JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ADVANCED DEEP LEARNING TECHNIQUES FOR FACIAL EMOTION RECOGNITION

M.AnushaRani[1*], B. PrasannaLakshmi[2], D.Hemalatha[3], N.AkhilaRani[4], A.Likitha[5].1*

Assistant professor: Vignan's Nirula Institute of Technology and Science for Women[2,3,4,5] B. Tech Scholar: Vignan's Nirula Institute of Technology and Science for Women

ABSTRACT:

Facial emotion recognition (FER) has become increasingly significant in fields such as human-computer interaction, mental health assessment, and social robotics. This paper explores the application of deep learning techniques, particularly Convolutional Neural Networks (CNNs), to enhance the accuracy and efficiency of emotion detection from facial expressions. We examine various datasets, including FER-2013 and CK+, which provide the necessary labelled images for effective model training and evaluation. The discussion also addresses key challenges, such as data imbalance, cultural differences in emotional expression, and the impact of environmental factors like lighting and occlusion on performance. The development of more robust and adaptable emotion recognition systems, paving the way for their integration into real-world applications across multiple domains.

1.INTRODUCTION:

The human life and work are developed by the application called information technology using Artificial Intelligence (AI) technique. Nowadays, a typical person is developed with AI in the form of monitor screens, security control systems, commanding the functions for social work by using voices and interface between computers and humans called Human Recognition Face [1-9]. A filtering mechanism for EEG signals is used to shape information about human emotions, where face expression recognition system is one way to do this process [10-21]. Nonverbal communication, including facial expression, is a significant means of conveying one's emotions, intentions, objectives, and ideas to others [21-27]. Facial appearance is the consequence of a facial sign or expression that reveals the location of the human face. Many computers today employ a method of estimating how satisfied a customer will be with a service based on the expressions portrayed on the customer's face. A customer's happiness level as shown by a selection made on a screen may not always be reliable [28-30].

Six human emotions have been identified as universally perceivable without regard to cultural background by psychologists Ekman and Friesen [31-35]. Currently, Deep NNs (DNNs), especially convolution neural networks (CNNs), have drawn attention in Facial Emotion Recognition (FER) by virtue of its inbuilt feature extraction mechanism from images [35-36]. A few works have been reported with the CNN to solve FER problems [37-39]. However, the existing FER methods considered the CNN with only a few layers, although its deeper model is shown to be better at other image-processing tasks. The facts behind this may be the challenges related to FER. Firstly, recognition of emotion requires a moderately high-resolution image, meaning to work out high dimension data. Secondly, the difference in faces due to different emotional states is very low, which eventually complicates the classification task [40-41].

2.LITERATURE SURVEY:

Albert Mehrabian discovered, in 1968, that in person-to-person interactions, verbal indicators account for 7% of all interactions, vocal indications account for 38%, and facial reactions account for 55% [42]**. Ekman and Friesen** [43], Electronics 2021, 10, 1036. And **Ekman** [44] have performed extensive research in human facial expressions and identified universal facial expressions: happiness, sadness, anger, fear, surprise, disgust, and neutral states. **Agrawal, A., Mittal, N**.: There are three phases to the FER system used for expression recognition. Facial recognition is the initial process. This phase focuses towards locating a face in an input image and identifying its constituent parts. Adaboost, Haar-cascade and Histogram of Gradients (HOG) are all good examples of suitable algorithms [45]**. Verma, G., Verma, H**.: The identified face is then subjected to either a geometric feature-based or an appearance feature-based extraction of its distinguishing characteristics. Emotions are then categorized using the method's retrieved characteristics in a classification phase [46].

Facial emotions provide rich non-verbal information in real-time, and their correct interpretation is critical to participation in one's social environment: **Collin et al.,** 2013 [47-48]. Facial emotion recognition (FER) deficits are evident and pervasive across various disorders in children; namely, psychiatric disorders (e.g., schizophrenia/psychosis, mood disorders and anxiety reviewed in **Collin et al.,** 2013, neurodevelopment disorders e.g., autism **Eussen et al.**, 2015; **Evers et al**., 2015; Taylor et al., 2015] and ADHD - although the evidence is mixed Buford et al., 2015, neurological conditions e.g., epilepsy **Edwards et al**., 2017, and acquired brain disorders (e.g., traumatic brain injury Mancuso et al., 2015 and brain tumors **Bonner et al., 2008** [49].

3.PROPOSED METHODOLOGY:

The most widely used method of image analysis is the convolutional neural network (CNN). In contrast to a multilayer perceptron (MLP), CNN has concealed layers known as convolutional layers. The proposed approach is based on the CNN framework with two levels. Background removal, which is utilized to draw out emotions from an image, is the initial stage that is advised. Here, the primary expressional vector is extracted using the standard CNN network module (EV). Finding pertinent facial points of significance leads to the creation of the expressional vector (EV). Changes in expression are closely tied to changes in EV. Using a basic perceptron unit on a facial image with the background removed, the EV is produced. We also include a non-convolutional perceptron layer as the final level in the FERC model that has been proposed.

By following this step-by-step methodology, we aim to address the challenges of facial expression recognition.

3.1 Data Collection:

In 2013, Yang et al. presented the 2013 Facial Expression Recognition dataset (FER2013). This dataset is available through Kaggle. The Facial Emotion Recognition (FER+) dataset is an expansion of the initial FER collection, in which the images were re-labeled as happiness, sorrow, neutral, angry and disgust. Because of its tremendous scientific and business significance, FER is crucial in the domains of computational vision and artificial intelligence. FER is a technique that examines facial movements across passive images, as well as videos, to disclose details about a person's state of mind.

3.2 Data Preprocessing:

- Resize all images to a uniform resolution to mitigate the impact of varying camera qualities.
- Apply histogram equalization to standardize lighting conditions across images.
- Augment the dataset by applying transformations such as rotation, flipping, and cropping to increase sample diversity and robustness

3.3CNN Algorithm:

Convolutional Neural Networks (CNNs) are a type of deep learning algorithm designed for processing image data. They consist of layers that perform convolution operations to extract features, followed by activation functions (like ReLU) to introduce non-linearity. Pooling layers reduce the spatial dimensions, and fully connected layers classify the features into categories. CNNs are efficient in learning hierarchical patterns, making them ideal for tasks like image classification, object detection, and facial recognition. Their ability to automatically learn relevant features and recognize objects regardless of their position has made them essential in computer vision applications.

3.4 Working of CNN:

Step1: Accepts the input images (e.g., grayscale facial images).

Step2: Apply convolution operations using filters to extract features from the images.

Step3: Perform max pooling to down-sample the feature maps, reducing dimensionality while retaining important features.

Step4: Convert the pooled feature maps into a one-dimensional vector.

Step5: A softmax function is applied in the output layer to produce probabilities for each class (e.g., different emotions). The class with the highest probability is chosen as the output.

4.Training Process:

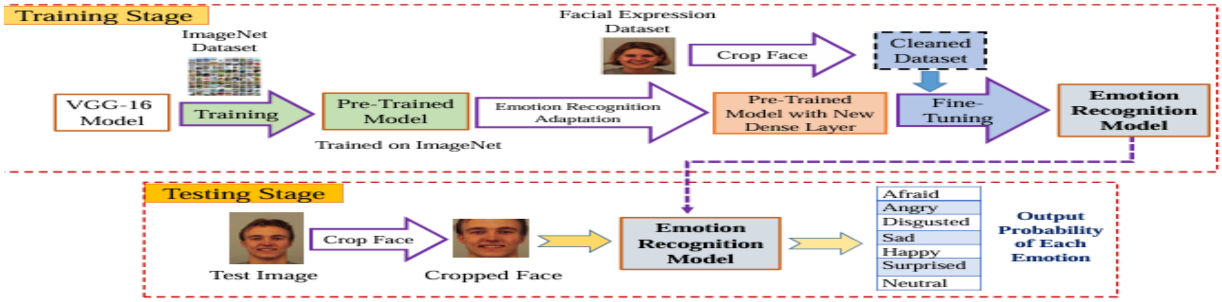

Figure-4.1 : Illustration of the proposed FER system based on transfer learning in deep CNN.

The following pseudocode outlines a basic process for emotion detection using Deep Face.

def predict_emotion_features(image):

 # Load the Deep Face model.

model = load_model("fer.h5")

 # Extract the features of the face in the image.

features = extract_features(image)

 # Predict the emotion features of the face.

emotion_features = model.predict(features)

 # Return the emotion features.

return emotion_features

The Deep Face model to identify emotion on the face is loaded using the load_model() method. The distinctive features of the human face in the image are extracted via the extract_features() method. These features may include the placement of the eyebrows, the contour of the lips, and the appearance of forehead wrinkles. Based on the data that are retrieved, the predict_emotion_features() algorithm forecasts the facial features associated with emotions.

5.Result:

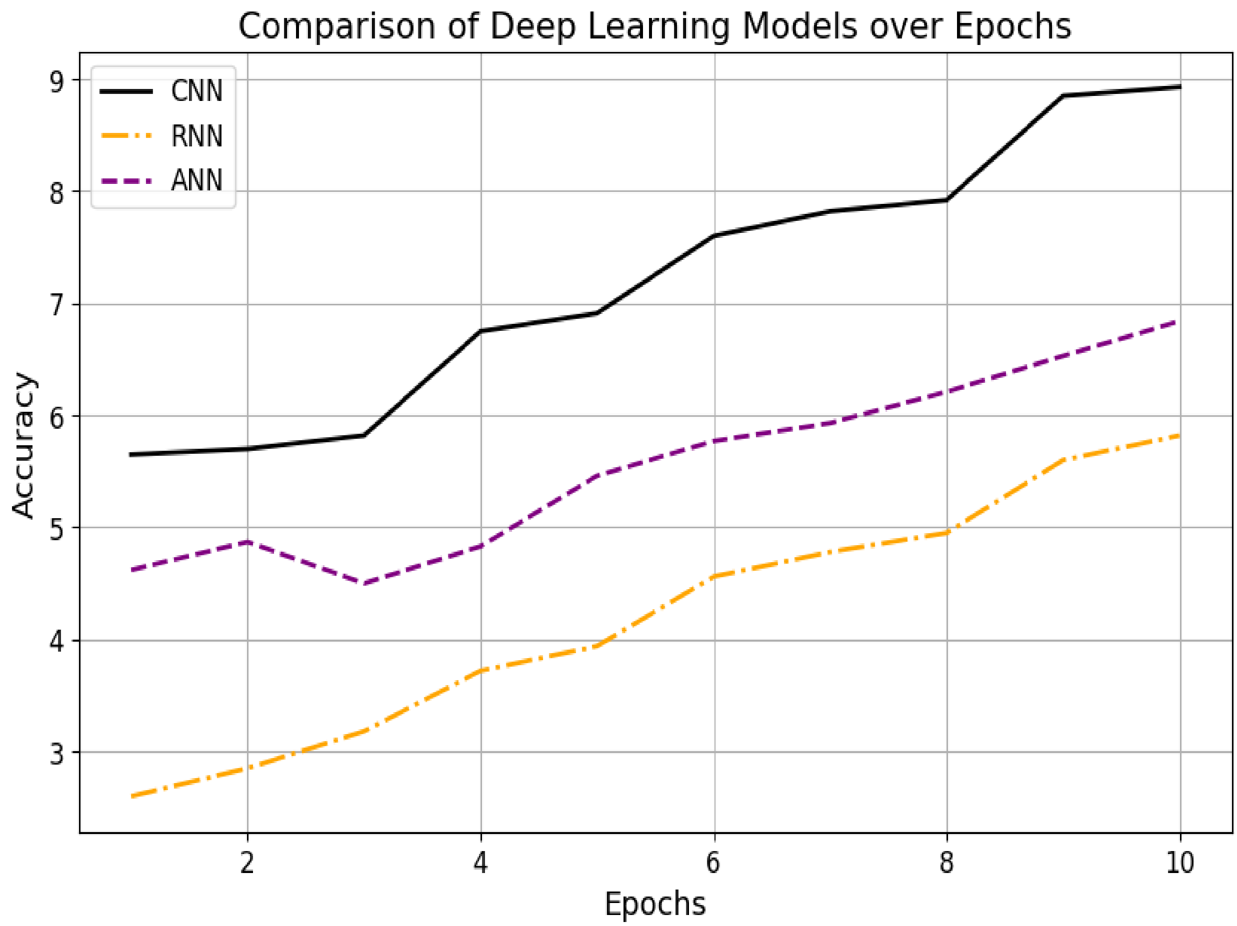

Figure-5.1 : Comparison of Deep Learning Models Over Epochs

The CNN model shows the highest accuracy, starting around 3 and steadily increasing to nearly 9 by the 10th epoch. This indicates that CNN performs significantly better than the other models as training progresses. The RNN model starts at a lower accuracy than CNN, around 4, and gradually increases to about 6. This suggests that while RNN improves over time, it does not reach the performance level of CNN. The ANN model has the lowest accuracy among the three, starting at approximately 3 and rising to around 5.5 by the 10th epoch. This indicates that ANN has the slowest improvement in accuracy compared to CNN and RNN.

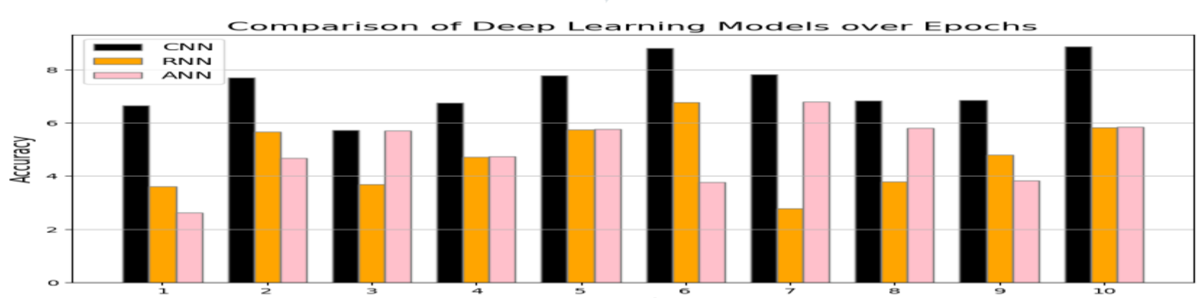

Figure-5.2 : Comparison of Deep Learning Models Over Epoch

X-Axis: Represents the number of epochs, ranging from 1 to 10.

Y-Axis: Represents the accuracy of the models, with values ranging from 0 to 10.

Black Bars: Represent the accuracy of the CNN model.

Orange Bars: Represent the accuracy of the RNN model.

Pink Bars: Represent the accuracy of the ANN model.

The graph allows for a visual comparison of how each model's accuracy changes over the epochs. It appears that the CNN model generally achieves higher accuracy compared to the other models across most epochs. The RNN and ANN models show varying performance, with some epochs where their accuracy is closer to that of the CNN. This visualization is useful for analysing the performance of different deep learning architectures over time, helping to identify which model may be more effective for a given task.

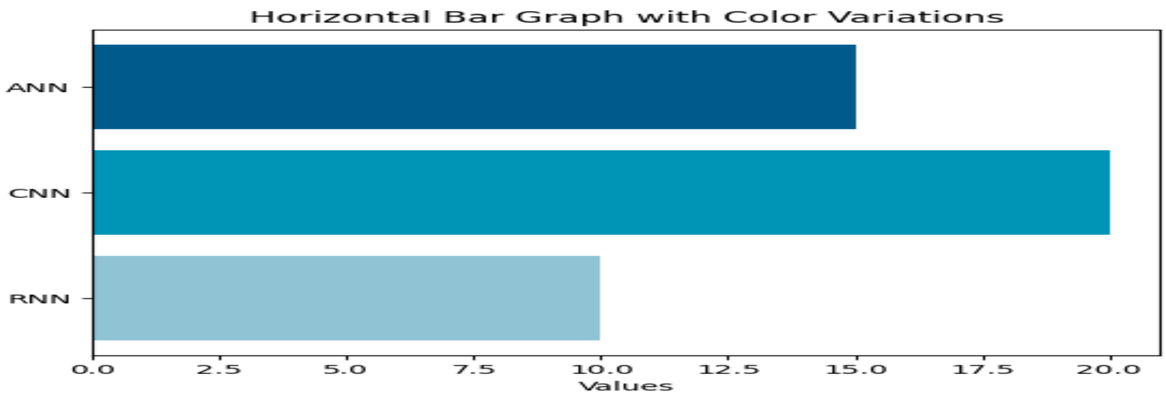

Figure-5.3 : Horizontal Bar Graph with Color Variations

The image shows a horizontal bar graph comparing the performance of three different neural network models—ANN (Artificial Neural Network), CNN (Convolutional Neural Network), and RNN (Recurrent Neural Network)—based on their accuracy or effectiveness in an emotion recognition. By analysing the heights of the bars, one can quickly assess which model performs best. For instance, if the CNN bar is significantly higher than the others, it suggests that CNN is more effective for this task. Understanding which models perform better can guide future research and development efforts, helping to refine algorithms and improve accuracy in facial emotion recognition systems. The use of colour variations enhances readability and helps in distinguishing between the models, making it easier to interpret the results at a glance. Overall, this graph serves as a valuable tool for evaluating and communicating the performance of different neural network architectures in the context of facial emotion recognition.

6.CONCLUSION:

Facial emotion recognition using deep learning has revolutionized the way we interpret human emotions from visual data. Leveraging advanced architectures like convolutional neural networks (CNNs), these systems have achieved remarkable accuracy, often surpassing traditional machine learning techniques. The ability to analyze subtle facial cues opens up diverse applications, including security, where emotion detection can enhance surveillance, and healthcare, where it aids in monitoring patients' emotional well-being. Marketing strategies can also benefit from understanding consumer emotions, allowing for more targeted campaigns. However, the rapid advancement of this technology raises significant ethical concerns, particularly around privacy, consent, and the potential for misuse. It is crucial to establish guidelines to navigate these issues responsibly. Furthermore, future developments in multimodal approaches—integrating facial recognition with audio and textual data—promise even greater accuracy and context. As the field continues to grow, ongoing research and dialogue about ethical implications will be vital.

Ultimately, while deep learning for emotion recognition offers exciting opportunities, a balanced approach is necessary to ensure its positive impact on society.

7.REFERENCES:

[1] Patibandla, R. S. M. L., & Narayana, V. L. (2021). Computational intelligence approach for prediction of COVID-19 using particle swarm optimization. In *Advances in Computational Intelligence and Data Analytics* (Vol. 923, pp. 175–189). Springer. https://doi.org/10.1007/978-981-15-8534-0_9

[2] Santhi Sri, K., Sandhya Krishna, P., Lakshman Narayana, V., & Khadherbhi, R. (2021). Traffic analysis using IoT for improving secured communication. In *Advances in Intelligent Systems and Computing* (Vol. 213, pp. 499–507). Springer. https://doi.org/10.1007/978-981-33-4443-3_48

[3] Narayana, V. L., & Bharathi, C. R. (2019). Multi-mode routing mechanism with cryptographic techniques and reduction of packet drop using 2ACK scheme MANETs. In *Advances in Intelligent Systems and Computing* (Vol. 104, pp. 649–658). Springer. https://doi.org/10.1007/978-981-13-1921-1_63

[4] Narayana, V. L., & Bharathi, C. R. (2018). Effective multi-mode routing mechanism with master-slave technique and reduction of packet droppings using 2-ACK scheme in MANETS. *Mathematical Modeling of Computer Systems*, 91(2), 73–76. https://doi.org/10.18280/mmc_a.910207

[5] Lakshman Narayana, V., Lakshmi Patibandla, R. S. M., Pavani, V., & Radhika, P. (2023). Optimized nature-inspired computing algorithms for lung disorder detection. In *Advances in Intelligent Systems and Computing* (Vol. 1066, pp. 103–118). Springer. https://doi.org/10.1007/978-981-19-6379-7_6

[6] Narayana, V. L., Sudheer, B. N., Maddumala, V. R., & Anusha, P. (2020). Fuzzy base artificial neural network model for text extraction from images. *Journal of Critical Reviews*, 7(6), 350–354. https://doi.org/10.31838/jcr.07.06.61

[7] Narayana, V. L., Bhargavi, S., Srilakshmi, D., Annapurna, V. S., & Akhila, D. M. (2024). Enhancing remote sensing object detection with a hybrid Densenet-LSTM model. In *Proceedings of the International Conference on Computer Science and Advanced Technology* (pp. 264–269). IEEE. https://doi.org/10.1109/IC2PCT60090.2024.10486394

[8] Narayana, V. L., Syamalatha, P., Vatsalya, P., Sricharitha, V., & Akhila, V. (2023). Multi-level node authorization using recurrent neural networks for secure health monitoring system. Proceedings of the IEEE International Conference on Smart Computing and Networking Applications (ICSNA), 1697–1705. https://doi.org/10.1109/ICSCNA58489.2023.10370543

[9] Gopi, A. P., Swathi, V., Harshitha, G. S., Swetha, B., & Alekhya, N. (2023). Prediction of paddy yield based on IoT data using GRU model in lowland coastal regions. In Proceedings of the 5th International Conference on Smart Systems and Inventive Technology (ICSSIT 2023) (pp. 1747-1752). https://doi.org/10.1109/ICSSIT55814.2023.10060935

[10] Arepalli, P. G., Naik, K. J., & Amgoth, J. (2024). An IoT-based water quality classification framework for aqua-ponds through water and environmental variables using CGTFN model. International Journal of Environmental Research, 18(4), Article 73. https://doi.org/10.1007/s41742-024-00625-2

[11] Gopi, A. P., Babu, E. S., Raju, C. N., & Kumar, S. A. (2015). Designing an adversarial model against reactive and proactive routing protocols in MANETs: A comparative performance study. International Journal of Electrical and Computer Engineering, 5(5), 1111-1118. DOI: 10.11591/ijece.v5i5.pp1111-1118

[12] Sravanthi, G. L., Devi, M. V., Sandeep, K. S., Naresh, A., & Gopi, A. P. (2020). An efficient classifier using machine learning technique for individual action identification. International Journal of Advanced Computer Science and Applications, 11(6), 513-520. DOI: 10.14569/IJACSA.2020.0110664

[13] Gopi, A. P., Durga Mani, P., Chandana, V. B., Sulthana, S. R., & Parameswari, P. P. K. (2024). Classification of fake news using enhanced capsule neural network. In 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation, IATMSI 2024. DOI: 10.1109/IATMSI60426.2024.10502837

[14] **Assed, M.M., Khaff, T.C., Belizario, G.O., Fatorelli, R., Rocca, C.C.D.A., de Pádua Serafm, A**.: Facial emotion recognition in maltreated children: a systematic review. J. Child Fam. Stud. 29(5), 1493–1509 (2020).

[15] **Manoharan, J.S.:** Design of an intelligent approach on capsule networks to detect forged images. J. Trend. Compute. Sci. Smart Technol. 3(3), 205–221 (2021).

[16] **Mehendale, N**.: Facial emotion recognition using convolutional neural networks (FERC). SN Appl. Sci. 2(3), 1–8 (2020).

[17] **Alreshidi, A., Ullah, M.:** Facial emotion recognition using hybrid features. Informatics 7(1), 1–8 (2020).

[18] **Li, K., Jin, Y., Akram, M.W., Han, R., Chen, J.: Facial** expression recognition with convolutional neural networks via a new face cropping and rotation strategy. Vis. Comput. 36(2), 391–404 (2020).

[19] **Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M.S.; Hasan, M.; Van Essen, B.C.; Awwal, A.A.S.; Asari, V.K.** A State-of-the-Art Survey on Deep Learning Theory and Architectures. Electronics 2019, 8, 292. [CrossRef] 20. **Sahu, M.; Dash, R**. A Survey on Deep Learning: Convolution Neural Network (CNN). In Smart Innovation, Systems and Technologies; Springer: Singapore, 2021; Volume 153, pp. 317–325.

[20] **Zhao, X.; Shi, X.; Zhang, S**. Facial Expression Recognition via Deep Learning. IETE Tech. Rev. 2015, 32, 347–355. [CrossRef] **Li, J.; Huang, S.; Zhang, X.; Fu, X.; Chang, C.-C.; Tang, Z.; Luo, Z.** Facial Expression Recognition by Transfer Learning for Small Datasets. In Advances in Intelligent Systems and Computing; Springer: Berlin/Heidelberg, Germany, 2020; Volume 895, pp. 756–770.

[21] **Khan, A.; Sohail, A.; Zahoora, U.; Qureshi,** A.S. A survey of the recent architectures of deep convolutional neural networks. Artif. Intell. Rev. 2020, 53, 5455–5516.

[22] **Kolen, J.F.; Kremer, S.C.** Gradient Flow in Recurrent Nets: The Difficulty of Learning Long-term Dependencies. In A Field Guide to Dynamical Recurrent Networks; Wiley-IEEE Press: Hoboken, NJ, USA, 2010; pp. 237–243.

[23] **Simonyan, K.; Zisserman,** A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.

[24] **Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z.** Rethinking the inception architecture for computer vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.

[25] **He, K.; Zhang, X.; Ren, S.; Sun, J**. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

[26] **Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A.** Going deeper with convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.

[27] **Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q.** Densely Connected Convolutional Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269.

[28] **Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M.S.; Hasan, M.; Van Essen, B.C.; Awwal, A.A.S.; Asari, V.K. A** State-of-the-Art Survey on Deep Learning Theory and Architectures. Electronics 2019, 8, 292.

[29] **Kakuba, S.; Poulose, A.; Han, D.S.** Deep learning-based speech emotion recognition using multi-level fusion of concurrent feature. IEEE Access 2022, 30, 125538–125551.

[30] **Chen, Y.; He, J.** Deep learning-based emotion detection. J. Compute. Common. 2022, 10, 57–71.

[31] **Schoneveld, L.; Othmani, A.; Abdelkawy,** H. Leveraging recent advances in deep learning for audio-visual emotion recognition. Pattern Recognit. Lett. 2021, 146, 1–7.

[32] **Agrawal, A., Mittal, N.:** Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy. Vis. Comput. 36(2), 405–412 (2020).

[33] **Verma, G., Verma, H.:** Hybrid-deep learning model for emotion recognition using facial expressions. Rev Socionetwork Strategies 14(2), 171–180 (2020).

[34] **Wang Y., Metoki A., Alm K.H., Olson I.R**. White matter pathways and social cognition. *Neurosci. Biobehave.*

[35] **Vo TH, Lee GS, Yang HJ, Kim SH. Pyramid** with super resolution for in-the-wild facial expression recognition. *IEEE Access.* 2020; **8**:131988–132001. doi: 10.1109/ACCESS.2020.3010018. [CrossRef] [Google Scholar]

[36] **Mehrabian,** A. *Nonverbal communication* (Aldine Transaction, 2007).

[37] **Ekman, P. Darwin, deception, and facial expression.** ***Ann. N. Y. Acad.*** *Sci.***1000,** 205–2 (Kortli & Jridi, 2020) (2006). [PubMed]

[38] **Farzaneh, A. H. & Qi, X.** Facial expression recognition in the wild via deep attentive center loss in *2021 IEEE winter conference on applications of computer vision (WACV)* 2401–2410 (IEEE, 2021).

[39] **Alnuaim AA, et al.** Human-computer interaction for recognizing speech emotions using multilayer perceptron classifier. *J. Healthc. Eng.* 2022; **2022**:6005446. doi: 10.1155/2022/6005446. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

[40] **Kumari, H. M. L. S.** *Facial expression recognition using convolutional neural network along with data augmentation and transfer learning* (2022)

[41] **Baidu; 2021. A. I. Baidu.** https://ai.baidu.com/ [Google Scholar].

[42] Face++ -Face++ Cognitive Services. *Face++* 2021. https://www.faceplusplus.com/.

[43] **Kuo C.-M., Lai S.-H., Sarkis M**. A compact deep learning model for robust facial expression recognition. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); June 2018; Salt Lake City, UT, USA. pp. 2121–2129.

[44] FER-2013.Available online: https://www.kaggle.com/datasets/msambare/fer2013 (accessed on 2 November 2023).

[45] Zhang S, Zhang Y, Zhang Y, Wang Y, Song Z (2023) 'A Dual-Direction Attention Mixed Feature Network for Facial Expression Recognition.' Electronics 12(17):1–15. https://doi.org/10.3390/electronics12173595.

[46] Majety, V. D., & Murali, G. (2018). Remote health watchdog framework for seizure patient using electronic sensors. *International Journal of Engineering and Technology(UAE, 7*, 783–785. https://doi.org/10.14419/ijet.v7i3.12132

[47] Alapati, N., Anusha, N., Joharika, P., Jerusha, N.J., Tanuja, P.(2023)Prediction of Parkinson's Disease using Machine Learning in 2023 2 nd *International Conference on Electronics and Renewable Systems(ICEARS)(pp.1357–1361).IEEE*

[48] *Naresh, A., Reddy, B.A., Reddy, G.P., Kumari, K.R., Vaishnavi, M.S.(2023)Melanocytic Pigmented Skin Lesion Detection and Classification using Hybrid Deep Features based on Fully Convolutional Network* in 2023 2 nd *International Conference on Electronics and Renewable Systems(ICEARS)(pp.1011–1018).IEEE*

[49] Pavani, Vellalacheruvu, and I. Ramesh Babu. "Three level cloud storage scheme for providing privacy preserving using edge computing." *International Journal of Advanced Science and Technology* 28, no. 16 (2019): 1929-1940.