Journal of Computation Physics and Earth Science
ISSN: 2776-2521 (online)
Volume 3, Number 2, October 2023, Page 50-56
https://journal.physan.org/index.php/jocpes/index

Literature Review: Performance Analysis of CNN, LBP, and Haar Cascade using FER-2013 for Facial Emotion Recognition

Fahar Rafif Arganto[1] Daffa Aly Meganendra[2]
[1]Undergraduate Program in Applied of Instrumentation Meteorology, Climatology Geophysics (STMKG)
[2]Informatics Department, Diponegoro University

Article Info

Article history:

Received September 12, 2023
Revised September 24, 2023
Accepted October 01, 2023

Keywords:

Facial expression recognition,
Local Binary Patterns,
CNN,
Haar Cascade,
FER-2013.

ABSTRACT

The rapid progress in artificial intelligence is transforming how humans and computers interact, with facial expressions being key markers of human emotions. Since facial expressions change dynamically during communication, they offer insights into emotional states and have attracted significant research interest. However, detecting emotions through facial recognition is challenging due to individual differences in expressions, varied lighting conditions, and different facial orientations. These challenges highlight the need for models that can effectively address these issues to improve detection accuracy. This literature review explores several commonly used algorithms for emotion detection via facial recognition, including Convolutional Neural Networks (CNN), Haar Cascade, and Local Binary Pattern (LBP), with the FER2013 dataset serving as the basis for analysis.

Corresponden Author:

Fahar Rafif Arganto,
Undergraduate Program in Applied of Instrumentation Meteorology, Climatology Geophysics (STMKG)
Tangerang City, Banten, Indonesia
Email: fahararg@gmail.com

1. INTRODUCTION

Rapid developments in artificial intelligence play a significant role in relation between human-computer interaction. The face plays an important role in conveying human emotions. Changes to facial expressions during communication are initial indicators of emotional state, making them a significant area of interest for most researchers [1]. The application of emotion detection based on facial recognition has challenges that include various special characteristics of facial expressions at the individual level, lighting conditions, and facial positions which are an object for the development of a model that will be able to overcome them so that a model becomes more accurate [2].

The application of emotion detection based on facial recognition has challenges that include various special characteristics of facial expressions at the individual level, lighting conditions, and facial positions which will be an object for the development of a model that will be able to overcome them so that a model becomes more accurate. In this literature review, there are several algorithms for detecting human emotions through popular facial recognition Methods such as CNN, Haar Cascade, and Local Binary Pattern (LBP) with the FER2013 as a dataset.

The application of this technology has the potential to help in various sectors, such as mental health, behavioral analysis, and surveillance systems, for employees who will be able to monitor the performance of BMKG employees who require high performance in accuracy and precision in intense work situations.

2. RESEARCH METHOD

Facial expression-based emotion recognition systems are often trained using machine learning techniques. A commonly used model is Deep Neural Network, especially Convolutional Neural Network (CNN), for classification and identification purposes.

a. Conculutional Neural Network

CNN (Convolutional Neural Network) are artificial neural network algorithm that is particularly useful in the field of image processing such as object recognition, image classification, face recognition and other computer vision-based tasks. CNN works by taking important features in the data, such as particular patterns on the face, and preventing overfitting through dropout and max polling [3].

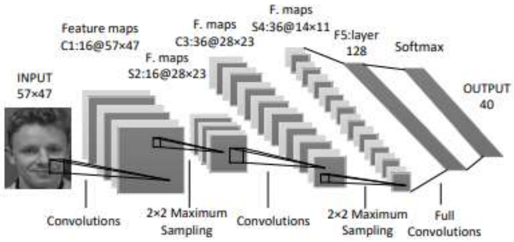

Fig. 1. CNN Architecture

The CNN architecture consists of an input layer, a first convolution layer to filter the input results with a kernel method, a ReLu layer to activate the activation function. followed by a pooling layer to simplify the feature map, another convolution layer with a corresponding pooling layer, then a fully connected (dense) layer [4]. Fig. 1 show the architecture of CNN.

b. Haar Cascade

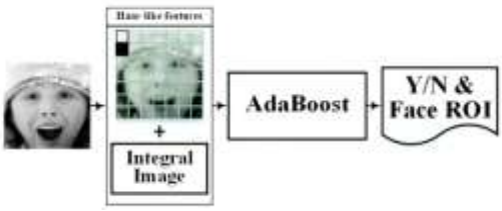

Fig. 2. Haar Cascade Dataflow

Haar cascades using Haar feature by subtracting the pixel value in the blank area from the pixel value in the write area.

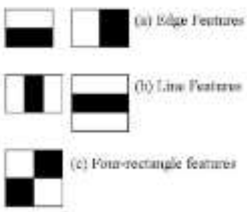

Fig. 3. Haar Cascade Feature

The face detector is based on a 24 x 24 grid. From this base detector, there are about 160,000 potential Haar-Like Features [7]. In Face recognition, the algorithm calculates the difference between pixel sums in these rectangular regions to detect object features, such as eyes or a nose in the case of face detection. Figure 4 shows the schematic of the detection cascade with N stages. The detection cascade is designed to reduce a large number of negative examples using minimal processing.

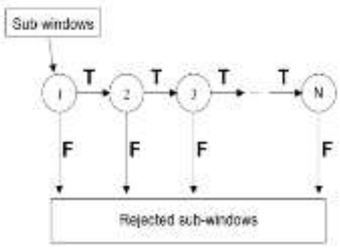

Fig. 4. Cascade Detection

Haar cascade classifiers use a multi-stage process to eliminate non-object regions early, ensuring efficient processing. Each stage contains a classifier that filters out non-target regions, only allowing those likely to contain the object to proceed to the next stage. This cascade improves the computational efficiency of object detection [8].

c. Local Binary Pattern

Local Binary Pattern (LBP) is an algorithm based on texture descriptors for texture classification and is widely used in image analysis, including facial expression recognition. LBP works by analyzing the pixel intensity pattern into a binary number (hence the term "binary pattern"[9].

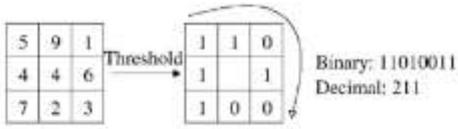

Fig. 4. Basic LBP operator

LBP was used as a general texture descriptor. The operator labels each pixel of the image by comparing the 3x3 neighborhood with the value of the center pixel and making it a binary number. The binary value is obtained by reading clockwise, starting from the upper left neighbor, as shown in the following Fig. 5.

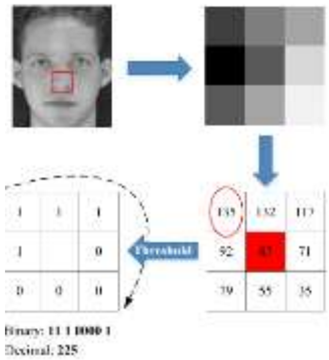

Fig. 5. LBP

To recognize facial expressions, the face image is divided into several small segments. The Local Binary Pattern (LBP) histogram is then computed for each segment. These histograms are combined into a single feature vector that captures spatial information, retaining details about local texture and the overall shape of the face [10].

d. FER-2013

FER-2013 is a datasheet containing 35,887 facial images that are differentiated based on seven emotional expressions namely: happiness, disgust, fear, anger, sadness, neutral, and surprise.This dataset includes images taken from the frontal position of the face with various expressions[11].

Fig. 6. Expression FER-2013

In Fig. 7 showing the distribution of facial expressions in the FER-2013 dataset. In this expression distribution, there is an uneven distribution of 7 expressions in the dataset where the disgusted expression has the smallest relative portion compared to other categories.This can lead to bias in certain classes and become a challenge in model development[12].

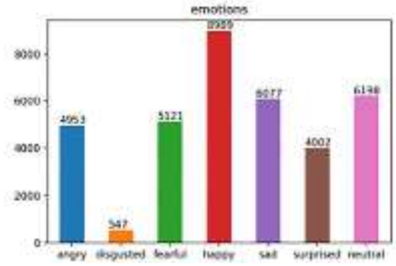

Fig. 7. Distribution of facial expression characteristics in FER 2013

e. Face Recognition

Facial expressions play a significant role in determining a person's emotional state. Parts of the face such as the eye area play a role in recognizing the emotions of sadness and fear. While the mouth is dominant in happiness and disgust. Analysis based on facial action units (AUs) shows that each emotion can show a specific combination of AUs. These findings can be applied to develop emotion recognition algorithms in computer vision [13].

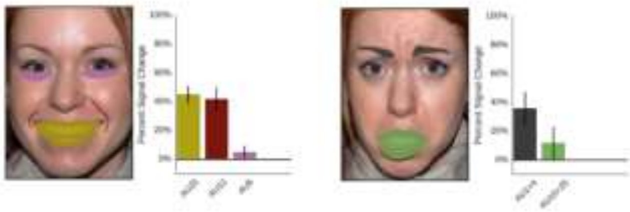

Fig. 8. Facial action Unit

Emotion classification processing is generally divided into data gathering, data pre-processing, advanced extraction, training, and validation, then model deployment [14].

After the data passes through pre-processing, it is continued at the input stage of the human face image, then the extraction of important features on the face and normalization of the feature vector which will then be classified according to the facial expressions that have been labeled on the datasheet to identify human emotions based on the trained model [15].

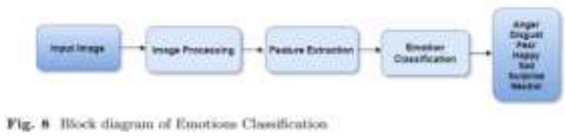

Fig. 9. Face recognition workflow

In the training and validation stage, the datasheet will be divided into two, namely the training set and the testing set to test the model.

3. RESULT AND DISCUSSION

The CNN model will be compared against the CNN+Haar Cascade and CNN+LBP models to find out the best performing model on the FER 2013 dataset.

3.1 CNN

The system consists of transmitter and receiver parts. The components in the transmitter system consist of sensors, GPS, microcontroller and telemetry module. Figure 3 shows the block diagram of the transmitter system.

- In Sarvakar et al. CNN was used to classify five major facial expressions with a two-part process, image background removal and facial expressive vector detection. The model was trained with the FER-2013 dataset consisting of approximately 35,000 images. The model was trained using 80% of the data and tested with 20% of the data. The model then predicts the emotion in the image based on the training that has been performed previously [3].
- Pranav, E et al. using Adam optimization algorithm to CNN by using 5 parameters of facial expressions in the dataset and got an accuracy rate of 78.04% [16].

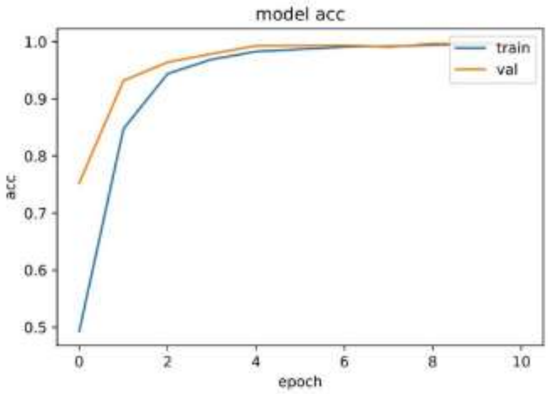

Fig. 9. Accuration rate with Adam optimizer

3.2 CNN+LBP

- Jumani et al. performed experiments using a combination of CNN and Local Binary patterns for face-based emotion identification using the FER2013 dataset with a training set of 28,709 images and a testing set of 3,589 images resulting in a training accuracy rate of 98% and 74% on testing [17]. The fig. (10) shows the accuracy rate on the testing and training sets.

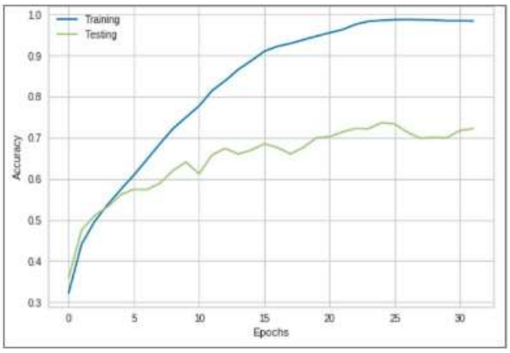

Fig. 10. Training and testing accuracy FER-LBPCNN

- Xu, Q et al. performed LBPCNN by expanding the dataset from FER 2013 with rotation variations up to 90 degrees giving a classification rate of 94.73% on seven facial expression parameters.[18]. Figure 11 shows that the results are higher than conventional models of CNN such as AlexNet (64.29%) and EmotionNet (66.71%) with the same dataset treatment.

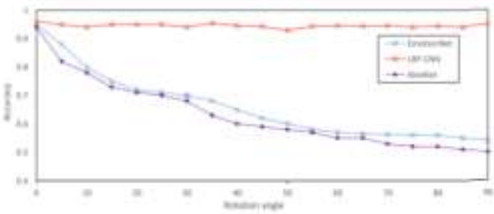

Fig. 11. LBCNN performance

3.3 CNN+Haar Cascade

- The integrated method between Haar Cascade and CNN works by dividing into training stage and testing stage. Haar cascade performs face segmentation and trains the CNN model to detect human emotions in the training stage and in the testing stage [19].

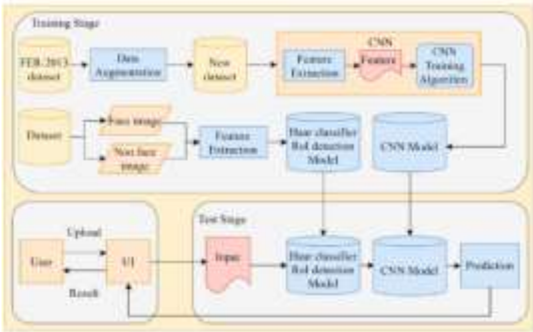

Fig. 12. CNN-Haar Cascaade mechanism

- Rasheed et al. Combined the Haar Cascade classifier and CNN methods by obtaining a validation performance accuracy of 65.59%, with a training time of about 87 seconds per epoch [15].

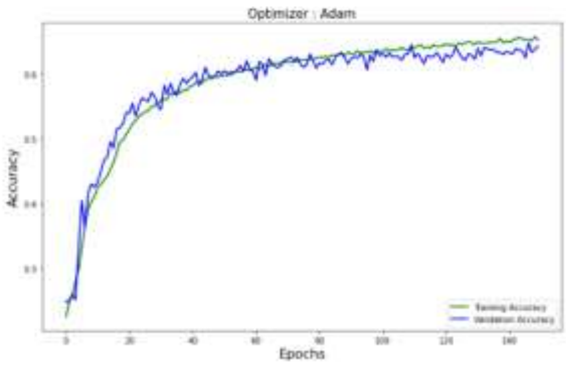

Fig. 13.CNN-Haar Cascaade accuracy rate

- Yeh et al.tested LBPCNN with FER2013 dataset using augmentation method Horizontal flipping and color adjustment on the dataset obtained accuracy rate of 68.36% and 64.34% respectively. This system has been tried to detect 75 different people and get 73.53% accuracy results and be able to correctly identify the emotions of 57 people with an accuracy rate of 55.8% [20].

4. CONCLUSION

Facial expression recognition is a complex and challenging task in the field of machine learning. Many methods and techniques have been developed to obtain good accuracy in this recognition task. A combination of LBP CNN, Haar Cascade CNN, and conventional CNN is discussed and gives various output accuracies. The treatment of the FER-2013 dataset is also affected in the accuracy validation outcome. From all the reviews, it is found that convolutional neural networks are used as the basic foundation for the development of emotion recognition through facial expressions as it provides high accuracy.

REFERENCE

[1] W. Mellouk and W. Handouzi, "Facial emotion recognition using deep learning: Review and insights," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 689–694. doi: 10.1016/j.procs.2020.07.101.

[2] X. Lu, H. Yu, and B. Shan, "Relationship between Employee Mental Health and Job Performance: Mediation Role of Innovative Behavior and Work Engagement," *Int J Environ Res Public Health*, vol. 19, no. 11, Jun. 2022, doi: 10.3390/ijerph19116599.

[3] K. Sarvakar, R. Senkamalavalli, S. Raghavendra, J. Santosh Kumar, R. Manjunath, and S. Jaiswal, "Facial emotion recognition using convolutional neural networks," *Mater Today Proc*, vol. 80, pp. 3560–3564, Jan. 2023, doi: 10.1016/j.matpr.2021.07.297.

[4] Z. Xie, J. Li, and H. Shi, "A Face Recognition Method Based on CNN," in *Journal of Physics: Conference Series*, Institute of Physics Publishing, Nov. 2019. doi: 10.1088/1742-6596/1395/1/012006.

[5] J. Li, K. Jin, D. Zhou, N. Kubota, and Z. Ju, “Attention mechanism-based CNN for facial expression recognition,” *Neurocomputing*, vol. 411, pp. 340–350, Oct. 2020, doi: 10.1016/j.neucom.2020.06.014.

[6] P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features,” 2001.

[7] P. A. Riyantoko, Sugiarto, and K. M. Hindrayani, “Facial Emotion Detection Using Haar-Cascade Classifier and Convolutional Neural Networks,” in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Mar. 2021. doi: 10.1088/1742-6596/1844/1/012004.

[8] “Evaluation_Of_Haar_Cascade_Classifiers_D (1)”.

[9] C. Shan, S. Gong, and P. W. McOwan, “Facial expression recognition based on Local Binary Patterns: A comprehensive study,” *Image Vis Comput*, vol. 27, no. 6, pp. 803–816, May 2009, doi: 10.1016/j.imavis.2008.08.005.

[10] C. Shan, S. Gong, and P. W. McOwan, “Facial expression recognition based on Local Binary Patterns: A comprehensive study,” *Image Vis Comput*, vol. 27, no. 6, pp. 803–816, May 2009, doi: 10.1016/j.imavis.2008.08.005.

[11] K. Sarvakar, R. Senkamalavalli, S. Raghavendra, J. Santosh Kumar, R. Manjunath, and S. Jaiswal, “Facial emotion recognition using convolutional neural networks,” *Mater Today Proc*, vol. 80, pp. 3560–3564, Jan. 2023, doi: 10.1016/j.matpr.2021.07.297.

[12] N. Zhou, R. Liang, and W. Shi, “A Lightweight Convolutional Neural Network for Real-Time Facial Expression Detection,” *IEEE Access*, vol. 9, pp. 5573–5584, 2021, doi: 10.1109/ACCESS.2020.3046715.

[13] M. Wegrzyn, M. Vogt, B. Kireclioglu, J. Schneider, and J. Kissler, “Mapping the emotional face. How individual face parts contribute to successful emotion recognition,” *PLoS One*, vol. 12, no. 5, May 2017, doi: 10.1371/journal.pone.0177239.

[14] E. S. Agung, A. P. Rifai, and T. Wijayanto, “Image-based facial emotion recognition using convolutional neural network on emognition dataset,” *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-65276-x.

[15] L. Rasheed, U. Khadam, S. Majeed, S. Ramzan, M. S. Bashir, and M. M. Iqbal, “Face Recognition Emotions Detection Using Haar Cascade Classifier and Convolutional Neural Network,” Nov. 28, 2022. doi: 10.21203/rs.3.rs-2048290/v1.

[16] *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE, 2020.

[17] S. Z. Jumani *et al.*, “Facial Emotion Identification Based on Local Binary Pattern Feature Detector,” *Indian J Sci Technol*, vol. 12, no. 28, pp. 1–6, Jul. 2019, doi: 10.17485/ijst/2019/v12i28/146447.

[18] Bing. Xu and Kefen. Mou, *Proceedings of 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC 2020) : June 12-14, 2020, Chongqing, China*. IEEE Press, 2020.

[19] R. Ravi, S. V. Yadhukrishna, and R. Prithviraj, “A Face Expression Recognition Using CNN LBP,” in *Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020*, Institute of Electrical and Electronics Engineers Inc., Mar. 2020, pp. 684–689. doi: 10.1109/ICCMC48092.2020.ICCMC-000127.

[20] J. F. Yeh, K. M. Lin, C. C. Chang, and T. H. Wang, “Expression Recognition of Multiple Faces Using a Convolution Neural Network Combining the Haar Cascade Classifier,” *Applied Sciences (Switzerland)*, vol. 13, no. 23, Dec. 2023, doi: 10.3390/app132312737.