ISSN No:-2456-2165

Real Time
Classification of Face Expression Recognition

Dwi Lydia ZA
Master of Informatics Engineering
Faculty of Computer Science
Sriwijaya University, Indonesia

Samsuryadi
Informatics Engineering
Faculty of Computer Science
Sriwijaya University, Indonesia

Dian Palupi Rini
Informatics Engineering
Faculty of Computer Science
Sriwijaya University, Indonesia

Abstract:- Classification of facial expressions is rapidly becoming an important part of computer systems, and interactions between humans and computers. Because the most expressive way of showing human emotions is through facial expressions. Classification of facial expressions is studied through several aspects related to the face itself. When facial expressions change, then the curves on the face such as eyebrows, nose, lips and mouth will automatically change. This study combines real-time facial expressions classification using the Principal Component Analysis (PCA) and Convolutional Neural Network (CNN) methods. This study showed higher results in 9,5% of previous studies using the CNN method.

Keywords:- *Classification, Facial Recognition, Image Recognition, Convolutional Neural Network, Principal Component Analysis.*

I. INTRODUCTION

The introduction of facial expressions or Face Expressions Recognition (FER) has been the topic of recent research on human-computer interaction. Human facial expressions provide important clues about human emotions and behavior. The introduction of facial expressions is very important for applications such as digital entertainment, customer service, monitoring of drivers and emotional robots. There have been extensive studies and methods developed [1] [2].

The majority of the proposed methods are evaluated with limited frontal FER, and their performance decreases when handling non-frontal and multi view FER cases [3]. To face such challenges, this study proposes the Principal Component Analysis (PCA) method and the Convolutional Neural Network (CNN) method to classify facial expressions in the case of multi-view and unrestricted environments with great efficiency and durability.

One method of classifying facial expressions that has quite good accuracy based on several previous studies, PCA (Principal Component Analysis) or often referred to as Eigenfaces is a multifunctional method to use, because Eigenfaces has many functions especially in face recognition such as predictions, deletions redundancy, data compression, dimensional reduction and feature extraction [4].

In the previous study, research was conducted using the Fisherface method with the Backpropagation Neural Network approach, where the test data used was the JAFFE dataset [5]. The face recognition system using Convolutional Neural Network has been implemented against Data Testing The Extended Yale Face Database B [6].

In another CNN study, implementing the Extended Local Binary Pattern as a texture classification was able to overcome the effect of light intensity on the image so that the image affected by light intensity can produce feature pattern extraction that is almost the same as the image with low illumination and configuration weighting initialization parameters with using standard spreads that can speed up convergence and stability rather than randomly initializing [7].

This study combines PCA with CNN for classification of facial expressions in real time using the SFEW 2.0 dataset with an accuracy of 70.4%. This result is higher than the previous study of 9.5% [3].

II. RELATED WORK

Research on facial expressions with a comparison of sparse coding method (comparing TLbSC algorithm with PCA & LPP algorithms using SVM classifier as feature training) with CK+, JAFFE and NVIE dataset [8]. Research on facial expressions with the implementation of raspberry pi II installed on robots to recognize emotions in real time interactive applications, using the Viola Jones Haar Cascade method, Active Shape Model (ASM) for feature extraction and hosting for real time classification [9].

Research on facial expressions using 4 databases namely; CK+, MMI, Oulu-CAISA and AFEW with the Spatio-Temporal Manifold (STM) method and Universal Manifold Model (UMM) [10]. Research on facial expression recognition uses the Principal Component Analysis (PCA) method using the JAFFE dataset, namely: angry, happy, sad, disgust, fear and surprise [11]. Research on image processing and image learning in the introduction of facial expressions [12].

Research on face recognition in real time uses Convolutional Neural Network with CNN model construction to a depth of 7 layers with input from the extraction results of Extended Local Binary Pattern with radius 1 and neighbor 15 showing facial recognition performance reaching an average accuracy rate of more than

89% in ± 2 frames per second [7]. Research on face recognition using the Convolutional Neural Network using the dropout process obtained the best results with an accuracy level of recognition as high as 89,73% [6].

Research on facial expression image recognition uses the Principal Component Analysis (PCA) and Extreme Learning Machine (ELM) algorithms using a dataset from JAFFE, with 210 images of facial expressions, consisting of 10 people with 7 different expressions, taking photos every 3 times where is the ratio of the data train : the dataset used is 4:1 [13]. Research on face recognition and detection for improved safety camera performance [14].

Research on facial expression recognition uses an Enhanced Random Forest Conditional Convolutional Neural Network. By using various CK+ public datasets, JAFFE, BU-3DEF and LFW databases [3]. In this study, how to use and implement the Principal Component Analysis (PCA) and Convolutional Neural Network (CNN) methods as classifiers of facial expressions in real time.

III. PROPOSED METHOD

The process in the classification of human facial expressions consists of three stages, namely face image detection, feature extraction and classification of facial expressions. The face detection process in this study used the Viola Jones method. Viola Jones method is a face detection method that provides face detection results with high accuracy [15]. The flow of the Viola-Jones method in detecting faces (can be seen in Figure 1).

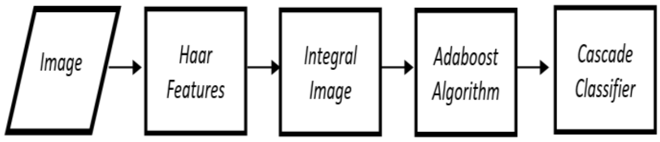

Fig 1:- Process Flow of the Viola-Jones Method

A. Stage of Face Image Detection

Input data is in the form of images that have face objects and frontal face positions, using haar features as object detectors and feature capture.

Then by using integral image to determine the presence or absence of hundreds of haars in an image. Through the Adaboost algorithm it is used to select important features and is used to practice classification. Features that have the greatest restrictions between objects and non objects are considered the best features. The next step is the cascade classifier, a method for combining complex classifiers in a multilevel structure that can increase the speed of object detection by focusing on the area of the image that has a chance. Figure 2 illustrates the structure of the cascade classifier.

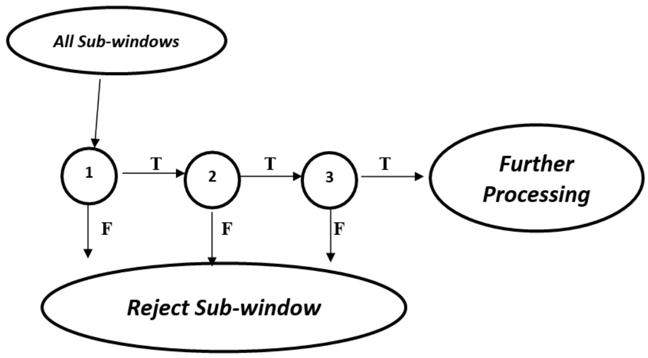

Fig 2:- Structure of the *Cascade Classifier*

B. Image Extraction

After the face image is pre-processed, the face image is extracted to get important features in the image. Feature extraction is the process of taking the characteristics found in objects in the image. In this study, feature extraction uses the Principal Component Analysis method and a new approach in the feature extraction process, namely Feature Learning and Convolutional Neural Network.

PCA is a way of identifying patterns in data, and then data is extracted based on similarities and differences. Since it is difficult to find patterns in data that has large dimensions, where large graphic images are not sufficient, PCA is a powerful method of analyzing that data [16].

Looking for Eigenface value (using the PCA method) that is a significant feature which is a principle component of a collection of faces in the database. Eigenface is obtained from the Eigenvector covariance matrix from the set of images in the database. This Eigenvector is a feature that describes variations between facial images. The stages in taking features with this method are: calculating the average value of the image, calculating the image covariance matrix, calculating the eigenvalue and eigenvector PCA, sorting the eigenvalue from the largest to small and eliminating the small eigenvalue and then determining the eigenface value to be taken.

$$\mu = \frac{1}{N}\sum_{k=1}^{N} X_k \qquad (1)$$

$$C = \sum_{k=1}^{n}(x_k - \mu)\,(x_k - \mu)^T \qquad (2)$$

$$Cu_n = \lambda_n u_n \qquad (3)$$

$$\frac{\sum_{i=1}^{M}\mu_i}{\sum_{i=1}^{M}\mu_k} = \mathrm{A} \qquad (4)$$

C. Image Classification

The classification process is the process of grouping objects into the Convolutional Neural Network method, precisely at the last layer of the Convolutional Neural Network, namely the fully connected layer in the appropriate class.

ISSN No:-2456-2165

In the CNN method, the data is transmitted with a network so that it becomes two-dimensional data, which can produce linear operations and the weighting parameters on CNN that are different. In the CNN linear operation method using convolution operations, although the weight is not one dimensional in size, it transforms into four dimensions which are set of convolution kernels as shown in Figure 3. The dimensions of the weight on CNN are:

- *Neuron Input* X *Neuron Output* X *Height* X *Width* (5)

Due to the nature of convolution, CNN can only be used in data sets that have two-dimensional structures such as image and sound.

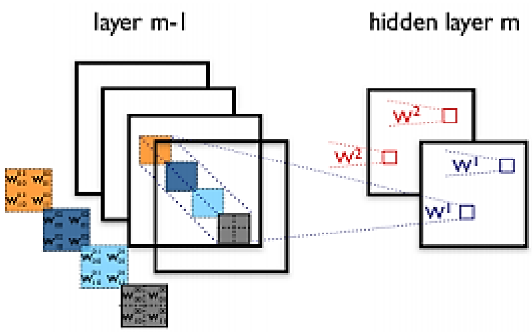

Fig 3:- Convolution Process on CNN

If we use a two-dimensional image I as our input, we might also want to use a two-dimensional K kernel:

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(m,n)K(i-m, j-n) \qquad (6)$$

Convolution is commutative, meaning that we can write equally:

$$S(i,j) = (K * I)(i,j) = \sum_m \sum_n I(i-m, j-n)K(m,n) \qquad (7)$$

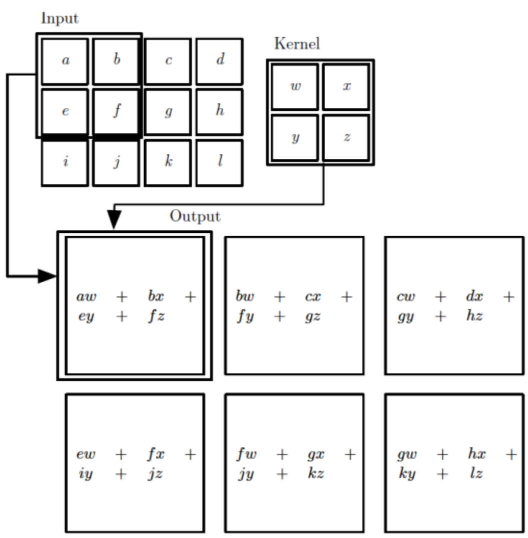

Fig 4:- 2-D convolution without flipping the kernel.

We limit output only to the position where the kernel is located entirely in the image, called "valid" convolution in some contexts. We draw a box with an arrow to show how to upper left element of the output tensor is formed by applying the kernel to the upper left area corresponding to the input tensor.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The method used has been implemented in the dataset of images take through Static Facial Expressions In The Wild (SFEW) 2.0 containing 1073 images with 7 different expressions; anger, disgust, fear, happiness, neutral, sadness and surprise [17].

In this study, testing and calculation of accuracy using confusion matrix, as in Table 1 below.

Confusion Matrix:

Predicted / Actual	Angry	Disgust	Fear	Happy	Sadness	Surprise
Angry	**0.72**	0	0	0.04	0.16	0
Disgust	0.13	**0**	0	0.174	0.565	0
Fear	0.143	0	**0.571**	0	0.286	0
Happy	0	0	0	**0.878**	0.041	0
Sadness	0	0	0	0.033	**0.9**	0
Surprise	0	0	0	0	0.143	**0.571**

Table 1:- Confusion matrix facial expressions from the SFEW 2.0 dataset [proposed]

Table 1 shows the calculation of the accuracy of the confusion matrix in the 6 facial expressions of the SFEW 2.0 dataset, where the expression of disgust is obtained by 0% accuracy, due to the expression generated and variations in expression. The highest accuracy was obtained at happy expressions of 43%, shown in figure 5 with the blackest colored box. The following is a confusion matrix graph from Table 1 where the actual predicted scale is limited to only 45%.

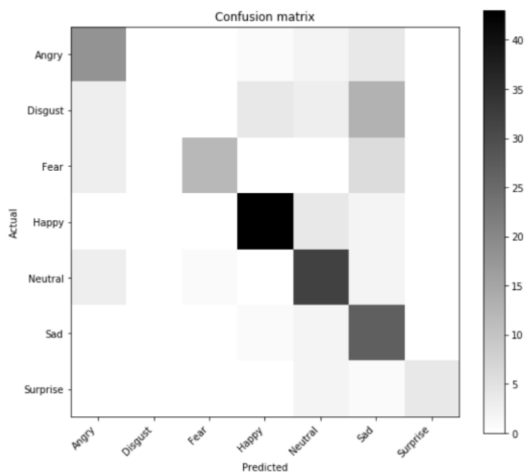

Fig 5:- Graph of confusion matrix facial expressions

In Table 2 shows the results of previous studies, with an average accuracy of 60,9%. This study uses the SFEW 2.0 dataset, a feature classifier method with extracts robust deep salient features of saliency-guided facial patches on CNN [3], as in the following table:

Confusion Matrix:

Predicted / Actual	Angry	Disgust	Fear	Happy	Sadness	Surprise
Angry	**0.559**	0.202	0.166	0	0.009	0.006
Disgust	0.137	**0.505**	0.087	0.02	0.111	0.139
Fear	0.055	0.106	**0.595**	0.016	0.184	0.004
Happy	0.002	0.16	0.01	**0.852**	0.004	0.012
Sadness	0.146	0.142	0.115	0	**0.572**	0.025
Surprise	0.019	0.121	0.156	0.013	0.036	**0.655**

Table 2:- Confusion matrix facial expressions from the SFEW 2.0 dataset [3]

In Figure 6 shows a graph of increasing accuracy in the training process and SFEW 2.0 validation/test dataset with an accuracy of up to 75% at epoch 160. The process of decreasing loss in the SFEW 2.0 training process and validation/test dataset is shown in Figure 7, where at each epoch loss has decreased until epoch 160 has the lowest loss (error) value of 0,8, which is shown in the decreasing graph movement. This proves that the method used in this study can reduce the level of loss in training data.

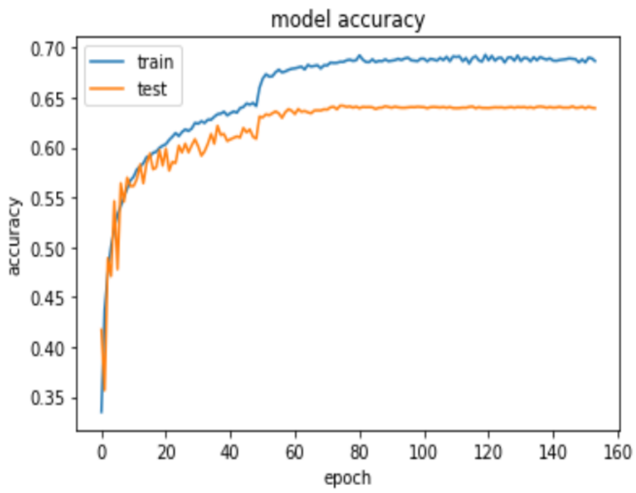

Fig 6:- Data accuracy model set SFEW 2.0

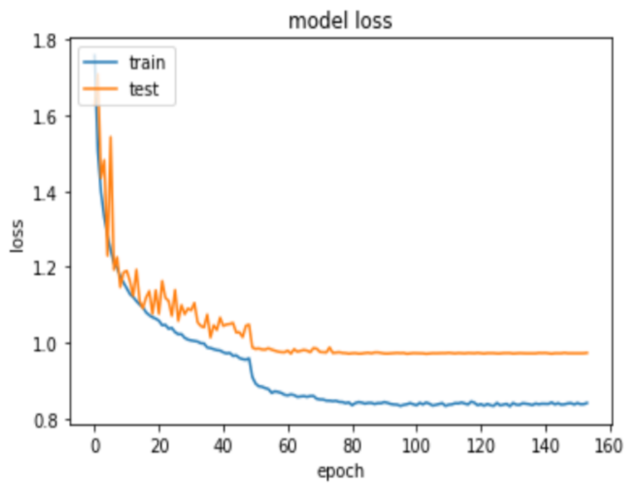

Fig 7:- SFEW 2.0 loss dataset model

V. CONCLUSION

The combined performance of Principal Component Analysis (PCA) and Convolutional Neural Network (CNN) resulted in 9,5% higher accuracy than previous research [3]. PCA functions as a feature extraction and feature selection, which can improve the performance of CNN to classify facial expressions rather than just using CNN.

REFERENCES

[1]. N. Nariswari, T. A. B. Wirayuda, and R. N. Dayawati, “WAVELET , ADAPTIVE BOOSTING FEATURE SELECTION , DAN SUPPORT VECTOR MACHINE,” pp. 0–7, 2011.

[2]. C. M. A. Ilyas, M. A. Haque, M. Rehm, K. Nasrollahi, and T. B. Moeslund, “Facial Expression Recognition for Traumatic Brain Injured Patients,” no. January, pp. 522–530, 2018.

[3]. Y. Liu, X. Yuan, X. Gong, Z. Xie, F. Fang, and Z. Luo, “Conditional convolution neural network enhanced random forest for facial expression recognition,” *Pattern Recognit.*, vol. 84, pp. 251–261, 2018.

[4]. I. N. T. A. Putra, I. B. G. Dwidasmara, and I. G. S. Astawa, “Design and Development of Realtime Attendance Systems Through Face Recognition Methods,” *J. Sains and Teknol.*, vol. 3, no. 2, p. 450, 2014.

[5]. Z. Abidin, “Development of Facial Expression Recognition System using Backpropagation Artificial Neural Networks (Case Study in MUG Database),” *J. Mat. Murni dan Terap.*, vol. 5, no. 1, pp. 21–30, 2011.

[6]. H. Abhirawan, Jondri, and A. Arifianto, “Face Recognition Using Convolutional Neural Networks (CNN),” *Univ. Telkom*, vol. 4, no. 3, pp. 4907–4916, 2017.

[7]. M. & B. S. Zufar, “Convolutional Neural Networks for Real-Time Face Recognition,” *J. Sains dan Seni*, vol. 5, no. 3, pp. 1–6, 2016.

[8]. Y. Yang, W. Zhang, and Y. Huang, “Cross-domain expression recognition based on sparse coding and transfer learning,” *AIP Conf. Proc.*, vol. 1839, no. May, 2017.

[9]. S. Palaniswamy and S. Tripathi, “Emotion Recognition from Facial Expressions using Images with Pose, Illumination and Age Variation for Human-Computer/Robot Interaction,” *J. ICT Res. Appl.*, vol. 12, no. 1, pp. 14–34, 2018.

[10]. M. Liu, S. Shan, R. Wang, and X. Chen, “Learning expressionlets on spatio-temporal manifold for dynamic facial expression recognition,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1749–1756, 2014.

[11]. A. Garg and V. Choudhary, “Facial Expression Recognition Using Principal Component Analysis,” *Int. J. Sci. Res. Eng. &Technology*, vol. 1, no. 4, pp. 39–42, 2012.

[12]. A. Koutlas and D. I. Fotiadis, "Image Processing and Machine Learning Techniques for Facial Expression Recognition," *Handb. Res. Adv. Tech. Diagnostic Imaging Biomed. Appl.*, pp. 247–262, 2011.

[13]. A. R. Wiyono, "MATH unesa COMPONENT ANALYSIS (PCA) DAN EXTREME LEARNING MACHINE (ELM)," vol. 6, no. 2, pp. 2–6, 2018.

[14]. T. Imanuel, "Human Face Detection and Recognition for Enhancing Security Camera Performance," 2018.

[15]. A. R. Syafira and P. Viola, "Face Detection System with Modification of Viola-Jones Method," *J. Emit.*, vol. 17, no. 01.

[16]. C. Zhou, L. Wang, Q. Zhang, and X. Wei, "Face recognition based on PCA image reconstruction and LDA," *Optik (Stuttg).*, vol. 124, no. 22, pp. 5599–5603, 2013.

[17]. A. Dhall, A. Dhall, S. Member, S. Lucey, and T. Gedeon, "Collecting Large, Richly Annotated Facial-Expression Databases from Movies," *IEEE Media*, vol. 6, no. 1, pp. 1–14, 2007.