IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CE-43 Scalable Machine Learning Model for Screening and categorizing Individual Depressive Disorder Intensity With Evidential Methods

Bhavana A. Zambare[1], Krishnakant P. Adhiya[2],

[1]Research Scholar,[2] : 0009-0001-6918-8078
SSBT's COET, Bambhori, Jalgaon, India,425001
Email:zambrebhavna@gmail.com
[2]Professor, ::
SSBT's COET, Bambhori, Jalgaon, India,425001
Email: adhiyakp@gmail.com:

Abstract – *Melancholy is a common mental disorder that has a significant impact on people's wellbeing and frequently results in serious emotional, cognitive, and physical deficits. More precise and scalable solutions are required since traditionaldiagnostic techniques, which rely on self-reported questionnaires and clinician interviews, are subject to subjectivity and prejudice. This work suggests a sophisticated machine learning-based method for using facial expression analysis to identify and categorize depression. By automatically evaluating and classifying depression severity, the process seeks to enhance the early detection of depressive disorders by utilizing cutting-edge deep learning algorithms. The study combines a number ofmachine learning models, such as EfficientNet, Vision Transformers, Visual Geometry Group (VGG16), and Residual Network (ResNet50). The goal of the suggested method is to automatically evaluate and categorize depression severity with more precision and dependability.*

Keywords— Depression detection, facial expression analysis, deep learning, emotion recognition.

INTRODUCTION

Over the world, mental disorders are considered the leading public health crisis. In accordance with the World Health Organization, one in every eight individuals faces mental health issues globally. Depression remains a significant mental health issue among the various mental health disorders, depression due to its relatively high prevalence and the depth with which it might strike an individual's life. Moreover, depression [9] is a mood disorder that has become the most disabling condition all over the world, affecting the thoughts and emotions of individuals. In essence, depression can be categorized as a severely degraded mood, a lack of complete interest in the activities that are used to be enjoyed, and various physical and cognitive symptoms that cause functional impairment [15] [4][11][12] [7]. The patients surviving depression face various challenges, including a lack of disease education, difficulty in acquiring medical care, and making the timely treatments hard for patients [16]. Hence, the development of a more effective system for analyzing the severity of depression becomes a crucial process. In recent decades, the depression diagnosis has been highly dependent on the expertise of psychiatrists, which is a time- consuming and labor-intensive process [13] [7]. Additionally, the clinical diagnostic systems for depression highly depend on the standardized psychometric instruments, such as the Hamilton Depression Rating Scale (HAMD-17), the Beck Depression Inventory-II (BDIII), and the Patient Health Questionnaire (PHQ-9). Yet, these tools are inherently limited by evaluator subjectivity and respondent bias, thus complicating objective and quantitative diagnosis [8]. Regarding these issues, more advanced automated detection strategies have been introduced recently. Moreover, the initial depression detection strategies were highly focused on facial behaviours, particularly analyzing the facial expressions and motion. In addition, an interpretable representation of motion dynamics for the estimation of depression severity has been introduced in [14], which showcased the correlation of facial dynamics with emotional state [4]. In the current decade, researchers have implemented various strategies for depression detection through diverse types of data, including Electroencephalogram (EEG) signals, speech, video, and text-based cues [7]. Various machine learning (ML) and deep learning (DL) frameworks have been introduced in recent decades. In essence, the diverse ML methods, including Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNN), and Support Vector Machines (SVM), have demonstrated a greater detection performance with enhanced precision and dependability of systems that identify depression. However, these models still faced limitations of scalability issues, a lack of ability to capture intricate data, a need for additional data, as well as computational complexities [17]. Additionally, the existing DL frameworks have encountered challenges, including overfitting and generalizability issues [1]. Moreover, the ML model, namely the extreme gradient boosting (XGBoost), denoted limitations of lack of transparency and class imbalance issues affected the models' effectiveness [4]. Further, the multimodal large modal-driven ensemble of expert networks (MLM-EOE) approach faced a challenge of a lack of ability to distinguish between the distinct negative emotions [6]. Therefore, to overcome these challenges, the research aims to propose an Incremental Learning Coupled with Evidential Deep Learning model for achieving enhanced depression detection performance.

LITERATURE LITERATURE REVIEW

The World Health Organization (WHO) says that about 300 million people around the world suffer from depression, making it one of the biggest health problems worldwide. Depression is a complex condition with symptoms like low energy, tiredness, and in severe cases, self-harm. Because of this complexity, there is a strong need for more accurate tools to diagnose it. This research uses machine learning to help identify serious mental health conditions like depression and improve the accuracy of diagnoses. The main goal is to improve mental health technology and support better outcomes for patients. A more complete and objective Judgment of a person's mental state can be achieved by Utilizing ML algorithms that draw from multiple inputs, including social media activity, physiological signals (such as EEG), speech patterns, and facial expressions [13].It has recently come to light that facial expression analysis is an effective method for identifying mental health issues, such as depression and emotional states. Emotions conveyed by a person's face are complex and nuanced, and they can be difficult to depict using other methods. Models that can reliably identify and understand these signals have been developed thanks to advancements in deep learning and computer vision, which provide important insights into an individual's psychological well-being [14].

The section explores the literature review of the previous depression detection frameworks, along with their major benefits and drawbacks. In [1], Dharma *et al.* presented a CapsuleThermNet framework for depression detection, which integrates a convolutional neural network (CNN) with the capsule Networks (CapsNet). The model offered a robust performance in detecting depression in their earlier stages. However, the model was prone to an overfitting issue, and its generalizability was impacted due to the dependence on smaller datasets. Choi, D. *et al.* [2] introduced a spatiotemporal spectral feature fusion network (STSFF-Net) model for depression detection. The model demonstrated more notable predictive outcomes than the baseline models by capturing the meaningful affective patterns associated with depression. Nevertheless, the framework failed to track the severity and progression of depression over time, which affected its ability to detect the early warning signs. In [3], Li, X. *et al.* suggested a two-stream network model, namely the SFTNet framework for detecting depression, which is a combination of a single-temporal network (STNet) and a full-temporal network (FTNet). The framework attained higher diagnostic accuracy and robustness through capturing both instantaneous facial expressions and long-term temporal dynamics. Nonetheless, the model struggled with limited generalizability because the sliding window augmentation technique did not apply to general data forms. Hussain, N., *et al* [4] introduced an ML model, namely the extreme gradient boosting (XGBoost) model, for detecting depression, which attained the greatest accuracy in detecting depression accurately. However, the lack of transparency and class imbalance issues affected the models' effectiveness.

Feng, Y., *et al* [5] introduced a cross-scale spatiotemporal enhancement Transformer (CSTEformer) approach for depression detection. The framework showcased an outstanding detection outcome through providing a culturally aware and dynamic solution for enhanced automated depression evaluation. Nevertheless, the model was computationally expensive and demonstrated poor performance in real-world settings. Lin, Z., *et al* [6] developed an MLM-EOE approach for automatic depression detection that effectively determined the characteristics of depression data over multiple states of emotions. Meanwhile, the model struggled with a major drawback of its inability to distinguish between the distinct negative emotions. In [7], Zhou, Y., *et al* implemented a multi-modal fused-attention network for depression detection that showcased a robust performance in identifying the severity levels of depression under sophisticated conditions. However, the models' performance on real-time data remained unverified, which limited their generalizability. An innovative cross-modal coattention network (CMCAN) approach was developed by Guo, W., *et al* [8] for depression detection, which delivered a robust automated depression detection outcome with a greater generalizability. Meanwhile, the models' effectiveness was affected due to issues with data privacy and security, as well as computational complexity.

The major challenges faced by the existing depression detection frameworks are listed below.

- → The CapsuleThermNet approach was prone to overfitting issue and its generalizability was affected due to the reliance on smaller datasets [1].
- → The lack of transparency and class imbalance issues within the XGBoost model impacted its overall effectiveness [4].
- → The MLM-EOE models' inability to distinguish between the distinct negative emotions affected the overall detection accuracy [6].
- → In the CMCAN model, the data privacy and security issues, as well as computational complexity, remained a major
- → challenge [8].

RESEARCH GAP

1. The major challenges faced by the existing depression detection frameworks are listed below.
2. The CapsuleThermNet approach was prone to overfitting issue and its generalizability was affected due to the reliance on smaller datasets [1].
3. The lack of transparency and class imbalance issues within the XGBoost model impacted its overall effectiveness [4].
4. The MLM-EOE models' inability to distinguish between the distinct negative emotions affected the overall detection accuracy [6].
5. In the CMCAN model, the data privacy and security issues, as well as computational complexity, remained a major challenge [8].

PROBLEM DEFINATION

Depression detection remains a crucial process for determining the symptoms and behavioral patterns related to depression, enabling early and timely intervention. In the modern era, various conventional depression detection strategies using ML and DL have been implemented. However, these approaches have struggled with significant drawbacks, including overfitting and generalizability issues [1], lack of transparency and class imbalance issues [4], data privacy and security issues, as well as computational complexity [8]. Hence, to address these issues, the research introduces the Incremental Learning Coupled with Evidential Deep Learning model for performing depression detection more effectively.

METHODOLOGY

The research intends to propose an Incremental Learning Coupled with Evidential Deep Learning Model for performing depression detection effectively.Initially, the input facial images for depression detection will be collected from the AffectNet+ Database [18] and RIKEN facial database [19], which is further preprocessed to standardize the quality of the input images. Consequently, the preprocessed images will be subjected to the feature extraction phase. Here, the features, such as Multiscale Features combining Spatial Feature Mapping (EfficientNet) and Temporal Feature Mapping (Temporal Attention), Distributed-scale Features, and Structural Features (Modified SIFT descriptor), will be extracted for achieving effective processing. Besides, the extraction of these features will enhance the models' ability to capture the hierarchical spatial details, temporal dynamics, as well as the fine-grained structural facial changes, allowing for enhanced performance. Followingly, the extracted feature vector will be transmitted as the input to the proposed Incremental Learning Coupled with Evidential Deep Learning Model for performing final decision-making processes. Within the model, the inclusion of incremental learning mechanism will allow the model to update its understanding of humans' facial expressions without forgetting previously learned knowledge.

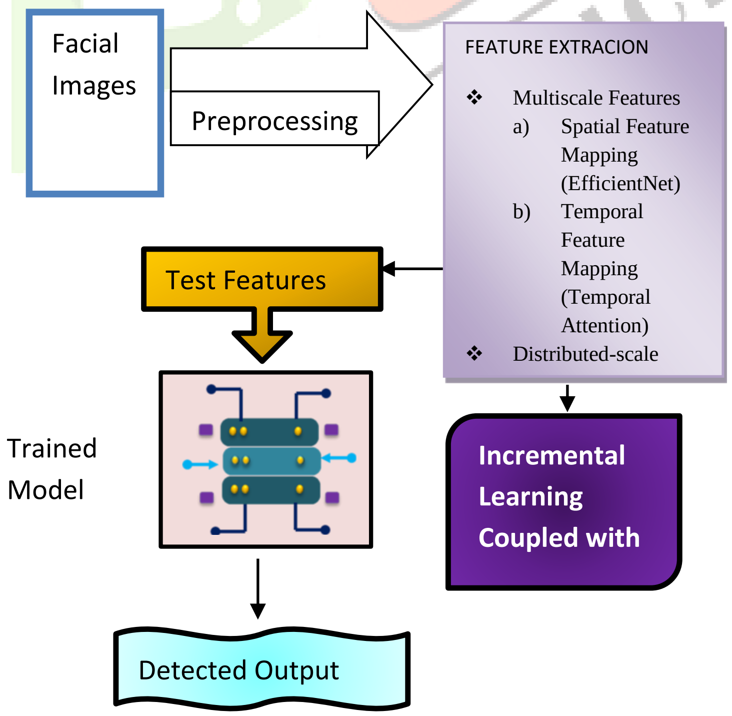

Fig. 1 Block Diagram of the Proposed Incremental Learning Coupled with Evidential Deep Learning Model for Depression Detection

Initially, the input facial images for depression detection will be collected from the AffectNet+ Database [18] and RIKEN facial database [19], which is further preprocessed to standardize the quality of the input images. Consequently, the preprocessed images will be subjected to the feature extraction phase. Here, the features, such as Multiscale Features combining Spatial Feature Mapping (EfficientNet) and Temporal Feature Mapping (Temporal Attention), Distributed-scale Features, and Structural Features (Modified SIFT descriptor), will be extracted for achieving effective processing. Besides, the extraction of these features will enhance the models' ability to capture the hierarchical spatial details, temporal dynamics, as well as the fine-grained structural facial changes, allowing for enhanced performance. Followingly, the extracted feature vector will be transmitted as the input to the proposed Incremental Learning Coupled with Evidential Deep Learning Model for performing final decision-making processes. Within the model, the inclusion of incremental learning mechanism will allow the model to update its understanding of humans' facial expressions without forgetting previously learned knowledge.

Feature Representation

The overall feature vector is expressed as:

$$F = ft\_spatialf(I) \oplus ft\_temporafl(I) \oplus ft\_structuralf(I) \quad (1)$$

where *ft_spatialf*, *ft_temportfl*, and *ft_structuralf* denote spatial, temporal, and structural feature extraction functions, respectively, and $\oplus$ represents feature concatenation.

EfficientNet Scaling

The compound scaling method is defined as:

$$Depth = a^\varphi, \quad Width = b^\varphi, \quad Resolution = c^\varphi \quad (2)$$

subject to:

$$a \cdot c^2 \cdot d^2 \approx 2 \quad (3)$$

where a, b, and c are constants, and φ is the scaling coefficient.

Temporal Attention Mechanism

The attention function is given by:

$$Att(X, Y, Z) = Softmax\left(({}^{T}SM) / \sqrt{d_k}\right) S \quad (4)$$

where X, Y, and represent query, key, and value matrices, respectively, and d_k is the dimension of the key vectors.

Modified SIFT Descriptor

The gradient magnitude and orientation are computed as:

$$SIFT(x,y) = \sqrt{T_x^2 + T_\gamma^2} \quad (5)$$

$$\theta(x,y) = \tan^{-1} (T_\gamma / T_x) \quad (6)$$

where T_x and T_γ denote image gradients along x and y directions.

Incremental Learning

The parameter update rule is expressed as:

$$\theta_{t+1} = \theta_t - \eta \; il(D\_new, \theta_t) \quad (7)$$

where θ_t represents model parameters at time I,l,n is the learning rate, and L is the loss function.

Evidential Deep Learning

The class probability based on evidence is defined as:

$$edl_i = el_i / \Sigma (el_j + 1), \quad j = 1 \text{ to } z \qquad (8)$$

where e_i denotes evidence for class i, and z is the number of classes.

Such a process will assist the model to overcome the challenge of data volatility and continuous evolution in human mental health, and will enhance the adaptability of the model. Additionally, the integration of evidential deep learning will enhance the models' capability to identify diverse types of emotions. Finally, the test features will be given to the trained model to assess the effectiveness of the model in performing depression detection, thus providing a detected outcome. Furthermore, the experiment of the proposed model will be conducted in the PyCharm software using the Python programming language, and the performance of the proposed model will be estimated using accuracy, sensitivity, specificity, and ROC. Figure 1 depicts the block diagram of the proposed model for depression detection.

CONCLUSION

During this portion of the research, I investigated numerous existing depression detection strategies used to identify depression. According to the literature analysis, various machine learning and deep learning algorithms have been developed for depression identification. However, these methodologies suffer a number of issues, including overfitting, a lack of generalization, and difficulty reliably detecting various emotional states. Based on a study of existing methodologies, a proposed Incremental Learning Coupled with Evidential Deep Learning model for depression identification has been developed.

REFERENCES

1. *Dharma, E.M., Prabowo, H., Trisetyarso, A. and Wiguna, T., 2025. CapsuleThermNet: A CNN-CapsuleNet Architecture for Early Detection of Suicide Risk in Depressed Patients Using Thermal Facial Imaging. IEEE Access.*
2. *Choi, D., Hong, J., Lee, J., Liu, Z., Zhang, G. and Jung, J., 2025. STSFF-Net: A two-stream network for depression detection from facial expressions. Neurocomputing, p.131126.*
3. *Li, X., Yi, X., Ye, J., Zheng, Y. and Wang, Q., 2024. SFTNet: A microexpression-based method for depression detection. Computer Methods and Programs in Biomedicine, 243, p.107923.*
4. *Hussain, N., Qasim, A., Mehak, G., Zain, M., Sidorov, G., Gelbukh, A. and Kolesnikova, O., 2025. Multi-level depression severity detection with deep transformers and enhanced machine learning techniques. AI, 6(7), p.157.*
5. *Feng, Y., Guan, Z., Che, C. and Liang, L., 2026. CSTEformer: A cross-scale spatiotemporal perception enhanced transformer for depression detection in Asian populations. Biomedical Signal Processing and Control, 117, p.109604.*
6. *Lin, Z., Wang, Y., Zhou, Y., Du, F. and Yang, Y., 2025. MLM-EOE: Automatic depression detection via sentimental annotation and multi-expert ensemble. IEEE Transactions on Affective Computing.*
7. *Zhou, Y., Yu, X., Huang, Z., Palati, F., Zhao, Z., He, Z., Feng, Y. and Luo, Y., 2025. Multi-modal Fused-attention network for depression level recognition based on enhanced audiovisual cues. IEEE Access.*
8. *Guo, W., Lin, Z., Li, X., Xu, Y., Zhang, S. and Yang, H., 2026. A Transformer-Based Cross-Modal Coattention Framework for Multimodal Depression Detection. International Journal of Intelligent Systems, 2026(1), p.1866250.*
9. *Atta, A., Elsayad, D., Ezzat, D., Amin, S. and ElGamal, M., 2024. Automated depression screening of clinical transcript optimized by grey wolf optimizer. International Journal of Intelligent Computing and Information Sciences, 24(3), pp.58-69.*
10. *Evans-Lacko, S.A.G.S., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Benjet, C., Bruffaerts, R., Chiu, W.T., Florescu, S., de Girolamo, G., Gureje, O. and Haro, J.M., 2018. Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the WHO World Mental Health (WMH) surveys. Psychological medicine, 48(9), pp.1560-1571.*
11. *Islam, M.R., Kabir, M.A., Ahmed, A., Kamal, A.R.M., Wang, H. and Ulhaq, A., 2018. Depression detection from social network data using machine learning techniques. Health information science and systems, 6(1), p.8.*
12. *He, L., Niu, M., Tiwari, P., Marttinen, P., Su, R., Jiang, J., Guo, C., Wang, H., Ding, S., Wang, Z. and Pan, X., 2022. Deep learning for depression recognition with audiovisual cues: A review. Information Fusion, 80, pp.56-86.*
13. *Pan, Y., Shang, Y., Liu, T., Shao, Z., Guo, G., Ding, H. and Hu, Q., 2024. Spatial–temporal attention network for depression recognition from facial videos. Expert systems with applications, 237, p.121410.*
14. *Kacem, A., Hammal, Z., Daoudi, M. and Cohn, J., 2018, May. Detecting depression severity by interpretable representations of motion dynamics. In 2018 13th ieee international conference on automatic face & gesture recognition (fg 2018) (pp. 739-745). IEEE.*
15. *Atta, A., ElSayad, D., Ezzat, D., Amin, S. and ElGamal, M., 2025. Efficacy of swarm-based neural networks in automated depression detection. Scientific reports, 15(1), p.25560. Zhang, X., Li, C., Chen, W., Zheng, J. and Li, F., 2025.*
16. *Optimizing depression detection in clinical doctor-patient interviews using a multi-instance learning framework. Scientific Reports, 15(1), p.6637.*
17. *Kour, H. and Gupta, M.K., 2022. An hybrid deep learning approach for depression prediction from user tweets using feature-rich CNN and bi-directional LSTM. Multimedia tools and applications, 81(17), pp.23649-23685.*
18. *AffectNet+ Database: http://mohammadmahoor.com/pages/databases/affectnetplus/ accessed on February 2026*
19. *RIKEN facial database: https://dmsgrdm.riken.jp:5000/uphvb/wiki/home/ accessed on February 2026.*