.2015.12.044.

[17] D. Bradley and G. Roth, "Adaptive Thresholding using the Integral Image,", vol. 12, no. 2, pp. 13–21, Jan. 2021, doi: 10.1080/2151237X.2007.10129236.

[18] H. Yang and X. A. Wang, "Cascade classifier for face detection," *J. Algorithms Comput. Technol.*, vol. 10, no. 3, pp. 187–197, Sep. 2022, doi: 10.1177/1748301816649073.

[19] E. Annlin, K. James, and S. Annadurai, "An Efficient Implementation of Weighted Fuzzy Fisherface Algorithm for Face Recognition Using Wavelet Transform," *J. Comput. Sci.*, vol. 8, no. 1, pp. 6–12, Oct. 2021, doi: 10.3844/JCSSP.2012.6.12.