Face Detection System Using Principal Component Analysis (PCA) Method with Eigenface Algorithm

# Fikri Yansyah1 and N H Waryanto2

1,2 Department of Mathematics Education, Faculty of Mathematics and Natural Science, Yogyakarta State University, Colombo Street, Yogyakarta, Indonesia

1 [fikri.yansyah2016@student.uny.ac.id](mailto:fikri.yansyah2016@student.uny.ac.id), 2 nur\_hw@uny.ac.id

**Abstract.** The use of PINs and passwords still have a low level of accuracy and are inadequate, so it is necessary to develop a security system using a face detection system. The purpose of this study is to detect faces using the PCA method with the eigenface algorithm and find out its accuracy. The data used is image data in the form of human face images with a system display using the Python GUI application. This method is used to simplify the observed variables by shrinking the number of pixels from a set of drawing spaces so that the new coordinate system can better draw distinctive features.

# Introduction

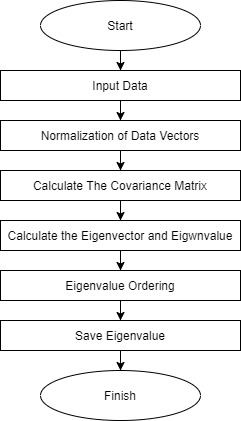
In PIN (Personal Identification Number) and password still can be found a few gaps that cause great harm such as the leak of user data within e-commerce in Indonesia. The data leak includes personal data of users such as username, date of birth, email address, location details, and other user’s personal data. One of the cause indications of the leakage is the security system is still using traditional methods PIN and password that haven’t ensure its security to protect someone’s personal data [1]. PIN and password still have low level of accuracy and easily being hacked that cause a huge loss. One of the solutions to this identity accuracy problem is can be detected and recognized by biometric technology in the form of face recognition system [2]. Face Detection is a process to find out is there a face in an image which is one of the very important initial phase before facial recognition carried out [3].

Currently, there have been many developments in the research of face recognition. Several methods of facial recognition system exist among research development on it, for instance Viola Jones Method, Hidden Markov Model (HMM), Gray Level Co-occurace Matrix (GLCM), Local Binary Pattern Histogram (LBPH), Principal Componen Analysis (PCA), etc [4]. However basically this technology is still can’t be stated perfect yet and needs improvement. Research needs to be developed to gain more accurate result and truly credible authentication. Some aspects needed in this implementation of biometric technology based facial recognition are increasing speed level and accuracy in face recognition [5]. One method of object detector that is used in detecting system is Principal Component Analysis (PCA) method [6]. This method significantly increases accuracy level of facial recognition through simplifying the observed variable by reducing dimension of a data without decreasing its characteristic significantly. Through this reduction, complexity of unnecessary face images can be removed therefore it can decrease computing time in detecting system.

# Research Method

This research development method use Principal Component Analysis (PCA) with Eigenface Algorithm [6]. PCA method with eigenface algortihm have chosen for this research because it can be used to increase the level of accuracy better.

* 1. *Face Detection (Training Phase).* Face detection is a process to find out is there a face in an image or not. This face detection is one of the early steps that is crucial before conducting face recognition. In the biometric technology’s security system, through recognising face form structure needs camera equipments in its identification. Which is device and face detection system act as detection code that works for someone’s face object [3]. In face detection system there are three components considered as the most credible and accurate components. These parts consist of eyes, nose, and mouth. In this phase, it will gain training data in the form of eigenface feature value as identity from each of input images.
  2. *Face Recognition (Testing Phase).* After system getting eigenface value of every images and saving the eigenface value in database, next phase is testing phase in the process of face recognition. Face recognition process is the process to identify an image suitable with image that has been through face adjustment detection process, which is in this process, face image training data in database will be examined its smallest or closest similarity rank. This similarity calculating process uses euclidian distance formula. Furthermore, system will generate the result of identification as “Hasil Nama”.
  3. *Feature Extraction PCA.* Feature extraction is the selecting feature from a form with the value obtained that will be analysed to the next process. This phase is conducted by counting the amount of dots or pixel found in every checking process with PCA Method. PCA is one of the results from applied linear algebra. This PCA method uses a way of reducing dimension from group or image space so that the base or new coordinate system can draw spesific feature from the groups better. This reduction is being conducted by removing correlation between free variable through transformation of origin/source free variable to the new free variable that is completely not correlated without removing important information in it (principal component). This reduction will make more efficient of computing time and complexity from face image that is unnecessary to be removed [7]. The chart of PCA feature value’s search step in the face detecting process [8] is presented by figure 1 as follows

**

**Figure 1.** The chart of PCA feature value’s search step in the face detecting process

*2.3.1. Turning image into a matrix.* First step in the process of face detection is managing a raw image. With measure of w= 200 Pixel and h= 200 *pixel*. After that matrix turned into 40.000 x *m*

*2.3.2.* *Measuring image average*. The next step is measuring image average that is the average of all training images in database. To get the image average there is formula as follows:

*k* = The number of training image matrix in the database

Average image is sum of which is the total of face 1, face 2, until face-n matrix as the amount of training data in database divided with amount of image above, then obtained a matrix with measure *m*x *k.*

*2.3.3. Normalize Data.* After getting the meanvalue, next step is normalize step of measuring difference between image training matrix and mean by subtracting image training with mean

;

Grouped into one matrix 𝐴 =[]

*2.3.4. Finding PCA covariance matrix.* Covariance matrix is all possible variation of column vector’s pair. To find covariance matrix we need to multiply matrix’s normalization result () with its transpose. The following is formula to find covariance matrix [9]:

C = Covarians Matrix

A = Repreaentation of the normalized training image matrix

AT= Transpose of Training Image Matrix

*2.3.5. Eigenvalue Searching.* Next step is eigenvalue searching by using Eigenface algorithm. Eigenface is one of algorithms in the face recognition with PCA method. This algorithm acts to reduce image input dimension by projecting it to subspace that is exist during the training. Its subspace is shaped lowest dimension that indicatesthat Eigenface is found during training phase.*.*

The way Eigenface algorithm works is when a training image is represented in a flatvector (vector combination) afterwards being combined into a single matrix. The matrix result generates *eigenvector* and then being extracted and saved in database. This following is formula to find *eigenvalue* and *eigenvector* [10]:

calculate eigenvalues and eigenvectors

*C x = x*

*C x = x*

Solution to calculate eigenvalue

*Det (C x*

Solution to calculate eigenvector

*(C x*

λ is an eigenvalue matrix, while is a matrix with dimensions *n* x *n* where *n* is a pixel image.

From the Covariant matrix, the Eigen values obtained are then sorted from the largest to the smallest. The eigen value is used to find the eigenvectors associated with each eigenvalue.

Next eigenvector will be searched by the formula

*(C x*

for =0

=0

The above matrix is multiplied and produces a new equation. Until the eigenvector results obtained in the form of a matrix. Then the value of each eigenvector is entered in sequence in one matrix

*2.3.6. PCA Feature Value.* This eigenface feature can be found by transforming source image into face space by using this equation [6]:

=𝐴

= Feature PCA (*eigenface*)

*A* = testing image normal matrix

After obtaining the PCA feature for each image, the next process will save the PCA value in the database

# Results and Discussion

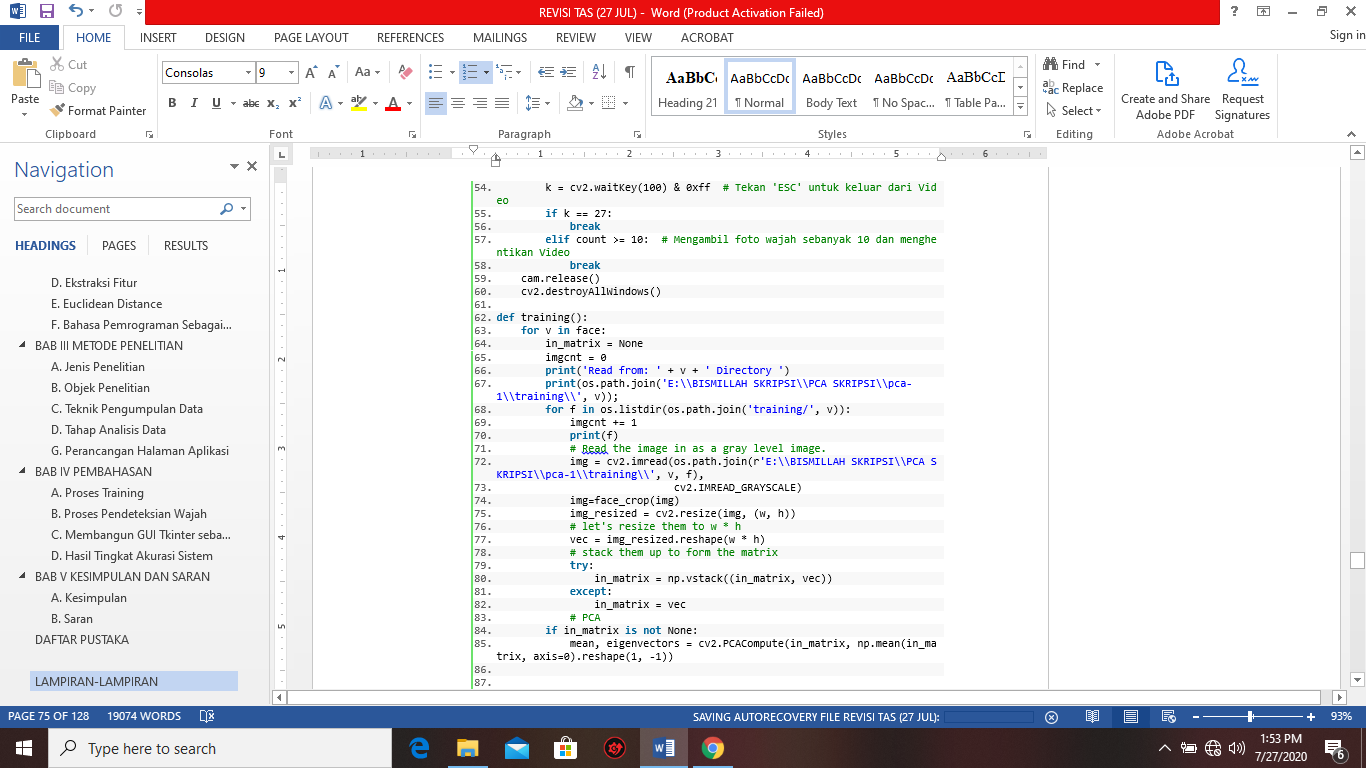
* 1. *Face Detection (Training Phase)*

*3.1.1 Image Input*. Train data of photo/image of human face a number of 20 human face objects which each of face object is taken 10 face images with different slope angle shot (-35º, 0 º, +35 º) lighting, and distance between face and camera (70 cm and 30 cm). Figure 2 is an example of training data obtained.



**Figure 2.** An example of training data obtained

*3.1.2. RGB Image to Grayscale Conversion.* Grayscale conversion process is used to detect the existence of face feature such as eyes, nose, and mouth, in an image by changing colored image (RGB) that is obtained from image input into gray (grayscale) step. Figure 3 is script being used to converse grayscale.

.

**Figure 3**. Script being used to converse grayscale

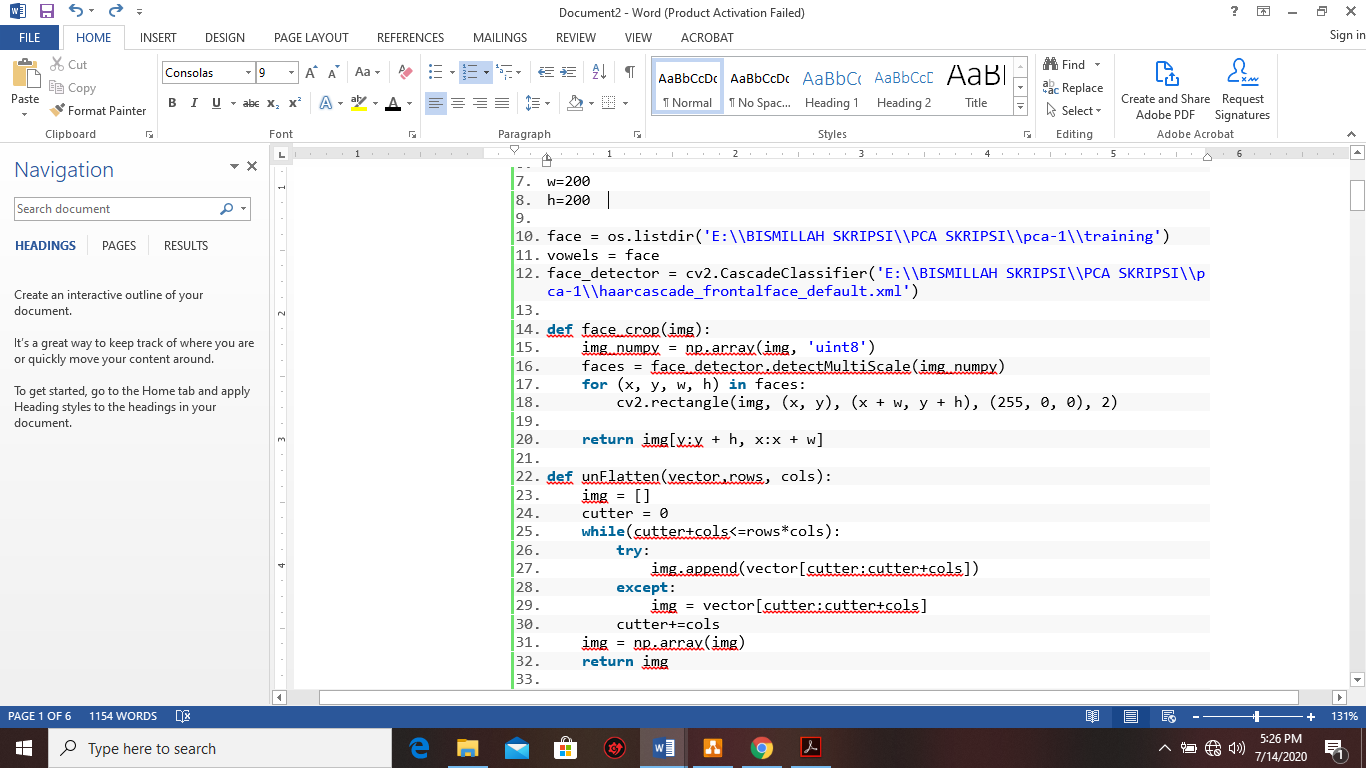
Figure 4 is an example result of converse RGB to Grayscale image

**Figure 4**. Result of converse RGB to Grayscale image



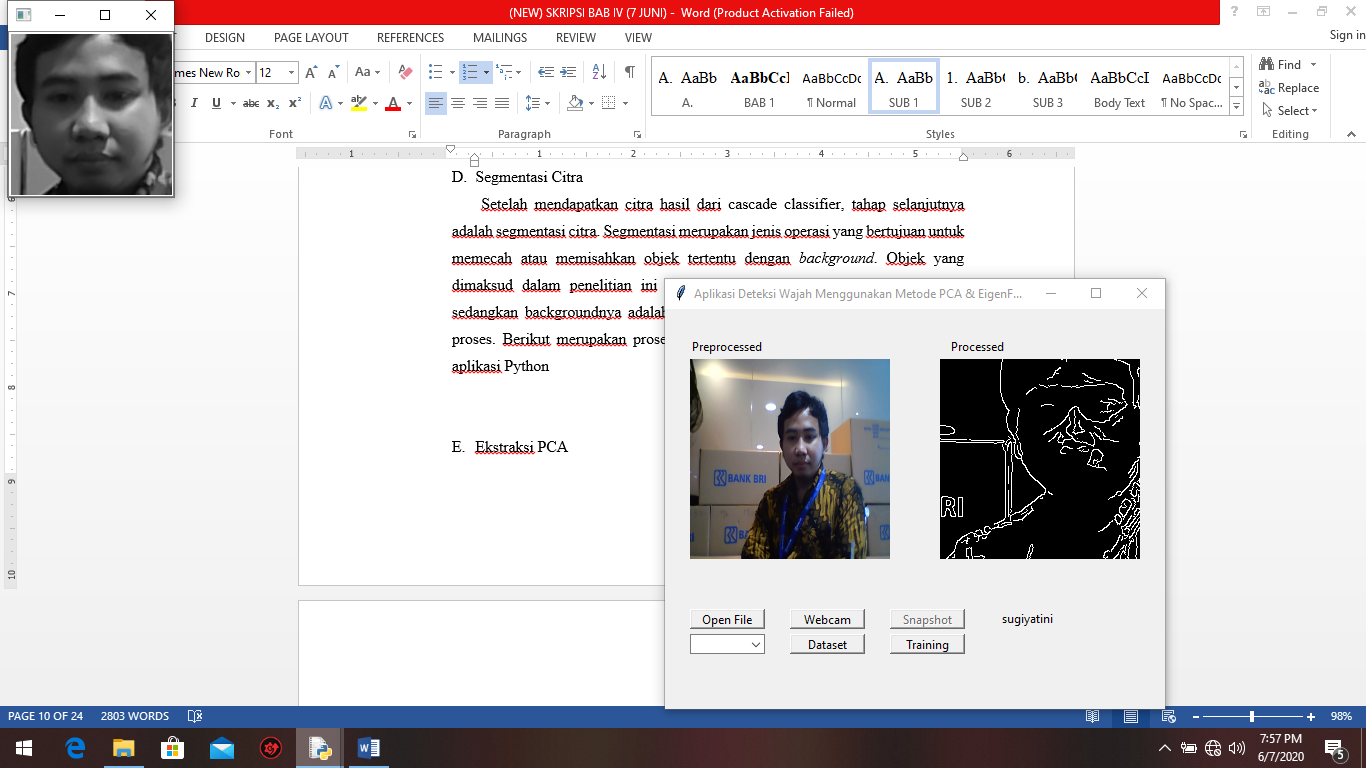
This conversion process used haar like feature method with calculation of image integral to find gap between dark area and light area on face. Final result of this process is haar feature value that will be used in image classification process by using *Haar Cascade Classifier.*

*3.1.3. Image Classification.* This method conducted calculation process from features value obtained from integral image step by organising with stratified classification form. Each of this stratified classification acts to reject image area that is not detected with face that has been trained in each level. Result of this classification process is true for image that fulfilled the value in all classification and false if it failed to fulfill. Final result of this classification is an image segmentation which used to separate between image object and background.

Figure 5 is a script that is used in the classification process

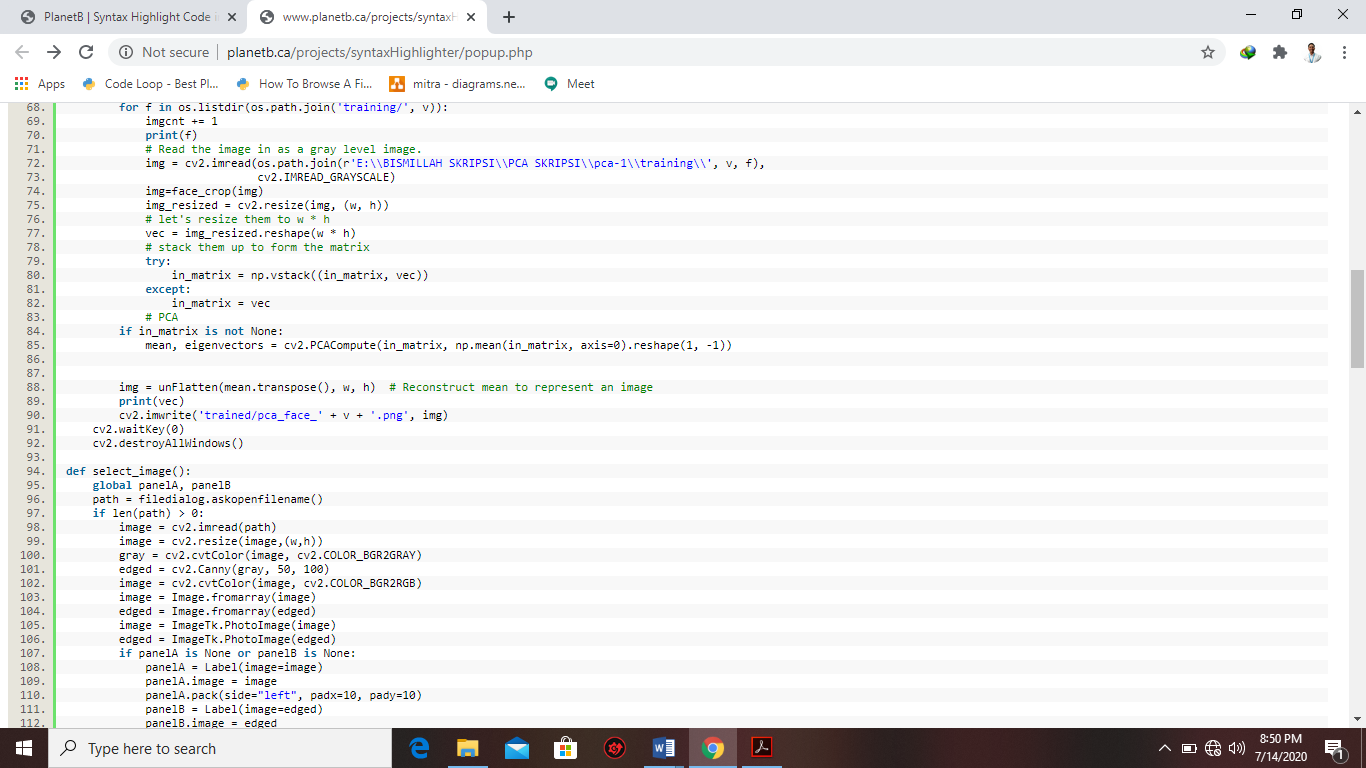
**Figure 5.** script that is used in the classification process

*3.1.4 Image Segmentation.* Segmentation is a type of operation which aims to split or separate certain object with background. The object referred in this research is human face that will be detected, meanwhile background is the background that is unnecessary in the processs. Figure 6 is an example result of image segmentation.



**Figure 6.** Result of image segmentation

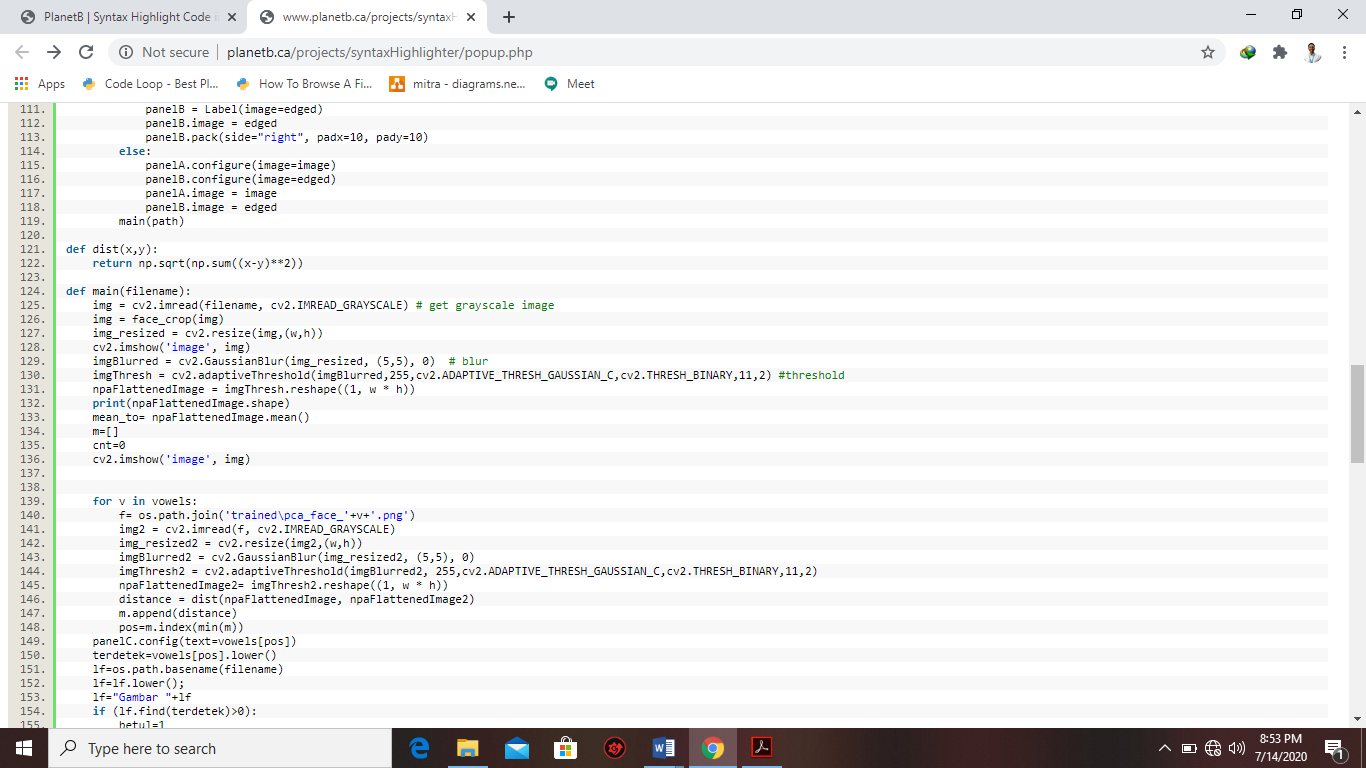
*3.1.5. Feature Extraction with PCA Method.* Image extraction process by PCA used vector called as *eigenvector* and values called as *eigenvalue* to get the most significant feature in dataset. This process started with turning image into matrix which contained image pixels sized *n* x *m* (200 x 200) and transforming it into matrix sized *n* x *k.* Afterwards system will find average value from all over image training in database. After that, it will find covariance matrix by multiply with its transpose. Then conduct eigenvalue searching by using *Eigenface* algorithm until get the feature of each imagesThe Figure 7 is a script that is used to find the value of the PCA feature.



**Figure 7**. Script that is used to find the value of the PCA feature

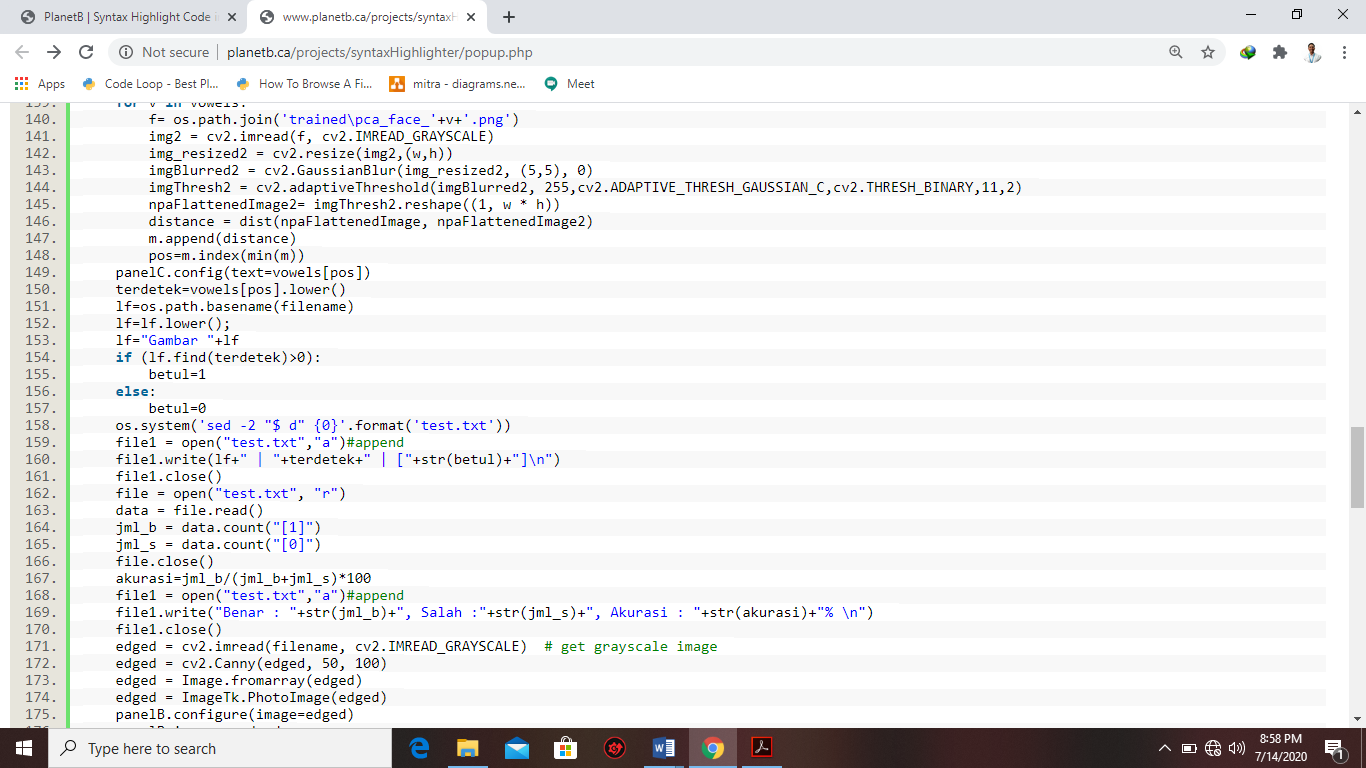
*3.1.6 Saving PCA Value (Eigenface).* Obtained eigenface value will be saved in database and used in training data matching step in image recognition process.

* 1. *Face Recognition (Testing Phase).* Detection process has the same steps like training process to get feature value from each test images. Starting from test image input step until getting PCA feature value. After getting PCA feature value, system will conduct step of matching with train data in database.

*3.2.1. Matching Phase.* The next phase after getting PCA value is the process in comparing one by one with image that has been trained. This process will find smallest or closest face similarity value by comparing one by one with train data in database. This similarity calculating process used *Euclidian distance* formula as image below*.*

**Figure 8.** Calculating process used Euclidian distance formula

*3.2.2. Displaying Test Image Recognition Result.* After getting image with smallest *Euclidian distance*, system then will display the answer of image identification result by using script in this Figure 9.



**Figure 9.** script that is used to find result of image identification

# Conclusion

Principal Component Analysis (PCA) and *Eigenface* algorithm can be used as an excellent face detection system method. This face detection process consist of: (a) taking the input image, (b) converting from an RGB image to grayscale image, (c) The process of calculating the difference between the pixel value in the bright area and the dark area value using the image integral calculation, (d) the image classification process using the haar cascade classifier, (e) the process of separating objects from the background to facilitate the PCA extraction process, (f) the process of searching for distinctive features through the PCA method and the eigenface algorithm, (g) Saving feature value of training data in database, (h) the process of matching feature values with testing data during application testing by comparing one by one with the images that have been trained using euclidean distance. The final result of the detection system is a name that matches the test image.

# References

1. Waryanto, N. H., Andayani, S., Fauzan, M., & Wulandari, T. (2019). Biometrics Patterns Of Eye Iris By Using Hidden Markov Model. *Journal of Physics : Conference Series*, 1-6.
2. Arya, S., Pratap, N., & Bhatia, K. (2015). Future of Face Recognition : A Review. *The 2nd International Symposium on Computer Vision and the Internet*, 578-585.
3. Wang, M., & Deng, W. (2019). Deep face Recognition: A Survey. *Beijing University of Post and Telecommunication*, 1-21.
4. Kumar, K., Wagan, A. A., & Khuhro, M. A. (2020). Texture Based Face Recognition Using GLCM and LBP Schemes. *Indian Journal of Science and Technology Vol. 13*, 1401-1411.
5. Omoyiola, B. O. (2018). Overview of Biometric and facial Recognition Techniques. *IOSR Journal of Computer Engineering (IOSR-JE)*, 1-5.
6. Suryadi, A. (2015). Sistem Pengenalan Wajah menggunakan metode Principal Component Analysis (PCA) Dengan Algoritma Fuzzy C-Means. *Jurnal Penidikan Matematika*, 58.
7. Kustian, N. (2016). Principal Componen Anlysis untuk Sistem Pengenalan Wajah. *Jurnal String Vol. 1 No. 2*, 197.
8. Erwin, Azriansyah, M., & Fachrurrozi, M. (2019). A study about Principal Component Analysis and Eigenface for Facial Extraction. *Journal of Physics: Conference Series*, 1-6.
9. Laltanpuia, R., Vanlalhruaia, & Chawngsangpuii, R. (2018). Face recognition Using PCA. *International Journal of Engineering Science Invention (IJESI)*, 17-28.
10. Wiryadinata, R., Istiyah, U., & Fahrizal, R. (2017). Sistem presensi Menggunakan Algortima Eigenface dengan Deteksi Aksesoris dan Ekspresi Wajah. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi Vol. 6, No.2*, 222.