PalArch's Journal of Archaeology of Egypt / Egyptology

Deep Convolution Neural Network for Facial Expression Recognition

[1]Soleen Basim Mohammed, [2]Adnan Mohsin Abdulazeez

[1]Department of Computer Science, Faculty of Science, University of Zakho, Duhok Kurdistan region, Iraq.
[2]Presidency of Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq

[1]Soleen Basim Mohammed , [2]Adnan Mohsin Abdulazeez, Deep Convolution Neural Network for Facial Expression Recognition ---Palarch's Journal Of Archaeology Of Egypt/Egyptology. ISSN 1567 - 214x
Keywords— facial expression recognition, principle component analysis, convolutional neural network

Abstract— The Facial Expression Recognition (FER) is still considered as an open research problem and proposing new techniques for more accurate recognition is challenging task. Thus, the main aim of this paper is a proposed algorithm for facial expression recognition FER. The proposed method is based on deep neural network, namely, Convolutional Neural Network (CNN) technique. The main objective of the proposed study is classifying emotions of facial expression into different types, natural, smile, surprise, disgust, squint and scream emotions are used in this study. The efficiency of the selected deep CNN for feature extraction and classification has been proven comparing other techniques. For feature extraction Principle Component Analysis (PCA) method, and K-nearest neighbor (KNN) for classification have been used. Experiments are carried out on the Extended Cohn-Kanada (CK+) and the Japanese Female Facial Expression (JAFFE) datasets to show the effectiveness of the proposed method. The evaluation results obtained height recognition rate 98.5%.
Keywords— facial expression recognition, principle component analysis, convolutional neural network, k-nearest neighbor.

I. INTRODUCTION

Facial expression consists of a lot of information, that it is significant way to express the emotion and exchange information [1]. Due to the evolution in science of technology, many people are trying to use image processing technology and machine vision technology to achieve automatic FER technology.

FER technology has several application for instance, behavioral science has used for social information providing such as origin, gender, and age [2]. Moreover, human machine interaction has been utilized in different fields; for example, computer interactive games, security-surveillance, and robots. A FER are recognized by humans virtually without effort while in computing techniques it is still considered as a challenging task for expression recognition. In the FER techniques, the optimal pre-processing stage, feature extraction stage, and classification stage are considered as main challenging. Particularly, under conditions of the input data variability, pose of the head, multiple sources of facial variability, environmental illumination, and clutter [3, 4].
Technologies of FER are usually depended on geometry appearance features for recognizing emotions [5]. A geometry features are extracted from the position, shape, and distances between the facial components for instance eyebrows, mouth, eye, and nose. On the other hand, the appearance features are dealing with the whole face, as well as it may deal with the specific regions using the texture caused by expression such as wrinkles, furrows, etc. [6].

Feature extraction is a crucial stage in FER because the quality of feature vectors could affect the performance of the classifier. Thus, significant feature should be extracted based on relevant techniques. More so, techniques for extraction features in low-level are considered as feature extraction while these techniques cannot extract features in high-level. Lately, Deep Learning (DL) techniques have been exploited for extracting significant features and they considered as promising techniques for extracting features in FER field. Moreover, both levels of features high and low could be extracted using DL techniques[7]. Due to the capability, DL techniques have been applied in image processing as well as in signal processing successfully.

In the recent age, the most popular DL technique that has been used for extracting significant features is CNN. Thus, good data representations could be extracted based on DCNN due to its popularity that stems from its ability from image data. The tasks of computation intensive of DCNN's are able to run on GPU that in high performance results in consuming very low power. For facial feature extraction, CNN broadly is used for determining the age, gender, and more so.

II. RELATED WORK

In computer vision, the Facial expression recognition (FER) field has become important and active in research. There are many works that have applied different algorithms and techniques for FER. The study of [8] introduced a method depending on using Local Binary Pattern (LBP) and CNN expression features based fusion technique. They used the LBP feature for facial expression image while where merge the abstract extracted features and learned based on CNN with the amended texture of LBP features for facial expression in the full connection layer of CNN technique. Their experiment used the both ORL face and CMU-PIE face databases. Based on the evaluation experiments the performance of accuracy of 91.28% has been achieved. To solve the problems of FER, the study of used deep CNN model to classify emotions in facial expression. On the FER technique, they used the method of transfer learning in order to fine-tune the model of face recognition to the facial expression model at the first step. Next step, to ameliorate the ability of classification, they used an amended Softmax loss function as well as a double activation layer in the introduced method. To evaluate the performance of their proposed study both datasets SFEW2.0 and FER 2013 are used. They obtained accuracy of 48.5% for SFEW2.0 dataset and 59.1% for FER 2013 dataset, the accuracy of both emotions disgust and natural was not good. The study of [11] presents a method for recognition of emotions in facial expression, they used the CNN technique for 2D plus 3D features depending on the FER method. The network that has presented was consists of two CNNs, one CNN has been used for 3D face shape while the second one has been used for face appearance with their colors in order to obtain efficacy and robustness of the proposed study. Thus, a network with 3 convolutional layers has been constructed which includes both layers max-pooling and normalization as well as with two fully connected layers. Moreover, the last layer used 6-way of softmax for the output. To evaluate the performance of their proposed work, the BU-3DFE dataset has been used, the rate of FER achieved 92%. The study [12]employed CNN to devise a FER. In the first step, they implement Haar-like features for detecting the face and used histogram equalization to reduce the interference caused by the different lighting conditions. Then, followed by constructing a four-layer CNN architecture, consist of 2 convolutional layers with 2 subsampling layers. In the last step, for multi-classification, they used the Softmax classifier for classification. For training and testing, they used both JAFFE and CK+ datasets to evaluate and simulate the recognition performance of the presented study. The accuracy performance of the proposed study achieved 76.7442% and 80.303% for JAFFE and CK+, respectively.For facial expression recognition, another approach proposed [13]based on the use of the Gabor filter features. After extracting regions of interest (ROI) Gabor features are extracted. Then, to select the most relevant features the PCA technique has been exploited. In classification, they used a Support Vector Machine (SVM) classifier for the classification of basic FE. Experiments have been conducted on both JAFFE and CK + datasets, the accuracy rate achieved 95.11% and 92.19% respectively. However, the limitation of the proposed study is the Gabor and SVM parameters are chosen manually. Automatic recognition of FER is considered as an actively combining research in recognizing emotions of facial expression. Therefore, the study of [14] automatic recognition of facial expression is considered as an actively combining research in recognizing emotion. Therefore, this study proposed to extend the DCNN approach to the FER task. Thus, based on identifying the events of action units of facial is considered as a portion of the system of facial action coding this task could be done. In the fully connected layers of CNN, they used up a regularization method known as (dropout) that is effective to reduce overfitting. In evaluation, they used CK+ datasets to evaluate their experiments. The expert of action units has validated as well as annotated the used dataset. The performance evaluation of the introduced study has been measured by using the ground truth of the dataset. They achieved an accuracy rate of 92.81% for the proposed study.In the study of [15], the authors proposed a method is to identify the entire distance learning process for students' understanding based on facial recognition. This study proposed a learning emotion recognition model to detect the input image, and they used the Haar Cascades method for eyes and mouth extraction. Then by using the NN classifier in training, different six emotional categories are obtained. Also applying Mean and median filters on image to make it smoother by removing unwanted noise. JAFFE database is used in evaluation for Experiments, the rate of accuracy reached around 87%. In the FER field, the dimensionality of features and selecting the most relevant features are very important steps. Because of a massive amount of memory as well as demanding preprocessing to process a whole image. Thus, the study of [16] proposed a method to overcome this problem, they used facial landmark identification for extracting features, and for classification, the CNN technique has been used as a classifier. To evaluate the performance of the proposed study, they used MUG, JAFFE, MMI, and CK datasets. In their evaluation, 80% of datasets are utilized in the training stage whereas 20% of them have been utilized in the testing stage. The accuracy rate of performance each of MUG, JAFFE, MMI, and CK datasets was 85%, 84%, 89%, and 85% respectively.

III. CONVOLUTION NEURAL NETWORK

The structure of CNNs has many layers, namely, as convolutional, pooling, rectified linear units, and fully connected [8, 9], [17]. Figure 1 shows an example of the structure of the CNN technique. The size of the input has been minimized to the end layers. Through sequential layers, and features are extracted from the low-level to the high-level [18].

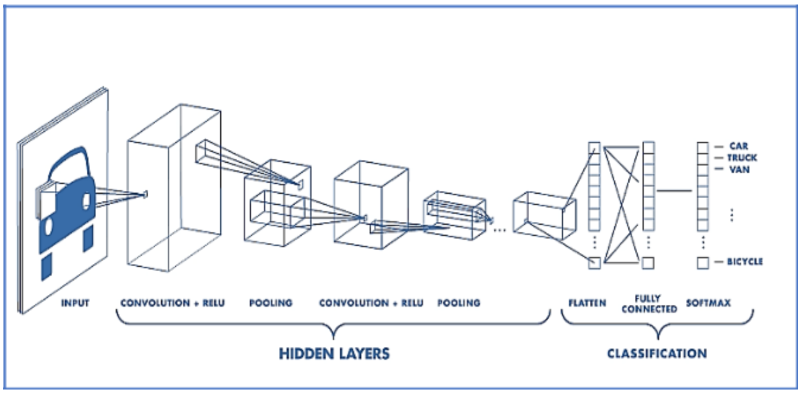


Figure 1: Architecture of CNN[16]

- Convolutional layer: This layer-based CNN technique deals with convoluting each image as input with kernels (filters). Once the output image has been obtained from this layer, it will be used as an input image for the second layer.
- Pooling layer or in other words down sampling is an operator to reduce the dimensionality of each activation map while retaining the most significant information, for example, maximum or maybe average.
- Rectified Linear Units ReLu: is a linear activation function, it is used the rectifier, f(x)= max(0,x) that will output the input directly if is positive and reconstitutes all the negative values by zero.
- Fully connected layer: multilayer perceptron is the final output layer in CNN, which could connect all neurons of the lower layer to each neuron of its layer. The features submit to the fully connected layers after it was alternated between the convolutional and subsampling operations.

IV. IMAGE PRE-PROCESSING

The first step in preprocessing an image is face detection. The Haar-like feature is one of the most classic features for face detection. Originally it was proposed in 1998 by Papageorgiou et al. [20], [21], and Haar-like feature known as the rectangular feature. As stated by Viola and Jones [22], they proposed real-time object detection algorithms to detect ER of human face detection. The reduced image scale is an important task that assists to minimize the information that could be learned by the network, additionally, it has the ability to make the training process faster as well as cost with less memory.

V. DATASETS

In our proposed method, the evaluation of experimental analysis has been evaluated on a popular FER dataset to show the performance of the proposed study. The used dataset called Japanese Female Facial Expression (JAFFE), and the extended dataset is called CohnKanade (CK+). The JAFFE dataset consists of 215 (grayscale) images of facial expression. This dataset holds seven different emotional categories of facial expression, namely, neutral, surprise, disgusting, smile, fear, sad, and anger which are taken from ten Japanese females [3]. The size of each image is 256 * 256 image resolution pixels, different emotion categories from JAFFE dataset for facial expression have been shown in Figure 2.

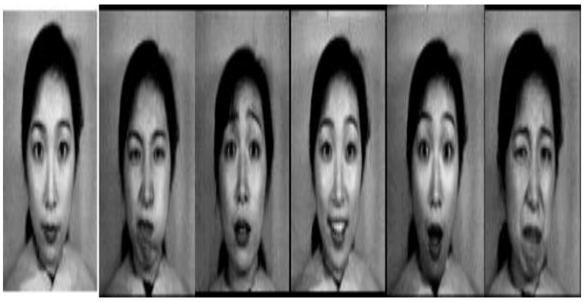


Figure 2: examples of JAFFE datasets

However, the CK+ dataset contains 123 subjects and 593 sequences images. Participants were of males and females, their age ranges between18 to 30 years. The resolution pixel of images is 640 x 480 or 640 x 490 [20]. different emotion categories from the CK+ dataset for facial expression have been shown in Figure 3.

Figure 3: examples of CK+ datasets

VI.PROPOSED METHOD

The proposed method for FER is based on the architecture of CNN which is able to recognize the emotion categories of facial expressions. The proposed FER system as shown in process diagram Figure 4, is divided into three different stages. The first stage is face detection by using the viola jones algorithm which is considered an important stage. The second stage is Feature extraction and selection of the most relevant features using PCA. In the final stage, the extracted features have been fed to the CNN technique for classification. In this study, the face has been detected, then the wanted region has been cropped which is used to extract from the unwanted region, after that, it has been utilized as input to CNN technique. This step has been done because using parts from facial (wanted region) as input inception layer could increase the probability of extracting a high level of features instead of using unwanted regions. More so, this step could reduce the time of training thereby increasing system accuracy

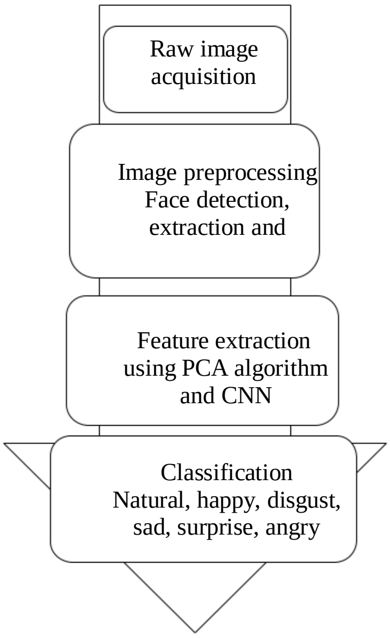


Figure 4: process diagram of the proposed FER system.

VII. Feature exraction using PCA

One of the common techniques in data science for extracting features is called Principle Component Analysis (PCA). This technique is considered as a way of reducing data dimensionality, where the process of this technique is data dimensional from high to low [23], [24]. In the data, the PCA captures big principal variability while disregarding the small variability. Additionally, the dataset dimensionality could be reduced by discovering a new set of variables that are considered as a smaller set of variables instead of the original one. The main aim of PCA, in high dimensional (n) dataset is finding the directions of maximum variance and project it onto a smaller subspace of dimensional with preserving most of the information [25]. Particularly, the process of the PCA technique utilizes an orthogonal transformation to transfer a set of variables that are potentially correlated with M into a set of variables that are uncorrelated with K are named as principal components.

The steps of the PCA algorithm [27], [28] as following below:

Step 1- read the data of each image face into the library NXN.
Step 2- calculate the average face (Ψ).

Step 3- the variation between the face from each image and the average face has been acquired $Ø_i = Γ_i – Ψ$.

Step 4- construct the covariance matrix C = ATA, Where A = [Ø1, Ø2, Ø3…, ØM].

Step 5- calculate the eigenvalues and eigenvectors of the covariance matrix.

Step 6- eigenvalues arrangement from largest to smallest M.

Step 7- keep the largest k eigenvectors (these are the k principal components) K < M.

Step 8- convert the data into a new space constructed by k eigenvectors.

Step 9- export processed data.

CNN is employing the concepts of weight sharing and receptive field by using these concepts, were reduced the number of trainable parameter and calculate the information that propagation through the layers by using convolution. Convolved the signal with a filter map with containing the shared weights to generate a feature map.

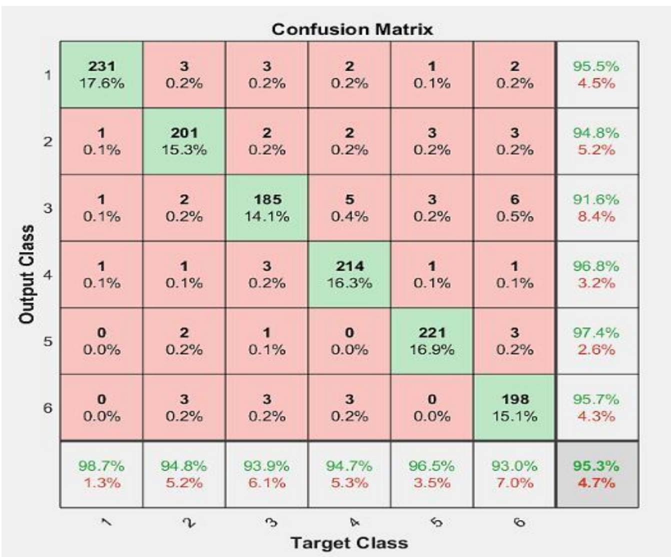
A grayscale image size 277*277 is passed through the architecture of CNN, which is consists of 4 layers of convolutional, 4 layers of max-pooling, 4 layers of dropout, and 6 fully connected layers. The first layer of CNN implements a convolution 16 kernel of size 5*5*1 with stride [13] and the ReLU activation function, then, it is followed by the layer of max-pooling with the kernel of size 2*2 and stride [25, 23]. Where this initiation layer is learned 16 various kernels and 16 outputs of feature maps. After that, the second layer of the CNN technique implements a new convolution to each of the obtained 16 output of feature maps from the previous layer by utilizing 32 various filters of 5*5*16 size followed by another max-pooling using a filter of size 2*2 as well as the stride of size [29]. Then, the 3rd CNN layer applies new convolution using 64 filter size 5*5*32, 2*2 max-pooling layer, and [29] of stride. The last convolution 4th one which 128 kernels of size 5*5*64 with the max-pooling size of 2*2 and stride [25]. Finally, the output is reunited to a fully connected layer in which all neurons of the former layer are connected to all neurons of its layer. Finally, classification has been done based on the maximum probability of class.

VIII. EXPREMINTAL RESULTS

The proposed system is runs on Intel(R) Core (TM) i5-7200 CPU @ 2.50GHz 2. GHz, RAM 8GB and developed on the MATLAB 2018 software. The performance evaluation of the proposed study has been done in term of accuracy based on two different datasets, namely, JAFFE and CK+ datasets. For classification datasets the network is trained using 70% of datasets for the training while 30% of the datasets for the testing of both CK+ and JAFFE datasets, Equation (1) has been utilized for accuracy implementation. The confusion matrix shows in Table 1 the recognition accuracy for six emotions of facial expression on the JAFFE dataset. More so, the accuracy of facial expression recognition for CK+ dataset based on confusion matrix has been presented in Table 2

$$= \frac{\text{Number of correct predictions}}{\text{Total number of predictions made}} \quad (1)$$

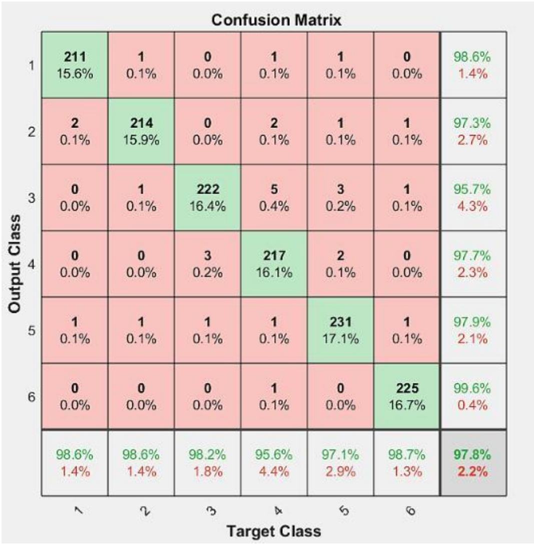
Table I. JAFFE Confusion Matrix



Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	
1	231 17.6%	3 0.2%	3 0.2%	2 0.2%	1 0.1%	2 0.2%	95.5% 4.5%
2	1 0.1%	201 15.3%	2 0.2%	2 0.2%	3 0.2%	3 0.2%	94.8% 5.2%
3	1 0.1%	2 0.2%	185 14.1%	5 0.4%	3 0.2%	6 0.5%	91.6% 8.4%
4	1 0.1%	1 0.1%	3 0.2%	214 16.3%	1 0.1%	1 0.1%	96.8% 3.2%
5	0 0.0%	2 0.2%	1 0.1%	0 0.0%	221 16.9%	3 0.2%	97.4% 2.6%
6	0 0.0%	3 0.2%	3 0.2%	3 0.2%	0 0.0%	198 15.1%	95.7% 4.3%
	98.7% 1.3%	94.8% 5.2%	93.9% 6.1%	94.7% 5.3%	96.5% 3.5%	93.0% 7.0%	95.3% 4.7%

Table II. CK+ Confusion Matrix


Confusion Matrix

Output Class \ Target Class	1	2	3	4	5	6	
1	211 15.6%	1 0.1%	0 0.0%	1 0.1%	1 0.1%	0 0.0%	98.6% 1.4%
2	2 0.1%	214 15.9%	0 0.0%	2 0.1%	1 0.1%	1 0.1%	97.3% 2.7%
3	0 0.0%	1 0.1%	222 16.4%	5 0.4%	3 0.2%	1 0.1%	95.7% 4.3%
4	0 0.0%	0 0.0%	3 0.2%	217 16.1%	2 0.1%	0 0.0%	97.7% 2.3%
5	1 0.1%	1 0.1%	1 0.1%	1 0.1%	231 17.1%	1 0.1%	97.9% 2.1%
6	0 0.0%	0 0.0%	0 0.0%	1 0.1%	0 0.0%	225 16.7%	99.6% 0.4%
	98.6% 1.4%	98.6% 1.4%	98.2% 1.8%	95.6% 4.4%	97.1% 2.9%	98.7% 1.3%	97.8% 2.2%

To show the efficiency of our proposed study, we have compared our study with some of the previous studies in the literature as illustrated in Table 3. Based on our comparison, it is obvious that our study outperformed all previous studies, where the performance evaluation of the study of [29], [30] have been done using different datasets. However, the study of [14],

evaluates their study based on using the same of our dataset, as it is shown in Table 3, our proposed study obtained higher accuracy results. Therefore, it is obvious that our study is efficient in recognizing emotions of facial expression for JAFFE and CK+ datasets. *Table III. Comparison with other methods*

Reference	Method	Datasets	Accuracy
[29]	Multi-scale CNN	FER2013	72.82 %
[31]	Deep Neural Networks	CK+ JAFFE	73.38% 45.07%
[14]	deep CNN	CK+	92.81%
[30]	Deep Neural Network Driven Feature Learning	MULTI-PIE BU3D-FE	85.2% 80.1%
Proposed method	deep CNN	CK+ JAFFE	98.5% 95.5%

IX. CONLUSION

Facial expression recognition (FER) based on different emotions has been used in data science widely. In this study, we present a new technique for FER using a deep learning technique, namely, CNN. In this study, a suitable architecture of CNN has been selected to deal with recognizing emotions of facial expression that consists of convolutional layer numbers, kernel and stride sizes, and pooling structures. For feature extraction, we have used the PCA algorithm. Additionally, for classification, we have applied a machine learning method which is called the KNN algorithm. While the experimental constructed on both of JAFFE database and the Cohn-Kanade database and rate of an average recognition was achieved by 98.5%.

REFERENCES

1. T. Zhang, “Facial Expression Recognition Based on Deep Learning: A Survey,” *Proc. 2nd Int. Conf. Intell. Interact. Syst. Appl.*, vol. 1, no. 1, pp. 1–7, 2017.
2. C. M. M. Refat and N. Z. Azlan, “Deep Learning Methods for Facial Expression Recognition,” *2019 7th Int. Conf. Mechatronics Eng. ICOM 2019*, pp. 1–6, 2019.
3. D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and D. A. Zebari, "Trainable Model Based on New Uniform LBP Feature to Identify the Risk of the Breast Cancer," *in 2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019: IEEE, pp. 106-111.
4. L. Nwosu, H. Wang, J. Lu, I. Unwala, X. Yang, and T. Zhang, “Deep Convolutional Neural Network for Facial Expression RecognitionUsing Facial Parts,” *Proc. - 2017 IEEE 15th Int. Conf. Dependable, Auton. Secur. Comput. 2017 IEEE 15th Int. Conf. Pervasive Intell. Comput. 2017 IEEE 3rd Int. Conf. Big Data Intell. Compu*, vol. 2018-Janua, pp. 1318–1321, 2018.
5. I. Leonid, K. Vladimir, and L. Anton, “Facial Expressio Recognition Algorithm Based on Deep Convolution Neural Network,” 2017.
6. M. Mohammadpour, “Facial Emotion Recognition using Deep Convolutional Networks,” pp. 17–21, 2017.
7. R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction," Journal of Applied Science and Technology Trends, vol. 1, no. 2, pp. 56-70, 2020.
8. F. Kong, “Facial expression recognition method based on deep convolutional neural network combined with improved LBP features,” 2019.
9. D. A. Zebari, D. Q. Zeebaree, A. M. Abdulazeez, H. Haron, and H. N. A. Hamed, "Improved Threshold Based and Trainable Fully Automated Segmentation for Breast Cancer Boundary and Pectoral Muscle in Mammogram Images," *IEEE Access*, 2020.
10. Y. Zhai, J. Liu, J. Zeng, and V. Piuri, “Deep Convolutional Neural Network for Facial Expression Recognition,” vol. 10667, pp. 211– 223, 2017.
11. X. Huynh, T. Tran, and Y. Kim, “Convolutional Neural Network Models for Facial Expression Recognition Using BU-3DFE Database,” vol. 514, 2019.
12. K. Shan, J. Guo, W. You, D. Lu, and R. Bie, “Automatic facial expression recognition based on a deep convolutional-neural-network structure,” *Proc. - 2017 15th IEEE/ACIS Int. Conf. Softw. Eng. Res. Manag. Appl. SERA 2017*, pp. 123–128, 2017.
13. B. Adil, “A novel approach for facial expression recognition,” *2019 Int. Conf. Netw. Adv. Syst.*, pp. 1–5, 2019.
14. C. Series, “Emotion recognition from facial expression using deep convolutional neural network,” 2019.
15. D. Yang, A. Alsadoon, P. W. C. Prasad, A. K. Singh, and A. Elchouemi, “ScienceDirect Procedia Computer Science An EmotionRecognition Model Based on Facial Recognition in Virtual Learning Environment,” *Procedia Comput. Sci.*, vol. 125, no. 2009, pp. 2–10, 2018.
16. N. P. Gopalan, “Facial Expression Recognition using Geometric Landmark Points and Convolutional Neural Networks,” *2018 Int. Conf. Inven. Res. Comput. Appl.*, no. Icirca, pp. 1149–1153, 2018.

17. A. Verma, P. Singh, J. Sahaya, and R. Alex, “Modified Convolutional Neural Network Architecture Analysis for Facial Emotion Recognition,” pp. 169–173, 2019.

18. D. Q. Zeebaree, A. M. Abdulazeez, O. M. S. Hassan, D. A. Zebari, and J. N. Saeed, "Hiding Image by Using Contourlet Transform" *TEST Engineering & management*, vol. 83, 24 May, pp. 16979 - 16990, 2020.

19. “A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way.” [Online]. Available: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53. [Accessed: 18-Dec-2019].

20. C. P. Papageorgiou, M. Oren, and T. Poggio, “A General Framework for Object Detection,” *Comput. vision, 1998. sixth Int. Conf.*, vol. 0, no. January, pp. 555–562, 1998.

21. Zeebaree, D. Q., Haron, H., & Abdulazeez, A. M. (2018, October). Gene selection and classification of microarray data using convolutional neural network. In 2018 International Conference on Advanced Science and Engineering (ICOASE) (pp. 145-150). IEEE.

22. A. Chellappa, M. S. Reddy, R. Ezhilarasie, S. Kanimozhi Suguna, and A. Umamakeswari, “Robust Real-Time Face Detection,” *Int. J. Eng. Technol.*, vol. 7, no. 2, pp. 29–32, 2018.

23. A. L. Machidon, O. M. Machidon, and P. L. Ogrutan, “Face Recognition Using Eigenfaces , Geometrical PCA Approximation and Neural Networks,” *2019 42nd Int. Conf. Telecommun. Signal Process.*, pp. 80–83, 2019.

24. S. Wang, “Feature Extraction and Face Recognition Algorithm,” *2017 13th Int. Conf. Nat. Comput. Fuzzy Syst. Knowl. Discov.*, no. 1, pp. 1910–1914, 2017.

25. X. Dxu, P. Jeudu, and S. Jpdlo, “a comparative review of various approaches for feature extraction in face recognition,” pp. 2705– 2710, 2016.

26. D. Q. Zeebaree, H. Haron, A. M. Abdulazeez, and D. A. Zebari, "Machine learning and Region Growing for Breast Cancer Segmentation," *in 2019 International Conference on Advanced Science and Engineering (ICOASE)*, 2019: IEEE, pp. 88-93.

27. R. Kaur, “Face Recognition Using Principal Component Analysis,” no. 2012, pp. 585–589, 2015.

28. R. M. Ebied, “Feature Extraction using PCA and Kernel-PCA for Face Recognition,” pp. 72–77.

29. E. Chen, Y. Tie, and Y. Gong, “Facial Expression Recognition with Multi-scale Convolution Neural Network,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9916 LNCS, pp. 376–385, 2016.

30. T. Zhang, W. Zheng, Z. Cui, Y. Zong, J. Yan, and K. Yan, “A Deep Neural Network-Driven Feature Learning Method for Multi-view Facial Expression Recognition,” *IEEE Trans. Multimed.*, vol. 18, no. 12, pp. 2528–2536, 2016.

31. G. Wen, Z. Hou, H. Li, D. Li, L. Jiang, and E. Xun, “Ensemble of Deep Neural Networks with Probability-Based Fusion for Facial Expression Recognition,” *Cognit. Comput.*, vol. 9, no. 5, pp. 597– 610, 2017.