



Deep Neural Networks for Automatic Facial Expression Recognition

Venkata Srinivasu Veeram^{1*}, Suban Ravichandran², Gatram Rama Mohan Babu³

¹ Department of Information Technology, Faculty of Engineering and Technology, Annamalai University, Chidambaram 608002, Tamilnadu, India

² Department of Information Technology, Annamalai University, Chidambaram 608002, Tamilnadu, India

³ Department of Computer Science & Engineering (AI&ML), R.V.R. & J.C. College of Engineering, Guntur 522019, Andhra Pradesh, India

Corresponding Author Email: vasuveeram@gmail.com

<https://doi.org/10.18280/ria.360520>

ABSTRACT

Received: 14 July 2022

Accepted: 18 October 2022

Keywords:

facial emotion recognition, conventional facial expression recognition; deep learning-based facial expression recognition

Out of all non-linguistic communications, one of the most popular is face expression and is capable of communicating effectively with others. We have number of applications of facial expressions in as sorted arenas comprising of medicine like psychology, security, gaming, Classroom communication and even commercial creativities. Owing to huge intra-class distinction it is still challenging to recognize the emotions automatically based on facial expression though it is a vigorous area of research since decades. Conventional lines for this approach are dependent on hand-crafted characteristics like Scale Invariant Feature Transform, Histogram of Oriented Gradient and Local Binary Patterns surveyed by a classifier which is applied on a dataset. Various types of architectures were applied for restored performance as Deep learning proved an outstanding feat. The goal of this study is to create a deep learning model on automatic facial emotion recognition FER. The proposed model efforts more on pulling out the crucial features, thereby, advances the expression recognition accuracy, and beats the competition on FER2013 dataset.

1. INTRODUCTION

In human communication 33% is done orally and 67% is done through nonverbal components according to various studies [1, 2]. Facial expressions are the most common and important mode of interpersonal communication. Mehrabian [1] classified the representation of emotions to be visual as 56%, vocal as 36% and verbal as the remaining 8%. First and important sign that transmits the emotion during a conversation is variations in facial expression that is reason which made the researchers attracted by this modality.

To categorize emotion detection various methodologies used are Major technologies that are used in emotion detection can be categorized as follows. Vision based affect detection, Posture based affect detection, Speech based affect detection, and Text based affect detection. Speech, gestures, text, facial expressions, blood, pulse, volume, and other features are all applied in emotion identification.

In this present work contrast is on vision-based technology. Facial expression can be Positive and Negative. As per Albrecht et al. [3] expressions are classified into two types namely basic expressions and non-basic expressions. Basic expressions include Fear, Anger, Sadness, Disgust, Happiness and Surprise. Where as the non-basic expressions are Boredom, Irritation, Despair, Shame, Excitement and Panic.

Facial Expression Recognition (FER) attempts to automatically recognize the facial expression by analysing the facial feature changes. Most of the facial recognition systems identify facial features by extracting landmarks from the subject's facial image. It outputs the information about the facial expression recognized so that it can be used further to

identify the person's mood. The methodologies used to recognise facial expressions are divided into two categories: image-based and model-based.

Image based approach extracts features from images without any prior knowledge about the object of interest. Using the given set of data, process of predicting the class for a data is said to be termed as classification. The types of learners in classification can be of two types: lazy and eager. Lazy learners wait with training data until testing data appear. Usually they take less training time, but more predicting time e.g. k-Nearest Neighbour classifier. Eager learners construct a model using training data before testing data appears. Usually they take long training time, but less predicting time. To verify the applicability of the classifier, many methods are available. The common methods used for this are hold out and cross validation. In holdout method, given set of data will be divided according to the ratio of 80:20 (training: testing). Training samples will be used to build a model whereas the testing samples will be used to test the predicting power of the model. The snowballing curiosity for person expression recognition for several areas like online games, teachers can use student's facial expressions to determine their learning condition, human-computer interfaces [4], animation [5], medicine [6, 7], security [8, 9], diagnosis of ASD in children [10].

Facial expressions [11-13], language [14], electroencephalogram [15] are the various characteristics utilized for emotion recognition [16]. This paper mainly focuses on various FER techniques with three major steps namely pre-processing, feature extraction and classification. Also it focuses on image-based FER techniques to solve problems like are chosen. Mostly FER systems meet the

problems of variation in lighting variations, illumination, skin tone variations and pose variation.

Deep neural networks [17] gives the maximum pull out for characteristics for adequate recognition of facial expressions [18, 19]. In Section 2 discusses the usage of Deep learning techniques adopted by various researches.

2. LITERATURE SURVEY

Deep CNN for FER [20] across quite a lot of accessible databases was proposed by Mollahosseini et al. [21]. In the proposed model augmentation data technique was applied on the images after extracting the facial landmarks. Convolution layers are applied locally to increase the local performance as well as reduce the overfitting problem.

Lopes et al. [22] discuss the effect of pre-processing data prior to training the model for a healthier expression recognizer. Before applying convolutional neural networks to the input, the data is already under went through expansion, spin correction, sampling with 2 completely coupled layers with 256 and 7 neurons. Author shows that combining CK+, JAFFE, BU-3DFE Pre-processing procedures are more efficient. The pre-processing processes used by Mohammadpour et al. [23] in which CNN model is discussed with two convolutional layers followed by max pooling which designate the count of action units stimulated. A convolution neural network planning in which 2 convolution layers are used successively in the beginning followed by sparse batch normalization to fit the model without overflow [24].

Li et al. [25] discuss the facial occlusion problem with a modified method of CNN, in which a training of convolution neural network with a procedure of automatic CNN is applied to the VGG-Net network. Real-world Affective Faces Database, FED-RO and Affect-Net are the databases used for training and testing the architecture.

Yolcu et al. [26] proposed a method for exposure of the crucial parts of the face. To detect mouth, eyebrow, and eye the author used three CNN each one to spot a part of the face. Images are passed through crop stage and the detection before applying CNN. Author concluded this method of acquainting the impressive feature gained from subsequent type of convolutional neural network for distinguishing emotion, as a better performer. Agrawal et Mittal [27] discuss the distinct influence of convolution neural network constraints on the emotion acknowledgment using FER 2013 database. CNN contains 2 successive convolution layers in which the max pooling layer achieve average of 65.77% accuracy. Jain et al. [28] proposed a Deep Convolutional Neural Network with two residual blocks, with four layers in each. Kim et al. [29] proposed a blend of convolutional neural network and Long Short-Term Memory for a spatio-temporal model. Also, in reference [30] an architecture called Spatio-Temporal convolutional with Nested LSTM, which is constructed on three deep learning networks for multi-lateral features.

3. DEEP NEURAL NETWORKS FOR FACE EXPRESSION RECOGNITION SYSTEM

One of the most important fields in the man-machine interface is emotion recognition via facial expressions. Facial ornaments, non-uniform illuminations, position fluctuations, and other factors complicate emotion identification.

Traditional techniques to emotion detection have the problem of feature extraction and classification being mutually optimised. Deep learning techniques are being used more frequently by researchers to address this issue. Deep-learning techniques are now becoming increasingly essential in categorization challenges.

Over the past span in contempt of traditional facial identification models researchers have focused to the deep learning approach due to its high automatic recognition capacity. In this context we proposed a face expression recognizer using deep learning which goes in a particular stream involving the important stages as data pre-processing, feature extraction and then the classification of emotions respectively.

Pre-processing involves various steps such as cropping, scaling, normalization, and face alignment.

For better performance before feature extraction pre-processing is applied in the facial expression recognizer. By applying cropping and scaling process on the face image, nose will be considered as the centre point. Sampling is used for reduction of image size by preserving the features of original image. Smoothing of the image is done by Gaussian filter. For reducing the illumination and variations normalization is done as a part of pre-processing on facial images. Much intelligibility to the given input images is observed by normalization method used for extraction. Localization is a pre-processing method and it uses the Viola-Jones algorithm. Using scale invariant feature transform flow algorithm face alignment is done as a part of pre-processing. In facial expression recognizer, the ROI segmentation process is more satisfactory as it identifies the facial organs accurately which are important for expression recognition.

Feature extraction is performed on images after pre-processing, and with the help of helper function the features are extracted. Once the features extraction is done, we build a DNN model for classification of the expressions. In this proposed model we will first train the model and then test the model for a set of images to recognize the emotion and there by the sentiment or feeling of a person.

The proposed method with deep neural network model to obtain amended exposure of emotion is shown in Figure 1.

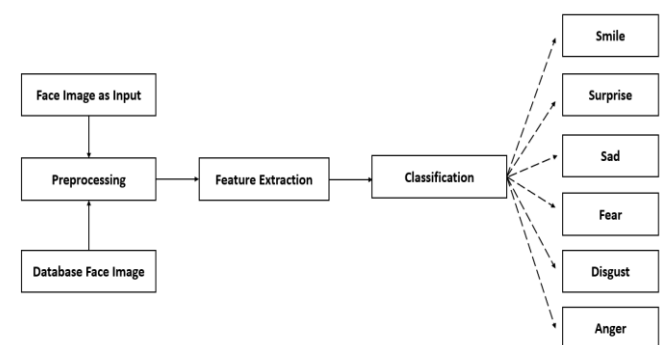


Figure 1. Face expression recognition system

In the proposed model we are using a sequential model method in keras to create our model for emotion detection, we are using dense, dropout, flatten, Con2D, and Maxpooling2D layers together to build a basic model that can actually be trained to classify various emotions. We are applying the following deep learning models for emotion classification as Random forest, Logistic regression, Support Vector Machine and Voting Classifier.

Support Vector Machines are a type of maximal margin hyperplane classification approach that guarantees strong generalisation performance by utilising the findings of statistical learning theory. SVMs have a high classification accuracy even when there is only a limited quantity of data available for training, which makes them particularly suited to a dynamic and interactive strategy for expression identification. Because of the often-subtle differences that exist between distinct expressions in our displacement-based data, such as "anger" and "disgust," as well as the wide range of possible variations in a particular expression when performed by different subjects, we decided to use SVMs as our primary method of classification.

Features extracted from the CNN model when combined with random forest classifier is a well-suited method. The realistic technique is to put the random forest into the last pool layer as the characteristic of the novel data will undergo some characteristic loss after passing through every layer of convolution neural network model which can be seen in Figure 2.

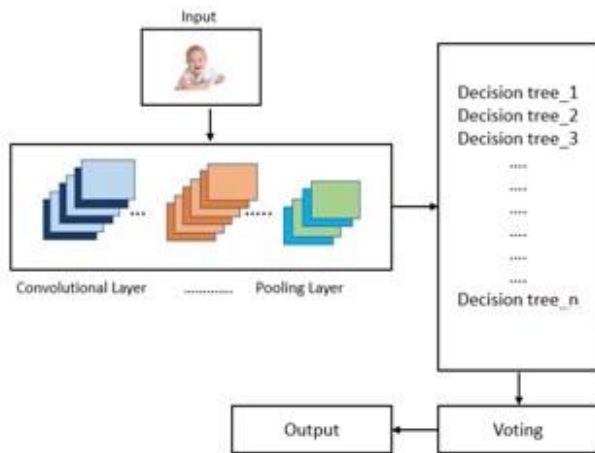


Figure 2. The structure of the new model

The first step in developing this model was to acquire convolutional neural network (CNN) features, and the second step was to integrate the CNN features with an upgraded random forest so that face expressions could be classified.

Cai et al. [24] discuss about the composition of multiple decision trees to form random forest classifier [RF]. The

ultimate result is resolved by voting on randomly selected decision trees.

A prospect selection-based method to regulate all the attained decision trees to reach the all requirements of virtuous and variety is proposed in this paper.

Improvement over conventional decision tree-based classifier is made by random forest classifier by overpowering some confines they have. Most importantly random forest method addresses the problem of overfitting by maintain good accuracy for both training and testing data, as well as it handles the missing data in a better way.

4. DATABASE

Numerous Facial Expression Recognizer datasets are now reachable to the scientists to fulfil the task, with variations in the count and dimensions of images, distinctions of the radiance, count and expression posture. Table 1 summarizes the FER databases used by various researchers in their works.

FER2013 (KAGGLE):



Figure 3. Sample images from the FER dataset for angry

Wu et al. [15] discusses about the dataset for facial expression recognition (FER-2013) presented in ICML 2013, with 35,888 images with a resolution of 48x48. Formerly, 28,709 images and 3589 images are present in training and test data respectively. The faces in database are robotically enumerated using Google image search API. The faces are labelled as any of the six cardinal expressions as well as neutral. FER dataset is composed with facial occlusion, unfinished faces, low divergence images and even with faces having eyeglasses.

Angry, disgust and happy sample images are shown in Figure 3, 4 and 5 respectively from FER database.

Table 1. A summary of some FER databases

Databases	Descriptions	Emotions
MMI [31]	2900 videos are used to identify the neutral, onset, apex, and offset.	Six primary emotions and one neutral feeling
MultiPie [32]	More than 750,000 photos were collected by 15 view and 19 illumination settings.	Anger, Disgust, Neutral, Happy, Squint, Scream, Surprise
SFEW [33]	700 photos with varying ages, occlusion, lighting, and head poses.	Six primary emotions and one neutral feeling
GEMEP FERA [34]	289 images sequences	Anger, Fear, Sadness, Relief, Happy
FER2013 [35]	Google image search yielded 35,887 grayscale images.	Six primary emotions and one neutral feeling
CK+ [36]	593 videos for posed and non-posed expressions	Six basic emotions, contempt and neutral
CASME II [37]	247 micro-expressions sequences	Surprise, Happy, Regression Disgust and others
JAFFE [38]	10 Japanese females posed for 213 grayscale photographs.	Six primary emotions and one neutral feeling
BU-3DFE [39]	2500 3D facial images captured on two view -45°, +45°	Six primary emotions and one neutral feeling
RAFD-DB [40]	30000 images from real world	Six primary emotions and one neutral feeling
Oulu-CASIA [41]	2880 videos captured in three different illumination conditions	Six basic emotions
AffectNet [42]	More than 440.000 images collected from the internet	Six primary emotions and one neutral feeling

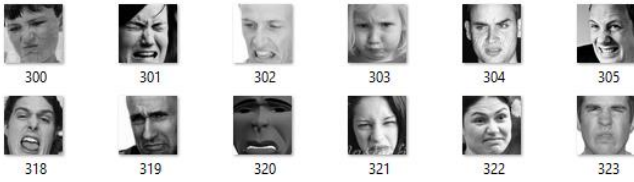


Figure 4. Sample images from the FER dataset for disgust

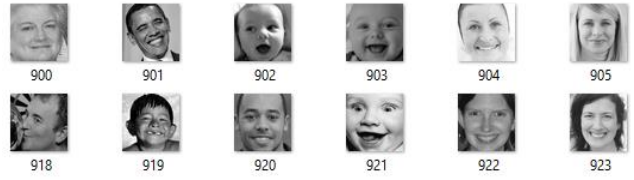


Figure 5. Sample images from the FER dataset for happy

5. RESULTS AND DISCUSSIONS

This proposed method presents a multiple Deep learning model based on voting mechanism of facial expression recognition method, to attain all categories of model of the time, understand the model of decision fusion.



Figure 6. Experimental results in FER2013

In Figure 6 each classifier in deep learning techniques [43] is trained on every region of interest for five times, every time by using twenty percent of data as test set and the remaining eighty percent of data as training set. Consequently, every split is considered as test set for single time and for the remaining four times as a training set. For the proposed model the FER-2013 dataset is considered which comprises of 35,887 distinct images, out of which 28,709 samples are considered as training set.

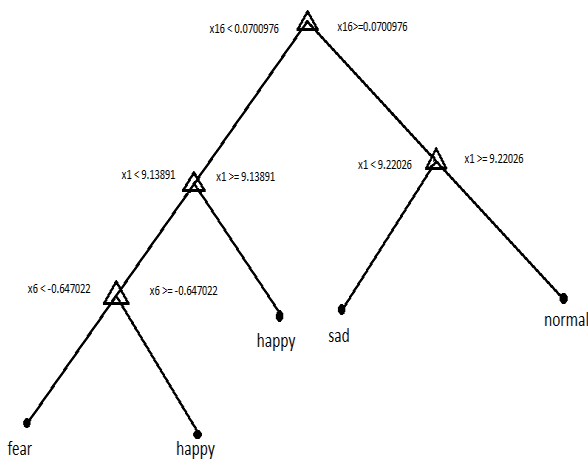


Figure 7. Generated decision tree models

The total dataset is classified as public and private, among which 3589 examples of public test is exploited for picking of optimal CNN model whereas to verify the accuracy rate

examples of private test are exploited. Decision tree models that were engendered in our progress is shown in Figure 7.

Table 2 depicts the various classifier performance results and it is observed that with Logistic regression model we got very low train accuracy which is only 31% and test accuracy as 20% it's very difficult to train such a data. If we look into the result, the test set accuracy is very low, it is a 25%. Logistic regression and SVM almost seem to be the same level of test accuracy. And coming to a random forest, we can look into the result is that we have accuracy of 47% on test data set which is really very good result. Next, we combined all the models using the voting classifier. We have the test score of 40% Accuracy on test data set which less than random forest. So, in the proposed work we found random forest classifier to be more effective for facial expression recognition.

Table 2. Classifier performance

Classifier	Accuracy train	Accuracy test	F1 score train	F1 score test
Logistic regression	0.31	0.23	0.23	0.18
Support Vector Machine	0.42	0.25	0.32	0.17
Random Forest	0.99	0.47	0.99	0.46
Voting Classifier	0.84	0.40	0.87	0.70

6. CONCLUSION

In this paper, we describe a fully deep neural network model for facial emotion recognition, which has been tested on public datasets to evaluate its performance. The work summarized the performance evaluations of the various classifiers. This work has a good amount of scope for improvement with the help of hyperparameter tuning. Deep learning for facial emotion recognition could be the effective initiation for many of the expression-based applications like online games, costumer feedbacks, learning status of students in online classes and many more. Facial expressions are simply one of many aspects that need to be encoded for human expressive behaviours in realistic applications. Combining many models into a high-level framework, which can provide supplementary information, may further boost robustness. Pure expression identification utilising only visible facial photos can produce promising results. This is a fascinating study area that also aids in the future study for clarifying the concept related to the facial expressions approaches with the aid of deep learning technique because of the strong complementarity between facial expressions and other modalities, such as infrared images, depth data from the 3D face replicas, and physiological data will take for good research in the field of deep learning.

REFERENCES

- [1] Mehrabian, A. (2017). Communication without words. In Communication theory, 193-200. Routledge.
- [2] Kaulard, K., Cunningham, D.W., Bülthoff, H.H., Wallraven, C. (2012). The MPI facial expression

- database-a validated database of emotional and conversational facial expressions. *PloS One*, 7(3): e32321. <https://doi.org/10.1371/journal.pone.0032321>
- [3] Albrecht, I., Schröder, M., Haber, J., Seidel, H.P. (2005). Mixed feelings: expression of non-basic emotions in a muscle-based talking head. *Virtual Reality*, 8(4): 201-212. <https://doi.org/10.1007/s10055-005-0153-5>
- [4] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., Taylor, J.G. (2001). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, 18(1): 32-80. <https://doi.org/10.1109/79.911197>
- [5] Aneja, D., Colburn, A., Faigin, G., Shapiro, L., Mones, B. (2016). Modeling stylized character expressions via deep learning. In *Asian Conference on Computer Vision*, pp. 136-153. https://doi.org/10.1007/978-3-319-54184-6_9
- [6] Edwards, J., Jackson, H.J., Pattison, P.E. (2002). Emotion recognition via facial expression and affective prosody in schizophrenia: A methodological review. *Clinical Psychology Review*, 22(6): 789-832. [https://doi.org/10.1016/S0272-7358\(02\)00130-7](https://doi.org/10.1016/S0272-7358(02)00130-7)
- [7] Chu, H.C., Tsai, W.W. J., Liao, M.J., Chen, Y.M. (2018). Facial emotion recognition with transition detection for students with high-functioning autism in adaptive e-learning. *Soft Computing*, 22(9): 2973-2999. <https://doi.org/10.1007/s00500-017-2549-z>
- [8] Clavel, C., Vasilescu, I., Devillers, L., Richard, G., Ehrette, T. (2008). Fear-type emotion recognition for future audio-based surveillance systems. *Speech Communication*, 50(6): 487-503. <https://doi.org/10.1016/j.specom.2008.03.012>
- [9] Saste, S.T., Jagdale, S.M. (2017). Emotion recognition from speech using MFCC and DWT for security system. In *2017 International Conference of Electronics, Communication and Aerospace Technology (ICECA)*, 1: 701-704. <https://doi.org/10.1109/ICECA.2017.8203631>
- [10] Leo, M., Carcagni, P., Distante, C., Spagnolo, P., Mazzeo, P.L., Rosato, A.C., Lecciso, F. (2018). Computational assessment of facial expression production in ASD children. *Sensors*, 18(11): 3993. <https://doi.org/10.3390/s18113993>
- [11] Meng, Q., Hu, X., Kang, J., Wu, Y. (2020). On the effectiveness of facial expression recognition for evaluation of urban sound perception. *Science of The Total Environment*, 710: 135484. <https://doi.org/10.1016/j.scitotenv.2019.135484>
- [12] Mollahosseini, A., Chan, D., Mahoor, M.H. (2016). Going deeper in facial expression recognition using deep neural networks. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1-10. <https://doi.org/10.1109/WACV.2016.7477450>
- [13] Liu, P., Han, S., Meng, Z., Tong, Y. (2014). Facial expression recognition via a boosted deep belief network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1805-1812.
- [14] Han, K., Yu, D., Tashev, I. (2014). Speech emotion recognition using deep neural network and extreme learning machine. In *Interspeech 2014*.
- [15] Wu, C.H., Chuang, Z.J., Lin, Y.C. (2006). Emotion recognition from text using semantic labels and separable mixture models. *ACM Transactions on Asian Language Information Processing (TALIP)*, 5(2): 165-183. <https://doi.org/10.1145/1165255.1165259>
- [16] Marechal, C., Mikolajewski, D., Tyburek, K., Prokopowicz, P., Bougueroua, L., Ancourt, C., Wegrzyn-Wolska, K. (2019). Survey on AI-based multimodal methods for emotion detection. *High-Performance Modelling and Simulation for Big Data Applications*, 11400: 307-324.
- [17] LeCun, Y. (1989). Generalization and network design strategies. *Connectionism in perspective*, 19: 143-155.
- [18] Khorrami, P., Paine, T., Huang, T. (2015). Do deep neural networks learn facial action units when doing expression recognition? In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 19-27.
- [19] Tzirakis, P., Trigeorgis, G., Nicolaou, M.A., Schuller, B.W., Zafeiriou, S. (2017). End-to-end multimodal emotion recognition using deep neural networks. *IEEE Journal of Selected Topics in Signal Processing*, 11(8): 1301-1309. <https://doi.org/10.1109/JSTSP.2017.2764438>
- [20] Zahara, L., Musa, P., Prasetyo, E., Karim, I., Musa, S. (2020). The facial emotion recognition (FER-2013) dataset for prediction system of micro-expressions face using the convolutional neural network (CNN) algorithm based raspberry Pi. *2020 Fifth International Conference on Informatics and Computing (ICIC)*, pp. 1-9. <https://doi.org/10.1109/ICIC50835.2020.9288560>
- [21] Mollahosseini, A., Chan, D., Mahoor, M.H. (2016). Going deeper in facial expression recognition using deep neural networks. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1-10. <https://doi.org/10.1109/WACV.2016.7477450>
- [22] Lopes, A.T., De Aguiar, E., De Souza, A.F., Oliveira-Santos, T. (2017). Facial expression recognition with convolutional neural networks: coping with few data and the training sample order. *Pattern Recognition*, 61: 610-628. <https://doi.org/10.1016/j.patcog.2016.07.026>
- [23] Mohammadpour, M., Khalilardali, H., Hashemi, S.M.R., AlyanNezhadi, M.M. (2017). Facial emotion recognition using deep convolutional networks. In *2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI)*, pp. 0017-0021. <https://doi.org/10.1109/KBEI.2017.8324974>
- [24] Cai, J., Chang, O., Tang, X. L., Xue, C., Wei, C. (2018). Facial expression recognition method based on sparse batch normalization CNN. In *2018 37th Chinese Control Conference (CCC)*, pp. 9608-9613. <https://doi.org/10.23919/ChiCC.2018.8483567>
- [25] Li, Y., Zeng, J., Shan, S., Chen, X. (2018). Occlusion aware facial expression recognition using CNN with attention mechanism. *IEEE Transactions on Image Processing*, 28(5): 2439-2450. <https://doi.org/10.1109/TIP.2018.2886767>
- [26] Yolcu, G., Oztel, I., Kazan, S., Oz, C., Palaniappan, K., Lever, T.E., Bunyak, F. (2019). Facial expression recognition for monitoring neurological disorders based on convolutional neural network. *Multimedia Tools and Applications*, 78(22): 31581-31603. <https://doi.org/10.1007/s11042-019-07959-6>
- [27] Agrawal, A., Mittal, N. (2020). Using CNN for facial expression recognition: A study of the effects of kernel size and number of filters on accuracy. *The Visual Computer*, 36(2): 405-412. <https://doi.org/10.1007/s00371-019-01630-9>

- [28] Jain, D.K., Shamsolmoali, P., Schdev, P. (2019). Extended deep neural network for facial emotion recognition. *Pattern Recognition Letters*, 120: 69-74. <https://doi.org/10.1016/j.patrec.2019.01.008>
- [29] Kim, D.H., Baddar, W.J., Jang, J., Ro, Y.M. (2017). Multi-objective based spatio-temporal feature representation learning robust to expression intensity variations for facial expression recognition. *IEEE Transactions on Affective Computing*, 10(2): 223-236. <https://doi.org/10.1109/TAFFC.2017.2695999>
- [30] Yu, Z., Liu, G., Liu, Q., Deng, J. (2018). Spatio-temporal convolutional features with nested LSTM for facial expression recognition. *Neurocomputing*, 317: 50-57. <https://doi.org/10.1016/j.neucom.2018.07.028>
- [31] Pantic, M., Valstar, M., Rademaker, R., Maat, L. (2005). Web-based database for facial expression analysis. In 2005 IEEE International Conference on Multimedia and Expo, 5. <https://doi.org/10.1109/ICME.2005.1521424>
- [32] Gross, R., Matthews, I., Cohn, J., Kanade, T. Baker, S. (2008). Multi-PIE. 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition. <https://doi.org/10.1109/AFGR.2008.4813399>
- [33] Dhall, A., Goecke, R., Lucey, S., Gedeon, T. (2011). Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark. In 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), pp. 2106-2112. <https://doi.org/10.1109/ICCVW.2011.6130508>
- [34] Valstar, M.F., Jiang, B., Mehu, M., Pantic, M., Scherer, K. (2011). The first facial expression recognition and analysis challenge. In 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG), pp. 921-926. <https://doi.org/10.1109/FG.2011.5771374>
- [35] Goodfellow, I.J., Erhan, D., Carrier, P.L., Courville, A., Mirza, M., Hamner, B., Bengio, Y. (2013). Challenges in representation learning: A report on three machine learning contests. In International Conference on Neural Information Processing, pp. 117-124. https://doi.org/10.1007/978-3-642-42051-1_16
- [36] Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I. (2010). The extended Cohn-Kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In 2010 IEEE Computer Society conference on Computer Vision and Pattern Recognition-Workshops, pp. 94-101. <https://doi.org/10.1109/CVPRW.2010.5543262>
- [37] Yan, W.J., Li, X., Wang, S.J., Zhao, G., Liu, Y.J., Chen, Y.H., Fu, X. (2014). CASME II: An improved spontaneous micro-expression database and the baseline evaluation. *PloS One*, 9(1): e86041. <https://doi.org/10.1371/journal.pone.0086041>
- [38] Lyons, M.J., Akamatsu, S., Kamachi, M., Gyoba, J., Budynek, J. (1998). The Japanese female facial expression (JAFFE) database. In Proceedings of Third International Conference on Automatic Face and Gesture Recognition, pp. 14-16. <https://doi.org/10.5281/zenodo.3451524>
- [39] Yin, L., Wei, X., Sun, Y., Wang, J., Rosato, M.J. (2006). A 3D facial expression database for facial behavior research. In 7th International Conference on Automatic Face and Gesture Recognition (FGR06), pp. 211-216. <https://doi.org/10.1109/FGR.2006.6>
- [40] Li, S., Deng, W., Du, J. (2017). Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2852-2861.
- [41] Zhao, G., Huang, X., Taini, M., Li, S. Z., Pietikäinen, M. (2011). Facial expression recognition from near-infrared videos. *Image and Vision Computing*, 29(9): 607-619. <https://doi.org/10.1016/j.imavis.2011.07.002>
- [42] Mollahosseini, A., Hasani, B., Mahoor, M.H. (2017). Affectnet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, 10(1): 18-31. <https://doi.org/10.1109/TAFFC.2017.2740923>
- [43] Veeram, V.V., Ravichandran, S., Babu, G.M. (2022). Deep neural networks for face recognition and feature extraction from multi-lateral images. *International Journal of Computer Science and Network Security*, 22(4): 700-704.