ISSN: 2633-4828 Vol. 5 No.2, June, 2023

International Journal of Applied Engineering & Technology

FACIAL EMOTION RECOGNITION TECHNIQUE BASED ON CLAHE, BILATERAL FILTER, AND DEEP LEARNING

Naveen Kumari
Assistant Professor, Department of Computer Science, Punjabi University Regional Centre for Emerging and Innovative Technology, Mohali (Punjab)
Email id: naveen_rcitm@pbi.ac.in
DOI: https://doi.org/10.5281/zenodo.22724052

ABSTRACT

Facial emotion recognition (FER) is a technique that identifies human emotions through facial expressions in images. This study introduces a new technique that enhances image quality using contrast-limited adaptive histogram equalization (CLAHE) and bilateral filtering, followed by classification using a deep convolutional neural network (CNN). The proposed approach achieved a maximum accuracy of 96.24% on the JAFFE dataset and comparative analysis reveals that the proposed facial emotion recognition technique achieves better results than the competitive models in terms of accuracy.

Keywords: *Emotions, Contrast-limited adaptive histogram equalization, Bilateral filter, Deep neural network*

1. INTRODUCTION

Facial Emotion Recognition (FER) is a subfield of computer vision and artificial intelligence that seeks to automatically detect human emotional states through facial expressions. A core ability of human intelligence is the extraction of visual information from images (Zhang et al., 2019). As artificial intelligence grows, it is believed that robots are able to comprehend human emotions. However, identifying emotions efficiently is an ill-posed problem (Gao et al., 2019). There are several ways of detecting human emotions, for instance by identifying faces, voices or physiological indicators. Speech and facial expressions are natural ways to recognize emotions (Zhang et al. 2016). Facial emotion recognition (FER) is a technique that identifies human emotions through facial expressions in images. Machine learning techniques enable data-driven analysis and predictive classification across different application domains (Taqvi, 2025). FER (facial Emotion Recognition) is an important problem in computer vision that seeks to recognize human emotional states from facial expressions. Applications include human computer interaction, intelligent surveillance, health care, education, robotics and emotional computing. Traditional FER approaches often rely on hand-crafted face characteristics. Deep learning methods may automatically learn discriminative features directly from the facial photos. Recent studies (S. Li et al 2022 and H. Ge et al 2022) have shown that convolutional neural networks (CNNs) and other deep neural architectures are effective for face emotion recognition. Deep convolutional neural networks are used for a lot of tasks in the area of computer vision.

However, changes of lighting, poor contrast, picture noise, facial expression and backdrop variables may degrade the FER performance. Therefore, an effective preprocessing stage is helpful prior to feature extraction and classification. In this paper, the suggested method utilizes the Contrast Limited Adaptive Histogram Equalization (CLAHE) and bilateral filtering for enhancing the quality of face pictures and then recognition of facial emotions using Deep Convolutional Neural Network (DCNN).

The Main contributions of this study are:

1. The facial emotion recognition technique is proposed.
2. The CLAHE (Contrast Limited Adaptive Histogram Equalization) is applied on the input photos to enhance the visibility of the images.
3. The bilateral filter is used to remove the noise.

4. A Deep convolutional network is used for emotion classification.
5. The JAFFE dataset is used for emotion classification.

The rest of the paper is organized as follows. Section 2 discusses related work. In section 3 provides an enhanced emotion recognition technique. Section 4 shows results and comparative analysis. Concluding remarks are given in Section 5.

2. RELATED WORK

Podder et al. (2023) developed a feature-boosted deep-learning approach with feature-boosting modules and skip connections to highlight expression-specific face characteristics. The suggested method has obtained 96.16% accuracy on the JAFFE dataset. Ahmad et al. (2023) suggested a model based on Stacked Sparse Auto-Encoder (SSAE-FER). Jeong et al. (2022) suggested a multi-head cross-attention network for facial expression recognition, emphasizing the significance of attention-based feature learning for real-world FER. Xue et al. (2022) suggested a coarse-to-fine cascaded network with smooth prediction for differentiating similar expressions in movies. Ngai et al. (2022) increased facial recognition utilizing a two-channel EEG and optical modality. Hung & Chang (2021) applied multilevel transfer learning for facial emotion recognition. Vijaya Lakshmi and Mohanaiah (2021) employed WOA and TLBO with MultiSVNN for classification. EEG and text fusion using SVM Zhang et al. (2020) employed CNN with emotion label distribution learning while Chen et al. (2018) used deep sparse autoencoders with softmax regression.

Many emotion identification approaches have been created and implemented to recognize human emotions in the literature (Hua et al. 2019). Li and Deng (2019) developed an artificial neural network (ANN) model for facial emotion recognition. Zhang et al., (2019) constructed spatial–temporal recurrent neural network to recognize the facial emotions. Kim et al. (2019) extracted adaptable facial features by hierarchical deep learning (HDL) to achieve superior outcomes. The existing literature shows that the decision tree (DT) (Lee et al. 2011 and Sun et al. 2019), support vector machine (SVM) (Varma et al. 2020 and Kar et al. 2019), random forest (RF) (Valstar et al.2016 and Puetal.2015), and artificial neural networks (ANNs) (Li and Deng 2019) are widely used to recognize human emotion. Jain et al. (2019) constructed deep neural networks (DNNs) with deep residual blocks. Wang et al. (2020) used a hybrid of CNN and RNN (CCNNRNN) for human emotion classification. kumari and bhatia (2020) discussed comparative study and analysis of various facial emotion recognition techniques .

Gupta et al. (2020) proposed a human emotion recognition model based on ResNet and an attention block (CRAB). Lakshmi and Ponnusamy (2021) extracted the features using the modified histogram of orientated gradients (HOG) and local binary pattern (LBP). Comparative study and analysis of various facial emotion recognition techniques discussed (kumari and bhatia 2020).

Wang et al. (2020) merged face and voice characteristics using CNN, RNN and LSTM utilizing decision-level fusion. Li and Lima (2021) used ResNet-50 for better robustness. Deng et al. (2019) employed cGANs, whereas Du et al. (2020) integrated CNN-based face features with Bi-LSTM heartbeat signals. Li et al. (2019) combined EEG and facial expressions to perform continuous emotion recognition. It has been found from the related work that most of the existing approaches produce good performance but suffer from the over-fitting issue. In addition, the existing models are not robust for images with poor visibility and noise.

3. PROPOSED METHODOLOGY

In this paper a facial emotion recognition technique is proposed for recognizing human emotions from facial images. The suggested technique may dynamically focus on important characteristics in the images during the training phase. Fig. 1 shows the workflow of the proposed technique for emotion recognition.

The main goal of the proposed technique is to increase recognition performance under challenging circumstances, including inadequate illumination, low contrast and remove noise. The facial emotion recognition dataset is first put into the MATLAB workspace. The input images are enhanced with CLAHE to boost the contrast and visibility. The bilateral filter is applied to remove noise but keep the important facial features. Then the preprocessed dataset is separated into training and testing sets. A deep convolutional neural network (DCNN) is trained using training images where convolution, batch normalization, ReLU activation and max-pooling operations are conducted in sequence. These processes are repeated in five convolutional layers to extract deep face features. Then the retrieved features are sent into a fully connected layer. The softmax and classification layers are used to detect the relevant emotion. Finally, the trained model is tested on the testing dataset with relevant metrics for performance.

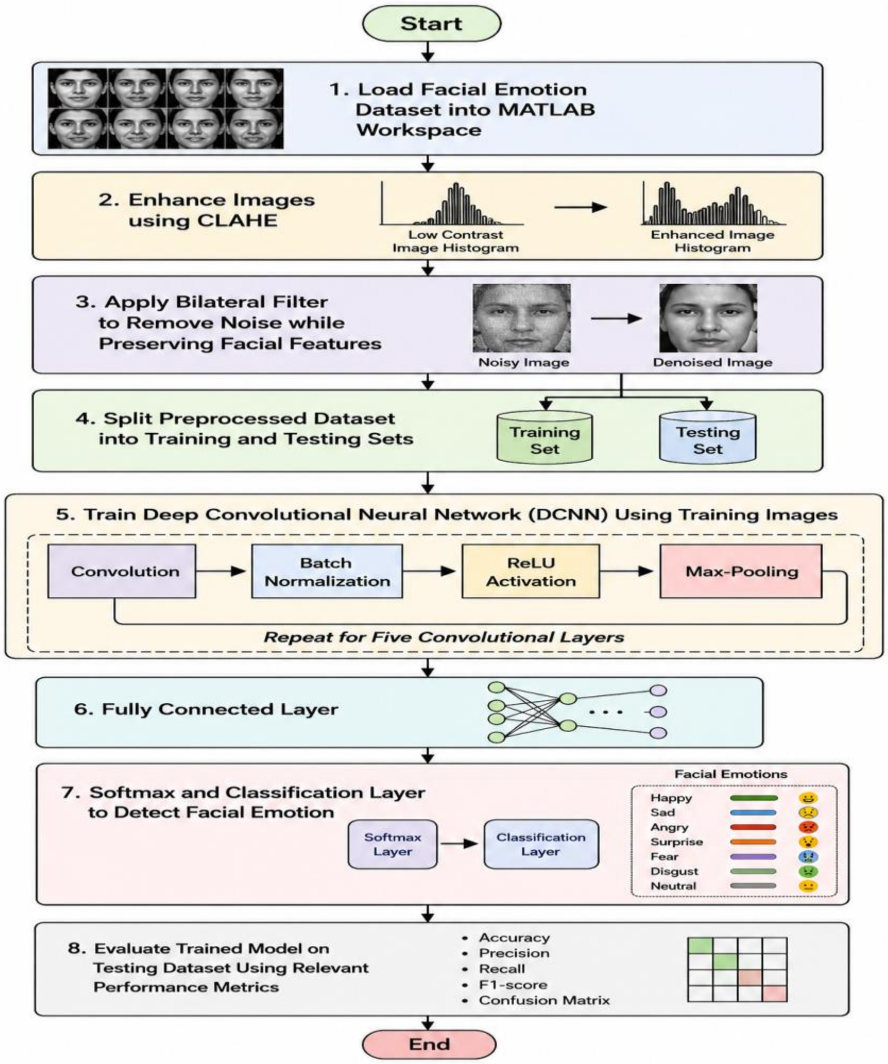

Figure 1: Workflow of the proposed technique for emotion recognition.

The proposed technique is applied to the JAFFE (M. J., Akamatsu et al.1998) face emotions dataset which is the most popular testbed for emotion recognition. The human emotions which are universally acknowledged are as follows: anger, neutral, disgust, happy, fear, surprise and sadness. Table 1 shows the JAFFE dataset emotion distribution and number of photos per class. The JAFFE (Japanese Female facial Expression) collection consists of 213 face photos of 10 Japanese female individuals with 7 facial emotions. The classes are labeled from 1 to 7 correspondingly.

Table1: JAFFE dataset emotion distribution

S. No.	Emotion	Number of Images

1	Anger	30
2	Neutral	30
3	Disgust	29
4	Fear	32
5	Happy	31
6	Sadness	31
7	Surprise	30
	Total	**213**

4. RESULTS

This section evaluates the performance of the proposed facial emotion recognition technique using different parameters like accuracy, precision, recall, F-measures. Table 2 presents evaluation metrics for the proposed facial emotion recognition technique.

Table 2: Evaluation metrics for the proposed facial emotion recognition technique

S. No.	Performance Metric	Formula	Description
1	Accuracy	(TP + TN) / (TP + TN + FP + FN)	Computes the overall proportion of correctly categorized facial images by the suggested model.
2	Precision	TP / (TP + FP)	Calculates the proportion of photographs accurately classified among all the images identified as a specific facial emotion.
3	Recall (Sensitivity)	TP / (TP + FN)	Measures the suggested model's ability to accurately identify photos pertaining to a certain facial emotion.
4	F1-Score	2 × (Precision × Recall) / (Precision + Recall)	It provides a balanced measure of precision and recall and is useful when emotion classes contain varied quantities of images.
5	Specificity	TN / (TN + FP)	Measures the ability of the model to accurately recognize images that are not part of a certain facial emotion.
6	Confusion Matrix	—	It shows the real and anticipated emotion classes and helps to identify the facial expressions that have been classified correctly and inaccurately.

Note: TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative.

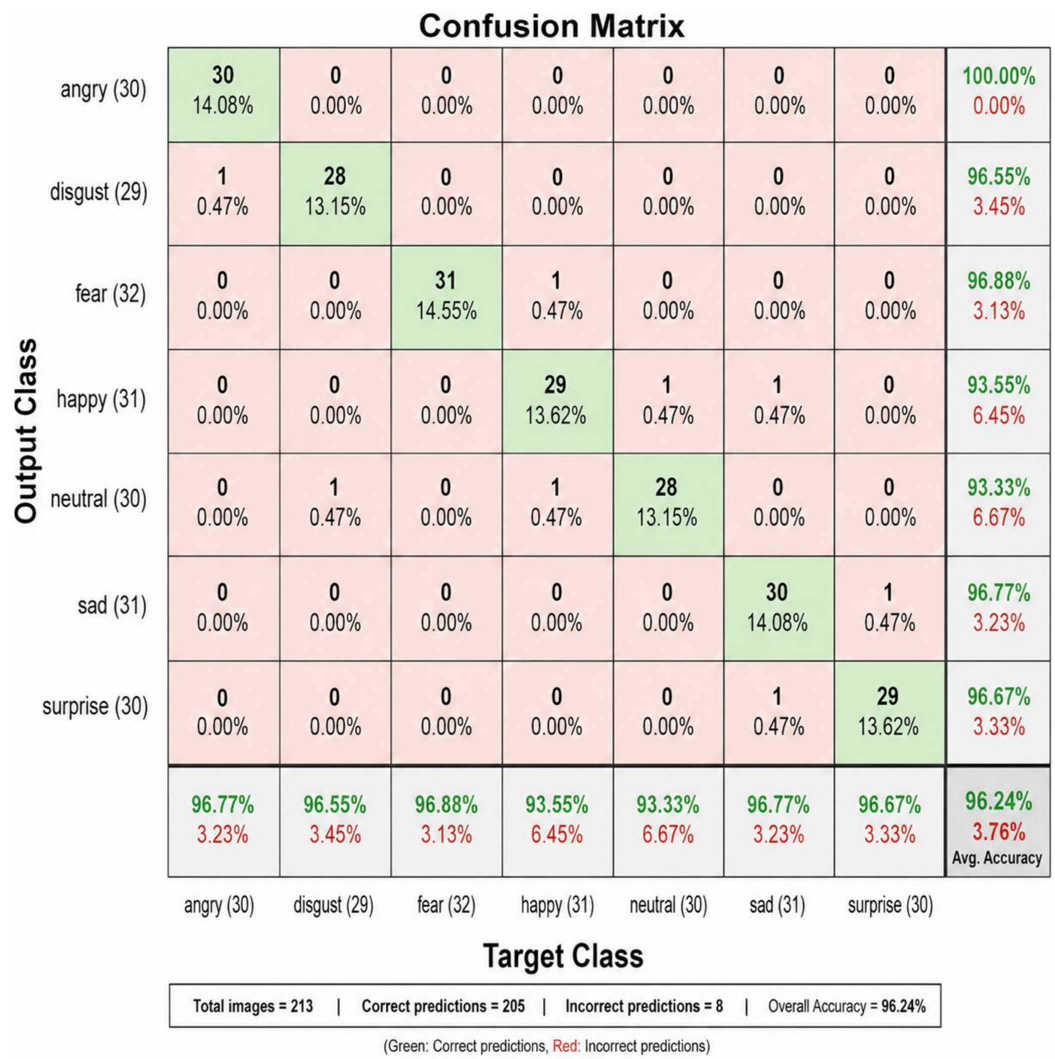

Figure 2: Confusion matrix of the proposed technique on the JAFFE dataset

Table 3: Result of Proposed technique

Emotion	Accuracy	Precision (%)	Recall (%)	F1-Measure (%)	Specificity (%)
Angry	100%	96.77	100.00	98.36	99.45
Disgust	96.55	96.55	96.55	96.55	99.46
Fear	96.88	100.00	96.88	98.41	100.00
Happy	93.55	93.55	93.55	93.55	98.90
Neutral	93.33	96.55	93.33	94.92	99.45
Sad	96.77	93.75	96.77	95.24	98.90
Surprise	96.67	96.67	96.67	96.67	99.45
Average	**96.24**	**96.26**	**96.25**	**96.24**	**99.37**

Figure 2 shows the confusion matrix of the proposed technique. The confusion matrix for the JAFFE dataset displays the suggested facial emotion recognition model's performance in seven classes: angry, disgust, fear, happy, neutral, sad, and surprise. Table 3 shows the result of the proposed technique achieves a maximum average accuracy, precision, recall, F1 – measure and specificity of 96.24%, 96.26%, 96.25%, 96.24% and 99.37%.

Comparative Analysis of the proposed technique with the existing deep learning techniques on the JAFFE dataset

S.No	Author	Deep learning Technique	Accuracy
1	Jain et al. (2018)	Hybrid CNN–RNN	94.91%
2	Islam et al. (2018)	Deep learning/CNN-based FER	93.51%
3	Jain et al. (2019)	Extended Deep Neural Network (DNN)	95.23%
4	Minaee et al. (2021)	Deep-Emotion / Attentional CNN	92.80%
5	Liang et al. (2021)	Patch Attention Layer (PAL) + CNN	93.38%
6	Ahmad et al. (2023)	Stacked Sparse Auto-Encoder (SSAE-FER)	92.50%
7	Podder et al (2023)	CNN + feature boosting + skip connections	96.16%
	Proposed Technique	**CLAHE + Bilateral Filter + CNN**	**96.24%**

The proposed technique has achieved maximum accuracy of 96.24%.The Comparative analysis reveals that the proposed facial emotion recognition technique achieves better results than the competitive models in terms of accuracy.

5. CONCLUSION

This study presents a facial emotion recognition technique based on contrast enhancement using CLAHE, noise reduction using bilateral filter, and emotion classification using a deep convolutional neural network. The suggested strategy is successful in improving the face images quality and keeping significant facial features. The proposed technique achieves 96.24% accuracy on the JAFFE dataset. The findings show that the proposed technique can recognize face emotions accurately and robustly, even in difficult settings such as low contrast and picture noise. Therefore, the proposed technique is a highly effective solution for the identification of facial emotions, as it achieves even greater accuracy than the competitive models.

REFERENCES

Ahmad, M., Saira, Alfandi, O., Khattak, A. M., Qadri, S. F., Saeed, I. A., Khan, S., Hayat, B., & Ahmad, A. (2023). Facial expression recognition using lightweight deep learning modeling. *Mathematical Biosciences and Engineering*, *20*(5), 8208–8225. https://doi.org/10.3934/mbe.2023357

Chen, L., Zhou, M., Su, W., Wu, M., She, J., & Hirota, K. (2018). Softmax regression based deep sparse autoencoder network for facial emotion recognition in human-robot interaction. Information Sciences, 428, 49–61. https://doi.org/10.1016/j.ins.2017.10.044

Deng, J., Pang, G., Zhang, Z., Pang, Z., et al. (2019). CGAN based facial expression recognition for human-robot interaction. IEEE Access. https://doi.org/10.1109/ACCESS.2019.2891668

Du, G., Long, S., & Yuan, H. (2020). Non-contact emotion recognition combining heart rate and facial expression for interactive gaming environments. IEEE Access, 8, 11896–11906. https://doi.org/10.1109/ACCESS.2020.2964794

Gao, L., Zhang, R., Qi, L., Chen, E., & Guan, L. (2019). The labeled multiple canonical correlation analysis for information fusion. IEEE Transactions on Multimedia, 21(2), 375–387.

Ge, H., Zhu, Z., Dai, Y., Wang, B., & Wu, X. (2022). Facial expression recognition based on deep learning. Computer Methods and Programs in Biomedicine, 215, 106621. https://doi.org/10.1016/j.cmpb.2022.106621

Gupta, A., Arunachalam, S., & Balakrishnan, R. (2020). Deep self-attention network for facial emotion recognition. Procedia Computer Science, 171, 1527–1534.

Hua, W., Dai, F., Huang, L., Xiong, J., & Gui, G. (2019). HERO: Human emotions recognition for realizing intelligent Internet of Things. IEEE Access, 7, 24321–24332. https://doi.org/10.1109/ACCESS.2019.2900231

Hung, J. C., & Chang, J.-W. (2021). Multi-level transfer learning for improving the performance of deep neural networks: Theory and practice from the tasks of facial emotion recognition and named entity recognition. Applied Soft Computing, 109, 107491. https://doi.org/10.1016/j.asoc.2021.107491

Islam, B., Mahmud, F., Hossain, A., Goala, P. B., & Mia, M. S. (2018). A facial region segmentation based approach to recognize human emotion using fusion of HOG & LBP features and artificial neural network. *2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT)*, 642–646. https://doi.org/10.1109/CEEICT.2018.8628140

Jain, D. K., Shamsolmoali, P., & Sehdev, 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

Jain, N., Kumar, S., Kumar, A., Shamsolmoali, P., & Zareapoor, M. (2018). Hybrid deep neural networks for face emotion recognition. *Pattern Recognition Letters*, *115*, 101–106. https://doi.org/10.1016/j.patrec.2018.04.010

Jeong, J.-Y., Hong, Y.-G., Kim, D., Jeong, J.-W., Jung, Y., & Kim, S.-H. (2022). Classification of facial expression in-the-wild based on ensemble of multi-head cross attention networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 2353–2358). IEEE/CVF.

Kar, N. B., Babu, K. S., Sangaiah, A. K., & Bakshi, S. (2019). Face expression recognition system based on ripplet transform type II and least square SVM. Multimedia Tools and Applications, 78(4), 4789–4812.

Kim, J.-H., Kim, B.-G., Roy, P. P., & Jeong, D.-M. (2019). Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. IEEE Access, 7, 41273–41285. https://doi.org/10.1109/ACCESS.2019.2907327

Kumari, N., & Bhatia, R. (2020). Comparative study and analysis of various facial emotion recognition techniques. In Decision Analytics Applications in Industry (pp. 157–172). Springer. https://doi.org/10.1007/978-981-15-3643-4_11

Lakshmi, D., & Ponnusamy, R. (2021). Facial emotion recognition using modified HOG and LBP features with deep stacked autoencoders. Microprocessors and Microsystems, 82, 103834. https://doi.org/10.1016/j.micpro.2021.103834

Lee, C.-C., Mower, E., Busso, C., Lee, S., & Narayanan, S. S. (2011). Emotion recognition using a hierarchical binary decision tree approach. Speech Communication, 53(9–10), 1162–1171. https://doi.org/10.1016/j.specom.2011.06.004

Li, B., & Lima, D. (2021). Facial expression recognition via ResNet-50. International Journal of Cognitive Computing in Engineering, 2, 57–64. https://doi.org/10.1016/j.ijcce.2021.02.002

Li, D., Wang, Z., Wang, C., Liu, S., Chi, W., Dong, E., Song, X., Gao, Q., & Song, Y. (2019). The fusion of electroencephalography and facial expression for continuous emotion recognition. IEEE Access, 7, 155724–155736. https://doi.org/10.1109/ACCESS.2019.2949707

Li, S., & Deng, W. (2019). Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition. IEEE Transactions on Image Processing, 28(1), 356–370. https://doi.org/10.1109/TIP.2018.2868382

Li, S., & Deng, W. (2022). Deep facial expression recognition: A survey. IEEE Transactions on Affective Computing, 13(3), 1195–1215. https://doi.org/10.1109/TAFFC.2020.2981446

Liang, X., Xu, L., Liu, J., Liu, Z., Cheng, G., Xu, J., & Liu, L. (2021). Patch attention layer of embedding handcrafted features in CNN for facial expression recognition. *Sensors, 21*(3), 833. https://doi.org/10.3390/s21030833

Lyons, M. J., Akamatsu, S., Kamachi, M., & Gyoba, J. (1998). Coding facial expressions with Gabor wavelets. In Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition (pp. 200–205). IEEE.

Minaee, S., Minaei, M., & Abdolrashidi, A. (2021). Deep-Emotion: Facial expression recognition using attentional convolutional network. *Sensors, 21*(9), 3046. https://doi.org/10.3390/s21093046

Ngai, W. K., Xie, H., Zou, D., & Chou, K.-L. (2022). Emotion recognition based on convolutional neural networks and heterogeneous bio-signal data sources. Information Fusion, 77, 107–117.

Podder T, Bhattacharya D, Majumder P, Balas VE. 2023. A feature boosted deep learning method for automatic facial expression recognition. *PeerJ Computer Science* 9:e1216 https://doi.org/10.7717/peerj-cs.1216

Pu, X., Fan, K., Chen, X., Ji, L., & Zhou, Z. (2015). Facial expression recognition from image sequences using twofold random forest classifier. Neurocomputing, 168, 1173–1180.

Sun, L., Fu, S., & Wang, F. (2019). Decision tree SVM model with Fisher feature selection for speech emotion recognition. EURASIP Journal on Audio, Speech, and Music Processing, 2019(1), Article 2.

Taqvi, S. (2025). Data-driven agricultural development in Assam through machine learning. *Journal of Artificial Intelligence in Agriculture and Applied Research, 1*(1), 34–48. https://agrpjournals.com/journals/journal/index.php/jaiaar/article/view/7/6

Valstar, M., Gratch, J., Schuller, B., Ringeval, F., Lalanne, D., Torres Torres, M., Scherer, S., Stratou, G., Cowie, R., & Pantic, M. (2016). AVEC 2016: Depression, mood, and emotion recognition workshop and challenge. In Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge (pp. 3–10).

Varma, S., Shinde, M., & Chavan, S. S. (2020). Analysis of PCA and LDA features for facial expression recognition using SVM and HMM classifiers. In Techno-Societal 2018 (pp. 109–119). Springer.

Vijaya Lakshmi, A., & Mohanaiah, P. (2021). WOA-TLBO: Whale optimization algorithm with teaching–learning-based optimization for global optimization and facial emotion recognition. Applied Soft Computing, 110, 107623.

Wang, X., Chen, X., & Cao, C. (2020). Human emotion recognition by optimally fusing facial expression and speech feature. Signal Processing: Image Communication, 84, 115831.

Xue, F., Tan, Z., Zhu, Y., Ma, Z., & Guo, G. (2022). Coarse-to-fine cascaded networks with smooth predicting for video facial expression recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 2412–2418). IEEE/CVF.

Zhang, T., Zheng, W., Cui, Z., Zong, Y., & Li, Y. (2019). Spatial-temporal recurrent neural network for emotion recognition. IEEE Transactions on Cybernetics, 49(3), 839–847. https://doi.org/10.1109/TCYB.2017.2788081

Zhang, T., Zheng, W., Cui, Z., Zong, Y., Yan, J., & Yan, K. (2016). A deep neural network-driven feature learning method for multi-view facial expression recognition. IEEE Transactions on Multimedia, 18(12), 2528–2536.

Zhang, Z., Lai, C., Liu, H., & Li, Y.-F. (2020). Infrared facial expression recognition via Gaussian-based label distribution learning in the dark illumination environment for human emotion detection. Neurocomputing, 409, 341–350.