New Student Drug Tests at College Using Principal Component Analysis Method

Agnes Chrisnalia¹, Edwar Ali², Mardainis³, Rahmiati⁴

Informatics Engineering Study Program, STMIK AMIK Riau^{1,2,3,4} agneschrisnalia@gmail.com, edwarali@sar.ac.id, mardainis@sar.ac.id, rahmiati@sar.ac.id

Article Info

Article History :

Received Aug 27, 2021 Revised Oct 07, 2021 Accepted Dec 21, 2021

Keywords :

Face Detection Drugs Principal Component Analysis College

ABSTRACT

Drugs are substances or illegal drugs that can endanger human life. Someone who consumes it in an inappropriate way will become dependent and even result in death. The physical characteristics of people who use drugs vary, but the more obvious characteristics are on the faces of drug users such as red eyes, stiff facial muscles, dark spots, pupils susceptible to light, sunken face shape, and dullness. The lack of physical characteristics of drug users due to similarities with other diseases makes it difficult for people to recognize them initially. However, for users whose face data has been tracked by the National Narcotics Agency, the facial data is stored in the dataset. This research was conducted with the aim of building a system that can detect and recognize prospective students whether they have ever been included in drug users recorded in the National Narcotics Agency dataset or not as one of the requirements for new student admissions to universities. The system built using the Principal Component Analysis method to process and extract images of the physical characteristics of drug users through the facial image data of drug users stored in the dataset. If the detected face has similarities with the characteristics in the dataset, it is necessary to suspect that the detected face is a drug user. The results of this study are the system is able to detect the faces of drug users using the Principal Component Analysis method with an accuracy of 90% and it is hoped that with this research the system can be one solution in helping universities as an identification effort to minimize drug use so that it can be an additional identification tool which strengthens someone detected using drugs.

© This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Corresponding Author:

Agnes Chrisnalia Informatics Engineering Study Program STMIK AMIK Riau Jl. purwodadi, Pekanbaru, Indonesia, 28294 Email : agneschrisnalia@gmail.com

1. INTRODUCTION

NAPZA (Narcotics, Psychotropics and Addictive Substances) is another term for drugs. Drugs are illegal drugs that can harm human health and can also make the user addicted, causing physical and psychological dependence when trying it. In drugs there are several categories of users, namely light,

medium and heavy users. When drugs enter the human body, they can affect body functions, especially in the brain, face, teeth, and skin.

according to [1], reasons for using drugs due to heavy work, socio economic abilities, and environmental pressures. Drug use in Indonesia is currently at an alarming stage because there are still many people who abuse it, especially in universities. Universities are one of the places to conduct transactions in drug trafficking and even drugs do not only touch students, but also lecturers and professors who abuse drugs. Drug abuse has many negative impacts, both psychologically, physically and socially. The physical characteristics that are affected by drug users are numerous, but the most dominant physical characteristics are clearly visible, namely on the faces of drug users such as a face that looks more droopy or gloomy, there are some black spots, scars, the muscles on the face are stiff, the eye part precisely in the pupil will be vulnerable to light, red eyes, the face shape looks dull, wrinkled and sunken. In plain view, drug users are still difficult to detect with the senses directly because of the similarity of characteristics with a disease.

Several previous studies such as research conducted by [2] implement Principal Component Analysis and Haar-Cascade Classifier methods [3] to monitor lecture attendance with a human face. Research conducted by [4] implementing the Principal Component Analysis method efficiently as a home security [5] to recognize the face image of the occupants of the house in real time and control the opening and closing of the prototype door.

Research conducted by [6] recognize and detect faces in real time and multiple faces using the Principal Component Analysis method with an accuracy rate of 90%. Research conducted by [7] face recognition as identifier [8] using the Principal Component Analysis method, the performance result is 80% in recognizing the input image correctly. Research conducted by [9] identify individuals based on lip print patterns that have been made using the Principal Component Analysis method.

Research conducted by [10] explained that face recognition and detection in real time using Principal Component Analysis was successfully carried out by being simpler and faster with an accuracy of 87.5%. Research conducted by [11] explained that the application of the Principal Component Analysis method into interactive media works well and could read and recognize numbers with an accuracy rate of 90%. So it can be concluded that the Principal Component Analysis method is a good method used to detect and recognize faces.

Based on these problems, researchers will provide a solution by designing and building a system of tools that can detect and recognize the faces of drug users in real time in the acceptance of prospective new students as one of the requirements for admission to higher education. Detecting the physical characteristics of drug users through facial image data of drug users using the Principal Component Analysis method. The Principal Component Analysis method is used to reduce characteristic feature information without having to eliminate the main characteristic feature information with the intention that the system performance will be faster and have high accuracy in extracting characteristic feature information on faces related to a set of data, so that the process of detecting the user's face drugs are more accurate. The face detection recognition process is divided into 2, namely training and testing, where at the initial stage of face training and testing will be processed to the pre-processing stage then look for features on the face of training and testing of drug users using the Principal Component Analysis method.

This tool system works by utilizing the Raspberry Pi [12] as an image processing center using the Principal Component Analysis method, Pi Camera 5 MP as input to capture images in the form of video in real time and Display Touch as an image output monitor that displays text notifications that state the face detection results from input whether drug users are detected or not detected. This system is more effective if it is implemented for drug test testing at the admission of new students to universities. The system built aims to detect and identify prospective new students whether they have ever been drug users and their faces are recorded on the BNN dataset which will be used as one of the requirements for new student admissions to universities. Where if the detected face has similarities with the characteristics based on the BNN dataset, it is mecessary to suspect that the face is a drug user and further action can be taken.

With the construction of this tool system, it is hoped that it can become a one solution in helping universities as an identification effort to minimize drug use so that it can be an additional identification tool that strengthens analysis that a person is detected as a drug user.

2. RESEARCH METHODS

The research method is a stage in conducting research that aims to get the desired results. The stages in conducting this research are as follows:

2.1. System Design

The following in Figure 1 is a block diagram system design in detecting the user's face for drug tests at universities in real time using Principal Component Analysis. The symbol description in Figure 1 is part of the required components such as the Pi Camera, Raspberry Pi 3 Model B, 3.5-inch Touch Display, and smartphone.

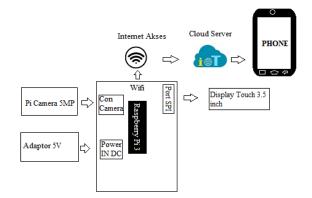


Figure 1. Block Diagram System

The system consists of input, process and output devices. In Figure 1 above, the Pi Camera 5 MP functions to capture video images in real time as a system input connected to the Raspberry Pi camera connector, a 5V DC adapter with a capacity of 3 Ampere is connected via the Raspberry Pi board's DC power input which functions as a voltage source for the entire circuit, and The Raspberry Pi 3 model B functions as the main control center microcontroller and image data processing process as well as connecting the system to the network and server to run the system from the program that was built.

Display touch 3.5 inch as an object viewer and the output of the Raspberry Pi connected to the SPI (Serial Peripheral Interface) port, and the cloud server as an internet of thing server to accommodate data from the device to the system.

2.2. Principal Component Analysis Process

The principle of Principal Component Analysis is to project an image into the eigenspace by finding the eigenvectors of each image and then projecting them into the eigenspace obtained. Principal Component Analysis procedure basically aims to simplify the observed variables by reducing their dimensions [13]. The steps in completing the Principal Component Analysis method [14] to detect the faces of drug users in real time are as follows:

Training Stage

- 1. Make a face image dataset into a flat vector matrix form.
- 2. Calculates the mean of the face image.

$$\Psi = \frac{1}{M} \sum_{n=1}^{M} \Gamma_n \tag{1}$$

Information:Ψ : Mean image trainingM : Number of training imagesΓ : Image training

3. Finding the eigenface value of each face image by reducing the training data matrix to the mean value.

$$\Phi_i = \Gamma_i - \Psi \tag{2}$$

- 4. Calculating the value of the covariance matrix $C = A^{T} A = \Phi_{m}^{T} \Phi_{n}$ (3)
- 5. Calculating eigenvalue and eigenvector

$$\mathsf{CV}_{\mathbf{i}} = \lambda_{\mathbf{i}} \mathsf{V}_{\mathbf{i}} \tag{4}$$

Identification Testing Stage

1. Calculate the eigenface to get the eigenvalues of the image.

$$\Phi_{new} = (\Gamma_{new} - \Psi) \tag{5}$$

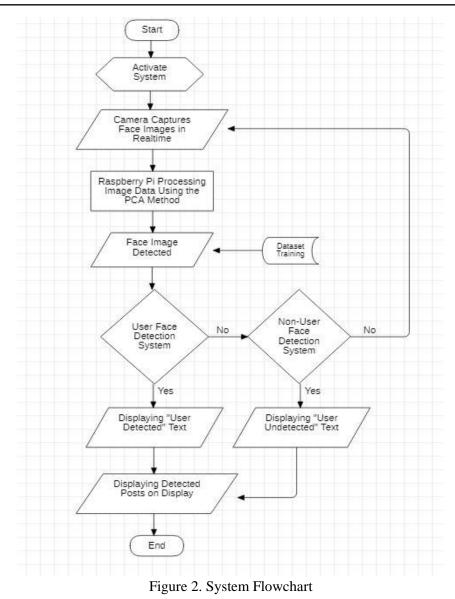
2. Finding the Euclidean Distance value is the shortest distance between the eigenface values of the training data and the eigenfaces of the test data.

$$ED = \Phi - \Phi_{new} \tag{6}$$

Euclidean distance [15] by calculating the smallest difference in the shortest distance between the eigenface training image values in the dataset and the eigenface test faces, where the training image which has the shortest distance to the testing image results from the calculation of the euclidean distance from the image weight matrix is considered suitable if the testing image and training image are in the same class. In other words, both are images of the same face.

2.3. Algorithm

The software used to create an image processing program to detect the faces of drug users is using python programming with the Open CV library on the Raspberry Pi. The flowchart of the testing algorithm in detecting the faces of drug users using Principal Component Analysis can be seen in Figure 2.



Flowchart of the system is built, starting from activating the system and then directing the camera at a person's face to capture facial images in real time, then the facial image data is processed by detecting the shape of the face and the identification process of facial characteristics is carried out as data matching between input and training data on the dataset using the Principal Component Analysis method by Raspberry Pi with an open CV, then calling the training data which stores several samples of facial image data of drug users in the dataset. The dataset is used as a match to the testing process. The results of matching the face image are used to determine whether the face is detected, if the face is detected by a drug user, the touch display displays information notification in the form of text with the sentence "User Detected" and if the face is not detected, the text information notification sentence will contain "User Undetected".

3. RESULTS AND DISCUSSION

3.1. Principal Component Analysis Method Calculation Results Training Stage

1. Create a face image dataset that takes the matrix pixel values into a flatvector form. The pixel value of the image matrix is searched first. The following uses 4 examples of facial image data that have been converted into a matrix into a flatvector form.

Agnes Chrisnalia, New Student Drug Tests at College Using Principal Component Analysis Method

2. Calculate the mean value of the face image.

3. Determine the eigenface value

Subtracting the rows in the training image data matrix value with the mean value. If the value is minus, then the value is replaced with zero.

$$\Phi_i = \Gamma_i - \Psi$$

$$\Phi = \Gamma - \Psi = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 0.75 & 0.5 & 0.75 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.75 & 0.5 & 0.75 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.75 & 0.5 & 0.75 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.75 & 0.5 & 0.75 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.75 & 0.5 & 0.75 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.75 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.75 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.25 & 0.5 & 0.25 & 0 & 0.5 & 0 & 0 & 0 & 0.5 \\ 0.25 & 0.5 & 0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 \end{bmatrix}$$

- 4. Calculates the value of the covariance matrix. This step is used to calculate the eigenvalue and eigenvector by multiplying the result of the eigenface value obtained by the transpose itself.
- 5. Calculate the eigenvalue and eigenvector of the covariance matrix. After the eigenvectors are obtained from the matrix obtained with the eigenvalues taken as nonzero, the resulting eigenvectors are then sorted by their eigenvalues from the largest value to the smallest value. The eigenvector with the largest eigenvalue is the principal component that has the largest variance in all images. The next eigenvector has the second best variant and so on. The eigenvector is generated by substituting the eigenvalues into the equation $\lambda I - Cv = 0$. The eigenvector of each eigenvalue is obtained based on each eigenvalue column and then reassembled into one matrix.

Identification Process Stage

The process by which new faces are entered and will be matched with faces in the dataset to find out which faces match the newly entered faces with the faces in the dataset. The identification step is to calculate the eigenface value for the test face matrix with the same steps as before, namely the face data

in the dataset to get the eigenface value of the new face. The determination of the eigenface and flatvector values for the image starts from the initial determination of the flatvector value minus the flatvector average (obtained from the training image) and obtains the eigenface for the test face.

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ \end{bmatrix} \\ \Phi_{new} &= \begin{pmatrix} \Gamma_{new} - \Psi \end{pmatrix} \\ \Phi_{new} &= \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} - \begin{bmatrix} 0.75 & 0.5 & 0.75 \\ 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{bmatrix} \\ &= \begin{bmatrix} 0.25 & 0 & 0.25 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.25 & 0 & 0.25 & 0 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.25 & 0 & 0.25 & 0 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{bmatrix} = \begin{bmatrix} 0.25 & 0 & 0.25 & 0 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.25 & 0 & 0.25 & 0 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0.25 & 0 & 0.25 & 0 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 & 0.5 \\ \end{bmatrix}$$

To obtain the level of similarity between several face images, the Euclidean Distance method is used to measure the distance. Euclidean Distance method is used to determine the shortest distance between the eigenface value of the training image in the dataset and the eigenface of the test face by calculating the smallest difference between the eigenface training image (Γ_i) in the dataset and the eigenface test face (Γ_{new}), after that add up the matrices of each euclidean distance.

Ŧ

Ŧ

$$ED = \Phi - \Phi_{new}$$

$$ED = \Phi - \Phi_{new}$$

$$ED = \frac{0.25 \ 0 \ 0.25 \ 0 \ 0.5$$

Based on the calculation results, the eigenface distance of face image one has the smallest distance value, which is 0.5. Because the distance between the eigenface of face one and the eigenface of the test face is the smallest, the identification results conclude that the test face is more similar to face one than face two, face three, and face four.

Agnes Chrisnalia, New Student Drug Tests at College Using Principal Component Analysis Method

3.2. Test Scheme

The test scheme uses the face objects of drug users that match in the dataset and human faces that do not match the dataset at random. The detected face object is facing the camera. The scheme of testing the face object can be seen in Figure 3.

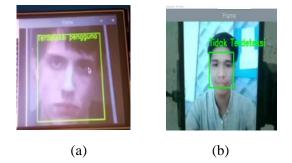


Figure 3. (a) User Detected Face (b) User Undetected Face

In Figure 3 the test scheme (a) presents a face image taken using the appropriate face in the dataset and (b) presents a face image taken using an inappropriate face in the dataset. The results of the testing scheme show that a face that is suitable and has character traits that meet the appropriate dataset will indicate that the face is detected by drug users.

3.3. Test Results

This test is carried out to determine the level of accuracy of the drug test detection system through the faces of drug users by using the training data contained in the dataset and different test data so that the desired results are achieved based on the experiments carried out. This test is carried out in real time with the Pi Camera connected to the Raspberry Pi. The following are the test results using the Principal Component Analysis method.

Table 1. First Test Results				
Tested Face Data	Processing Time (seconds)	Results	Success Percentage	
	3.15	User Detected		
	2.95	User Detected		
	3.51	User Detected	10/10*100	
	2.65	User Detected		
	15.39	User Detected		
	3.63	User Detected	= 100%	
	2.25	User Detected		
	1.94	User Detected		
	4.06	User Detected		
	3.53	User Detected		

In table 1 the results of testing on samples of facial data that match the dataset with image retrieval testing carried out 10 times and recording the time calculation using a stopwatch shows the percentage success rate of accuracy is 100%.

Table 2. Second Test Results				
Tested Face Data	Processing Time	Results	Success	
	(seconds)		Percentage	
	1.36	User Undetected		
	2.25	User Undetected		
	2.57	User Undetected	9/10*100 =90%	
	1.64	User Undetected		
	1.55	User Undetected		
	2.01	User Undetected		
	2.10	User Detected		
	3.06	User Undetected		
	2.51	User Undetected		
	1.62	User Undetected		

In table 2, the results of testing on samples of facial data that do not match the dataset with image retrieval testing were carried out 10 times. The test results show the percentage of success rate of accuracy is 90%.

Based on the tests carried out, the success rate of the system's accuracy in detecting human faces using the Principal Component Analysis method was successfully detected and worked as expected. Although the tests in table 1 and table 2 have quite good face detection accuracy, at any time if retested the results can be different. This is because there are several factors that can affect the level of accuracy of success during testing, namely lighting, facial image training data collection, and the similarity of eigenface values between the training data face image and the test face image so that the system that recognizes the different test face sample data is closest to the face sample data training data in the dataset will be recognized as facial images in the dataset.

4. CONCLUSION

Based on the results of the tests that have been carried out, it can be concluded that: System testing in image retrieval to detect drug tests through the faces of drug users and those who do not use drugs in real time using the Principal Component Analysis method can be carried out properly and as expected with an accuracy rate of 90%. This system can be implemented for drug testing on college students. The success and accuracy of the face detection system for drug users using the Principal Component Analysis method is successful by displaying notification text sentences that are detected by the user and not detected according to the facial characteristics in the dataset.

ACKNOWLEDGEMENTS

The author thanked STMIK Amik Riau for providing support and assistance for this research so that it can run well.

REFERENCES

- [1] BP Hariyanto, "Prevention and Eradication of Drug Trafficking in Indonesia," *J. Sovereign Huk.*, vol. 1, no. 1, pp. 201–210, 2018.
- [2] C. Suhery and I. Ruslianto, "Human Face Identification for Lecture Attendance Monitoring System using Principal Component Analysis (PCA) Feature Extraction," *J. Education and Researchers. information.*, vol. 3, no. 1, pp. 9–15, 2017.
- [3] Susanti, M. Azmi, E. Ali, Rahmaddeni, and YS Wijaya, "Comparison of Boolean Model and Vector Space Model in Searching Text Documents," *digits. Zo. J. Teknol. inf. commune.*, vol. 11, no. x, pp. 268–277, 2020.
- [4] S. Monika, A. Rakhman, and Lindawati, "house security with face recognition system in real time using principal component analysis method," *pros. 4th snatif*, pp. 395–401, 2017.

- [5] KD Septian, SJI Ismail, and A. Sularsa, "Prototype of a Face Recognition Security System Based on Principal Component Analysis (Pca)," *e-Proceeding Appl. science.*, vol. 5, no. 2, pp. 1340– 1349, 2019.
- [6] A. Zein, "Multi Face Detection And Recognition In Real Time Using Principal Component Analysis (Pca) And Eigenface Methods," *Technol. inf. ESIT*, vol. XII, no. 1, pp. 1–7, 2018.
- [7] A. Budi, Suma'inna, and H. Maulana, "Recognition of Facial Image as Identifier Using Principal Component Analysis (PCA) Method," vol. 9, no. 2, pp. 166–175, 2016.
- [8] MA Prastya, "Human Face Recognition System Using Viola-Jones Algorithm And Principal Component Analysis," *Ubiquitous Comput. its Apps. J.*, vol. 2, pp. 85–92, 2019.
- [9] H. Himmah, B. Hidayat, and F. Oscandar, "Individual Identification Based On Liver Print Patterns With Principal Component Analysis And Radial Base Function For Forensic Applications," *e-Proceeding Eng.*, vol. 5, no. 1, pp. 469–479, 2018.
- [10] A. Darmawan, "Real-Time Face Recognition Mobile Application Based on Principal Component Analysis," *Ubiquitous Comput. its Apps. J.*, vol. 2, no. 1, pp. 57–66, 2019.
- [11] H. Dafitri, MS Asih, and RI Astuti, "Interactive Media for Recognizing Numbers with Fingers Using the PCA (Principal Component Analysis) Method," *Query J. Sist. inf.*, vol. 3, no. 2, pp. 57– 65, 2019.
- [12] A. Hanuebi, S. Sompie, and F. Kambey, "Face Recognition Application to Open Doors Based on Raspberry Pi," *Tech. information.*, vol. 14, no. 2, pp. 243–252, 2019.
- [13] N. Kustian, "Principal Component Analysis For Face Recognition Systems Using Eigenface Method," J. String, vol. 1, no. 2, pp. 193–202, 2016.
- [14] Salamun and F. Wazir, "Designing Face Recognition System With Principal Component Analysis Method," *RABIT J. Teknol. and Sis. inf. UNIVRAB*, vol. 1, no. 2, pp. 61–77, 2016.
- [15] YB Muhammad Rizki Muliawan, Beni Irawan, "Implementation of Face Recognition Using the Eigenface Method in the Attendance System," J. Coding, Sis. computer. camel, vol. 3, no. 1, pp. 41–50, 2015.

BIOGRAPHY OF AUTHORS



Agnes Chrisnalia is a student of the 2016 STMIK Amik Riau Informatics Engineering Study Program. Her research topics include digital image processing.



Edward Ali, M. Kom obtained a Bachelor's degree and a Master's degree in Information Technology from UPI UPTK Padang. He has been a lecturer at STMIK Amik Riau since 2004. His current research interests include deep neural learning and deep network learning.



Ir Mardainis, M. Kom obtained a Bachelor's degree and a Master's degree in Information Technology from UPI UPTK Padang. He has been a lecturer at STMIK Amik Riau since 2008. His current research interests include security systems and mobile computing.



Rahmiati, M. Kom obtained a bachelor's degree in Information Technology from UPI UPTK Padang and obtained a Master's degree in Informatics Engineering at the Benarif Indonesia College of Informatics Engineering. She served as vice chairman2 Non-Academic Fields in Finance at STMIK Amik Riau. Her current research interests include the analysis and design of information systems.