Early Detection of Health Kindergarten Student at School Using Image Processing Technology

Muhammad Furqan Rasyid[1], Zahir Zainuddin[2], Andani[3]
[1,2,3]Department of Electrical Engineering, Faculty of Engineering, Universitas Hasanuddin Makassar,
Indonesia, 92119

{rasyidmf17d@student.unhas.ac.id[1], zahir@unhas.ac.id[2], andani@unhas.ac.id[3]}

Abstract. The health early detection system for kindergarten students helps teachers to monitor the health conditions of students. For this reason, the purpose of this system is to propose a system that can detect the health of kindergarten students so that teachers can concentrate more on teaching. This paper presents a technique of combining facial recognition and movement classification for health classifications. For expression recognition, this system uses the PCA method to extract features, then the Euclidean distance algorithm is used to calculate between Eigen Face images and test images. Classification of movements, this study uses two classifications namely active and inactive. For recognition of facial expressions, this system obtains an accuracy of 83.75%.

Keywords: *Kindergarten, PCA, EigenFace, Euclidean Distance.*

1 Introduction

Kindergarten is a level of education before the level of basic education which is a coaching effort aimed at children from birth to the age of six years. Growth and development have increased rapidly at an early age, ie from 0 to 5 years. Monitoring the growth, development, and disruption of child development is carried out in basic health care facilities and in kindergartens. Not only teaching but also in a school, a teacher must also look at the health conditions of his students, as a substitute for the role of parents who function to monitor the health status of their children [1]. This must be considered because the health condition of students can affect the growth of children under five if not handled properly and quickly. From this problem, researchers tried to build a system that could monitor the health conditions of students in schools so that the teacher could focus more on teaching.

There have been many types of research on expressions using video technology [2]. Not only that, a lot of research has been done to look at health conditions. one of these studies used movement monitoring in infants to provide information to doctors [3]. Other studies use video recordings of people who are exercising to analyze motor performance and evaluate their fitness status [4].

This research builds a system that can recognize students' health conditions. This application is built with a system of recognition of facial expressions and combined with the classification of movements. In this case, is the child active or inactive? After that, the two decisions are processed by the system to determine the health condition of the child. there are six human facial

ICOST 2019, May 02-03, Makassar, Indonesia

DOI 10.4108/eai.2-5-2019.2284609

expressions[5], but this system only takes four facial expressions that are most often displayed by students. Expressions consist of smile, happiness, sleepiness, and sadness. then the results of expression recognition are combined with the classification of movements. to determine health conditions, 8 conditions were obtained to determine the child's health condition. This study uses Cohn Kanade and Jaffe data sets then classified according to 4 expressions.

2 Technical Review

This paper applies the technique of early detection of health conditions carried out by continuously monitoring students.

2.1 Face Detection

Face detection is a challenge with increasing use in several applications. This is the first step for facial recognition, facial analysis and other detection of facial features [6]. The difference between face detection and recognition is that face detection exists to identify faces from images and find faces. Face recognition makes the decision "whose face is that?", Using image database [7]. In this case, using the cascade method. This method allows the background area of the image to be disposed of quickly while spending more calculations on the area such as promising objects [8].

Fig. 1. Face Detection

2.2 Expression Recognition

Face detection is one of the most studied problems in vision computing [3]. Face recognition systems can secretly take pictures of a person's face when he is present in a specified area [9]. Starting with facial recognition, images are processed to recognize facial expressions. In this study, we use 4 expressions that are often displayed by students in schools, namely: “ Happy”, “sleepy”, ”sad”, and “smile”. Despite many types of facial expressions, these four expressions are the expressions most often highlighted by kindergarten students.

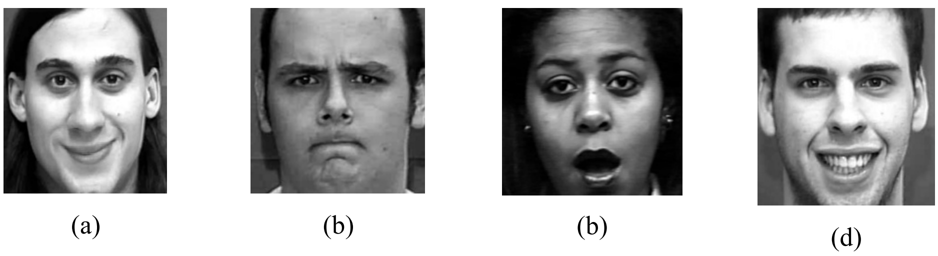


Fig. 2. Exspression Classification :(a) Happy, (b) Sad, (c) Sleepy, (d) Smile

2.3 PCA

The function of principal component analysis (PCA) is to summarize many independent variables (X) which have an indication of correlating or influencing each other into one or more new variables which contain a combination of the previous independent variables (X), which would allow the elimination of multicollinearity problems when a regression model is formed (main component regression).

PCA works on large data sets ($m \times n$) and then transforms them to size ($m \times k$) by maintaining data that is not redundant, or not correlated, or not related. Why is the data uncorrelated? because these data have an important role in the variation of the original data.

This method is effective in eliminating irrelevant information in the image and overcoming their shortcomings[10]. Fig. 3. illustrates the steps of the PCA process.

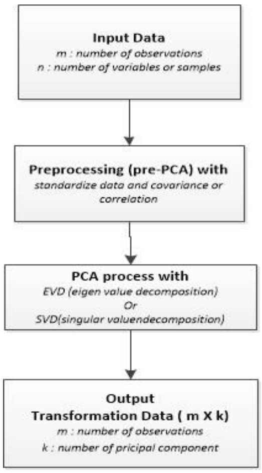


Fig. 3 . step of PCA

The number of the main components formed is equal to the number of original variables. Dimension reduction (simplification) is done by the percentage criteria of data diversity explained by the first few main components. If the first few main components have explained more than 75% of the diversity of the original data, then sufficient analysis is carried out up to the main component.

If the main component is derived from the normal multivariate population with random vector $x = (x_1, x_2, \cdots, x_p)$ and the average vector $\mu = (\mu_1, \mu_{2,\ldots,} \mu_p)$ and the covariance matrix Σ with the eigenvalue $\lambda 1 \geq \lambda 2 \geq \cdots \geq \lambda p \geq 0$ obtained a linear combination of the main components as follows then Var $(y_i) = e^{i\prime} \sum e_i$and Cov$(y_i, y_k) = e^{i\prime} \sum e_i$ where $i, k = 1, 2, \ldots, p$.

The requirement to form the main component which is a linear combination of variables X in order to have the maximum variant is to choose the feature vector (eigen vector), $e = (e_1, e_2, \ldots, e_p)$ so that $Var(y_i) = e_i' \Sigma e_i$ maximum and $e_i{}' e_i = 1$.

- The first main component is a linear $e_1'X$ combination that maximizes $Var\ (e_1'X)$ with the condition $e_1'e_1 = 1$.
- The second main component is a linear $e_2'X$ combination that maximizes Var $(e_2'x)$ with the condition $e_2'e_2 = 1$.

- The main component of i is a linear $e_i'X$ combination that maximizes $Var\ (e_i'X)$ with $e_i'e_k$ = 1 and $Cov\ (e_i'e_k\ )\ =\ 0$ for $k\ <1$.

Between these main components are not correlated and have the same variation as the characteristic root of Σ. The characteristic root of a variety of matrices Σ is a variant of the main component Y, so the matrix of the various types of Y is:

$$\Sigma \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_p \end{bmatrix}$$

The total diversity of the origin variable will be equal to the total diversity explained by the main components, namely:

$$\sum_{j=1}^{p} var(\chi_i) = tr(\Sigma\ ) = \lambda_1 + \lambda_2 + \cdots + \lambda_p = \sum_{j=1}^{p} var(y_i)$$

Depreciation of the dimensions of the original variable is done by taking a small number of components that can explain the greatest part of the diversity of data. If the main component taken is q component, where $q\ < p$, then the proportion of total diversity that can be explained by the main component of i is:

$$\frac{\lambda_1}{\lambda_1+\lambda_2+\cdots+\lambda_p} \qquad i = 1, 2, \cdots, p$$

Decreasing the main component of the correlation matrix is done if the data is first transformed into a standard form Z. This transformation is carried out on the data that the unit of observation is not the same. If the variable observed size is on a scale with a very wide difference or the unit of measurement is not the same, then the variable needs to be standardized.

The raw variable (Z) is obtained from the transformation of the origin variable in the following matrix:

$$Z = \left(v^{1/2}\right)^{-1}(\chi - \mu)$$

$v^{1/2}$ is a standard deviation matrix with the main diagonal element is $(a_{ii})^{1/2}$ while the other element is zero. Expectation value $E\ (Z)\ =\ 0$ and the diversity is :

$$Cov(Z) = (v^{1/2})^{-1} + \sum (v^{1/2})^{-1} = \rho$$

Thus the main component of Z can be determined from the characteristic vector obtained through the original variable correlation matrix ρ. To find the root characteristics and determine the weighting vector is the same as in the matrix. While the trace matrix correlation ρ will be the same as the number of p variables used.

The choice of the main component used is based on its root value, that is, the main component will be used if the root characteristics are greater than one.

2.4 Eigen Face

Main Component Analysis (PCA) forms the basis of eigenfaces. The dimensions of the image matrix are reduced using PCA. for example, if the face image is represented in g-dimensional space, PCA uses linear transformations and aims to get h-dimensional subspace, which answers the maximum variant in g-dimensional space and where h is too small according to g. The average center image is calculated by reducing the normalized training image from the calculated average image. If W is a matrix of the average training drawing centered on W_i $(i = 1, 2, \ldots L)$ and L is the number of training images, the covariance matrix D is calculated from W as in Equation 6.

$$D = W\, W^T \tag{6}$$

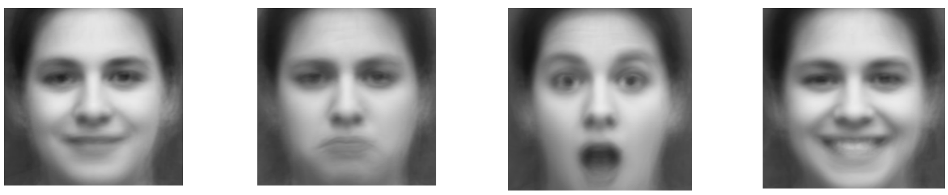
The size of the D covariance matrix can be reduced using D = WTW instead. The eigenvector ei and the eigenvalue λi are obtained from the covariance matrix[11].

$$z_i = E^t W_i (i = 1,2, \cdots, L) \tag{7}$$

In the Equation 7, z_i represents the new feature vector of the new lower dimensional subspace[12].

There is a negative aspect of this method, trying to maximize inter and intra-class scattering. Scattering between classes is good for classification while intra-class scattering is not. On the face of recognition, if there is a lighting variant, this increases intra-class scattering very high, even the class looks tarnished and causes a low classification[13].

In research, all images are trained using eigenface and then stored based on their classification.



(a) (b) (c) (b)

Fig. 4. Eigenface (a) Happy, (b) Sad, (c)Sleppy, (d) Smile

2.5 Euclidean Distance

Array of images that have been trained compared to the test image array using the euclidian distance method. The Euclidean distance is a calculation of the distance from 2 points in the Euclidean space. If $U = (x_1, y_1)$ and $V = (x_2, y_2)$ are two points that want to know the distance, then the Euclidean distance between u and v is [14].

$$EU = (U, V) + \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$ □□□

2.6 Motion Recognition

Recognition of motion is used for vision and processing of computer images, such as event detection, object tracking, behavior recognition, etc. [15]. In this research, using 2 classifications of movement, namely: "Active" and "Not active".

3 System Design

Based on the two classification results, the introduction of expressions and movements, the system makes a decision to determine the health condition.

262 images for training data were classified into 4 expressions. 50 pictures of facial expressions happy , 77 pictures of facial expressions sleepy, 67 images facial expression sad, and 68 images facial expression smile.

Fig. 5. early health detection system

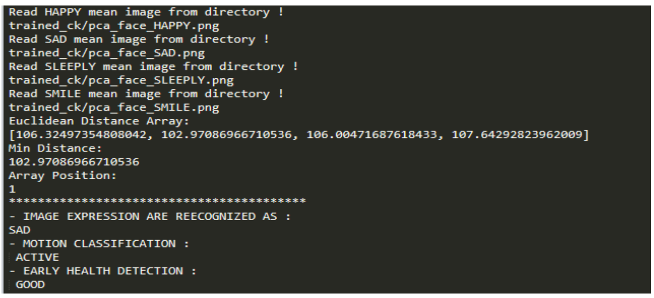


Fig. 6. Result

4 Result And Discussion

To detect proper health conditions, we use three parameters based on the combined results. The following label shows all the appropriate output combinations.

Table 1. Table Of Health Condition

HEALTH PARAMETER		OUTPUT
EXPRESSION	MOTION CLASSIFICATION	
Happy	Active	Good
Happy	Inactive	Good
Sleepy	Active	Good
Sleepy	Inactive	Sick
Sad	Active	Sick
Sad	Inactive	Sick
Smile	Active	Good
Smile	Inactive	Sick

In the table above it can be seen that there are 8 combinations. Of these eight combinations, it can be seen that there are 4 conditions in which health is detected in "Good" and 4 conditions in which health conditions are detected "Sick". This system detects a sick health condition, if the movement is "inactive" as many as 3 conditions, namely when the expression "sleepy", "sad", "smile" and when the movement "active" and expression "sad".For "healthy" conditions, if the movement is "inactive" as many as 1 condition when the expression "happy", and there are 3 times the movement "active" that is when the expression displays "happy", "sleepy", "smile".

4.1 Accuration

For recognition of facial expressions, we calculate the accuracy using the confusion matrix.

Table 2. Confusion Matrix Table Of Expression Recognition

		ACTUAL CLASS			
		HAPPY	SLEEPY	SAD	SMILE
PREICTED CLASS	*HAPPY*	14	0	1	2
	SLEEPY	0	18	0	2
	SAD	3	2	19	0
	SMILE	3	0	0	16

$$Accuracy = \frac{14 + 18 + 19 + 16}{80} \times 100\%$$

$$= 83{,}75\ \%$$

References

[1] J. Coto *et al.*, "Parents as Role Models: Associations Between Parent and Young Children's Weight, Dietary Intake, and Physical Activity in a Minority Sample," *Matern Child Health J*, Jan. 2019, https://doi.org/10.1007/s10995-018-02722-z

[2] M. W. Sullivan and M. Lewis, “Emotional Expressions of Young Infants and Children: A Practitioner’s Primer,” *Infants & Young Children*, vol. 16, no. 2, pp. 120–142, Apr. 2003, doi: 10.1097/00001163-200304000-00005
[3] Long Xu, I. Y.-H. Gu, A. Flisberg, and M. Thordstein, “Video-based tracking and quantified assessment of spontaneous limb movements in neonates,” in *2015 17th International Conference on E-health Networking, Application & Services (HealthCom)*, Boston, MA, USA, 2015, pp. 517–522, doi: 10.1109/HealthCom.2015.7454557
[4] C. Lee, Y. Terada, Y. Liu, and B. Lee, “A motion accuracy evaluator based on body parts movement by MapReduce video processing,” in *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, 2014, pp. 644–647, doi: 10.1109/BHI.2014.6864446
[5] P. Ekman and W. V. Friesen, “Constants across cultures in the face and emotion.,” *Journal of Personality and Social Psychology*, vol. 17, no. 2, pp. 124–129, 1971, doi: 10.1037/h0030377
[6] K. Dang and S. Sharma, “Review and comparison of face detection algorithms,” in *2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence*, Noida, India, 2017, pp. 629–633, doi: 10.1109/CONFLUENCE.2017.7943228
[7] L. Dinalankara, “FACE DETECTION & FACE RECOGNITION USING OPEN COMPUTER VISION CLASSIFIRES,” p. 19.
[8] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Kauai, HI, USA, 2001, vol. 1, pp. I-511-I–518, doi : 10.1109/CVPR.2001.990517
[9] K. C. Nwosu, “MOBILE FACIAL RECOGNITION SYSTEM FOR PATIENT IDENTIFICATION IN MEDICAL EMERGENCIES FOR DEVELOPING ECONOMIES,” *Journal for the Advancement of Developing Economies*, Feb. 2018, doi: 10.13014/K2DF6PD6
[10] M. Zhongli, L. Qianqian, L. Huixin, M. Zhongli, and L. Zuoyong, “Image representation based PCA feature for image classification,” in *2017 IEEE International Conference on Mechatronics and Automation (ICMA)*, 2017, pp. 1121–1125, doi: 10.1109/ICMA.2017.8015974
[11] A. Ozdil and M. M. Ozbilen, “A survey on comparison of face recognition algorithms,” in *2014 IEEE 8th International Conference on Application of Information and Communication Technologies (AICT)*, Astana, Kazakhstan, 2014, pp. 1–3, doi : 10.1109/ICAICT.2014.7035956
[12] M. Sharkas and M. A. Elenien, “Eigenfaces vs. fisherfaces vs. ICA for face recognition; a comparative study,” in *2008 9th International Conference on Signal Processing*, Beijing, China, 2008, pp. 914–919, doi : 10.1109/ICOSP.2008.4697276
[13] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection,” *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, vol. 19, no. 7, p. 10, 1997.
[14] M. D. Malkauthekar, “Analysis of euclidean distance and manhattan distance measure in face recognition,” in *Third International Conference on Computational Intelligence and Information Technology (CIIT 2013)*, Mumbai, India, 2013, pp. 503–507, doi : 10.1049/cp.2013.2636

[15]T. Oh, J. Lee, and I. S. Kweon, "Real-time motion detection based on Discrete Cosine Transform," in *2012 19th IEEE International Conference on Image Processing*, 2012, pp. 2381–2384, doi : 10.1109/ICIP.2012.6467376