Original Article
ISSN (Online): 2582-7472

ShodhKosh: Journal of Visual and Performing Arts
July-December 2023 4(2), 5059–5064

IMAGE AND VIDEO DEPRESSION CLASSIFICATION METHODS BASED ON DEEP LEARNING: A REVIEW

Pratiksha Deshmukh [1], Harshali Patil [1]

[1] Department of Computer Engineering Thakur College of Engineering Technology, Mumbai, Maharashtra, India

Corresponding Author

Pratiksha Deshmukh,
pratiksha.deshmukh@thakureducation.org

DOI

10.29121/shodhkosh.v4.i2.2023.6092

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ABSTRACT

Depression is the main psychological disorder, and it is diagnosed using the psychiatric evaluation and self-assessment questionnaires. However, these approaches are inefficient as they only diagnose depression in its final stages. In order to overcome this issue, deep learning-based methods have gained popularity to diagnose the depression in the early stage by analyzing the text, image, video, and biomedical signals. In this paper, image and video depression classification methods are studied and analyzed based on deep learning. Initially, in this paper, steps are required in the image/video depression classification method based on deep learning explained. Followed by a recent study of the depression classification method based on the dataset, pre-processing method, deep learning method, expected outcome, and inference drawn from it. Finally, we have defined the open research challenges in order to enhance the depression classification methods using the metaheuristic algorithms.

Keywords: CNN, Deep Learning, Depression, Diagnose

1. INTRODUCTION

People's brains are impacted psychologically by the contemporary lifestyle, which leads to depression and emotional distress [1]. One common mental illness that impacts a person's thinking and mental development is depression. The World Health Organization estimates that about 300 million individuals worldwide suffer from depression, and around 1 billion people have mental illnesses [2]. When someone is depressed, they are more likely to consider suicide. Approximately 800,000 individuals kill themselves every year. Therefore, it is necessary to implement a comprehensive approach to address the stress of mental health issues. Depression may negatively impact a person's socioeconomic status. People who are depressed are less likely to interact with others. Depression can be relieved via psychological therapies and counseling [3]. However, there are limitations to diagnosing depression based on an individual's subjective assessment. Depression may be misdiagnosed if its symptoms are misinterpreted as those of other conditions. Symptoms that are not indicative of depression are referred to as "masked depression" [3]. Atypical

depression is hard to identify and simple to misdiagnose, yet it's serious. In these situations, there is a chance that the illness worsens due to delayed treatment. Furthermore, a lot of people with depression are worried about how the public would view their mental illness, therefore they are hesitant to disclose it to professionals. Thus, a reliable and consistent way to differentiate between healthy people and depressed patients is required. The most widely used method for creating a depression detection model now is deep learning.

A recent development in machine learning is deep learning. Its goal is to create and mimic the human brain's neural network so that it can be studied and applied [4]. It mimics how the human brain interprets text, audio, and visual information. This kind of learning is unsupervised. It is also known as a deep neural network as its idea originates from artificial neural network research. Several hidden layers make up a multi-layer perceptron, a kind of deep learning structure. Deep learning forms abstract high-level semantic features by integrating low-level characteristics to represent attribute categories or features to identify dispersed data features. Next, the term "depth" refers to the relationship between machine learning and deep learning.

1) The focus of deep learning is on the model's depth; often, there are five, six, or more hidden layers.
2) Emphasize the significance of feature learning: it is widely recognized that the extraction of image features is the most critical component of the pattern recognition system in the field of recognizing images. The system's ultimate recognition rate is directly impacted by how well features are extracted. Deep learning can get the best feature expression via a process of layer-by-layer feature space translation.
3) Deep learning can learn features from data automatically: The ML features are manually created. Using designers' existing knowledge and parameter modification expertise makes it harder to take use of the benefits of big data.

In this paper, we have studied and analysed the image/video depression classification method based on deep learning. Initially, in this paper, the main steps are utilized for classification purposes are explained. After that, recent research articles are studied and analysed based on various factors, such as dataset, pre-processing method, deep learning algorithm, expected outcome, and inference drawn from it. Finally, the inference drawn from the previous studies and future scope is defined in order to enhance the depression classification method.

The remaining paper has four sections. Section 2 shows an overview of the image/video depression classification method, and recent methods are developed in it using deep learning. Next, Section 3 shows the open research challenges for the image and video depression classification method. Section 4 draws the conclusion.

2. IMAGE/VIDEO DEPRESSION CLASSIFICATION METHODS

Figure 1 shows the flowchart of the image/video depression classification method based on the deep learning algorithm. The detailed description of these steps is given below [5].

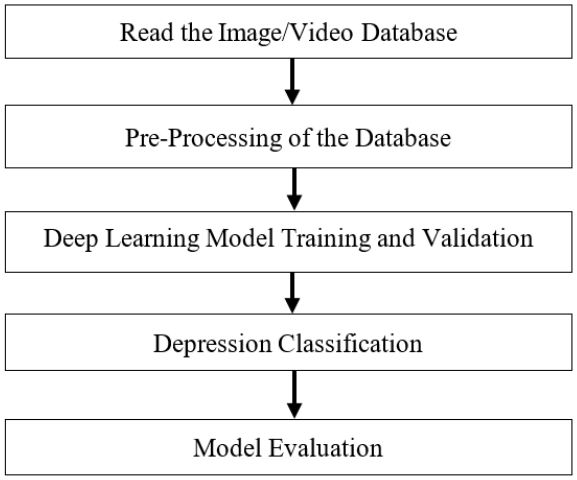

Figure 1 Flowchart of Image/Video Depression Classification Method

- Read the Image/Video Database: In this step, the standard dataset of the image is read for depression classification. On the other hand, in the video database, the frames of the video are extracted using the key frame extraction method for depression classification.
- Pre-Processing of the Database: In this step, the pre-processing on the image/video frame is done to enhance its characteristics using various approaches such as scaling, filtering, enhancement, and standardization.
- Deep Learning Model Training and Validation: In this step, the dataset is split into training and validation ratio. In the literature, 60:40, 70:30, and 80:20 is the most considered ratios. After that, the deep learning model is trained and tested on this dataset.
- Depression Classification: In this step, the deep learning algorithm classifies the depression by analysing the video frame/image characteristics.
- Model Evaluation: In this step, the depression classification method is evaluated using the various parameters in order to validate the performance of it over the existing methods. In the literature, accuracy, precision, recall, and F-score the most preferred parameters. The description of these parameters and how it is determined is given in Table 1 [6].

Table 1 Evaluation Parameters

Parameter	Equation
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F-Score	$\frac{2 * (\text{Recall} * \text{Precision})}{(\text{Recall} + \text{Precision})}$

Note TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative

Next, depression classification methods based on deep learning are evaluated in Table 2. The factors considered to evaluate the methods are the dataset, pre-processing method, deep learning method, expected outcome, and inference drawn.

Table 2 Depression Classification Methods based on Deep Learning

Reference	Dataset	Pre-Processing Method	Deep Learning Method	Expected Outcome	Inference Drawn
Melo et al. [7]	AVEC2013 and AVEC2014	Multi-task cascade convolutional neural network for detection facing landmark. Followed by, cropping and alignment of the face.	ResNet	Resnet+ Expectation Loss: RMSE: 8.25 and 8.23 for the AVED2013 an ADEC2014 dataset whereas MAE for the both dataset is 6.30 and 6.15, respectively.	In their work, the ResNet method is optimized using the ADAM optimizer. In this method, three losses are evaluated softmax, Euclidean lose, and expectation loss. The result shows that the ResNet with expectation loss outperforms over the existing methods.
Rustugi et al. [8]	FER2018 Open Source Database	-	Combinational Deep Neural Network (CNN and RNN)	In their work, visual analysis is performed.	In their work, two neural network is combined and final prediction is done by analysing the score of both models.

Kong et al. [9]	In their research, the depression dataset was collected from Department of Neurology (Affiliated to Hospital of Shandong University of Traditional Chinese Medicine). Further, the dataset was collected for the period of October 2021 to May 2022.	Cropping and scaling	FCN CNN, VGG11, VGG19, ResNet50, and Inception version of CNN	Accuracy value of FCN is 98.23%, VGG11 of 94.40%, VGG19 of 97.35%, and Resnet of 94.99%, and Inception of 97.10%. On the other hand, precision value of FCN is 98.11%, VGG11 of 96.15%, VGG19 of 98.15%, and Resnet of 98.03%, and Inception of 96.20%	In their work, the depression classification is done by utilizing the various approaches. Further, attention mechanism is taken into consideration for analyze the long-distance images.
Cheng et al. [10]	ESRD Patients Dataset (2018-2020)	Image Standardization (Registration, segmentation, and noise reduction), Feature Extraction based on Correlation Coefficient	Capsule Network	The correction rate is 82.47%, recall rate is 83.69%, and accuracy value is 88.79%.	In their work, MRI images are evaluated for depression classification. Further, in their work, Capsnet is used over CNN due to stronger generalization ability along with lesser training is required. Further, the benefit of the Cspsnet is that this network performs the optimization of network parameter in an iteration way which enhances the accuracy of the model.
Rajawat et al. [11]	AVEC2013, AVEC2014, AVEC2017, DAIC, Dementia Bank Database	-	Fusion Fuzzy Logic with CNN	Accuracy: 94.3%	A multi-model is presented by analysing the spatial and temporal information.
Yang et al. [12]	In their research, the CT images are taken considered and collected from a hospital in Mianyang citty.	SSIM for remove duplicate images, median filtering, cropping, and enhancement	MobileNetV2, ResNet50	Accuracy: 97.79%	In their work, fusion of deep learning method is done for depression classification.

Note

MTCNN: Multi-Task Cascade Convolutional Neural Network, CNN: Convolutional Neural Network, FCN: Fully Connected CNN, SSIM: Structural Similarity Index Measure, RNN: Recurrent Neural Network

3. OPEN RESEARCH CHALLENGES

This section presents the open research challenges in order to enhance the current methods proposed for depression classification based on deep learning.

- In the literature, convolutional neural network (CNN) is the most preferred deep learning algorithm over others for depression classification. However, the performance of the CNN algorithm is dependent on the parameters utilized in it, such as training learning rate, kernel length, and activation feature of the convolutional layer. In the previous studies, no authors have done the hyper-tuning of the parameters in order to enhance the accuracy of the method.
- Further, in the literature, traditional algorithms are utilized in the pre-processing method in order to enhance the characteristics of the database, segmentation, and feature extraction. However, we have found that the characteristics of the database image/video frame are different from each other in terms of contrast and noise. Therefore, traditional algorithms need to enhance by hyper-tuning the parameters of it which is not claimed by any authors in the literature.

The above challenges in the pre-processing method and deep learning algorithm can be enhanced by utilizing the metaheuristic algorithm. The metaheuristic algorithm is a category of the optimization algorithm and is employed over optimizations in order to search for the best solution for complex problems [13]. In order to accomplish this goal, initially, the population parameter of the metaheuristic algorithm is initialized. For the initialization purposes, it randomly takes the values in the lower and upper bounds of the parameter and puts these values in the desired algorithm and evaluates based on the objective function. Next, identify the best population at which the optimal value that achieves the desired objective function. Next, the metaheuristic algorithm is executed for a fixed number of iterations in which new populations are generated and evaluated based on the objective function and compared with the best population. If the evaluation of the new population is superior to the best population, then the best population is updated.

4. CONCLUSION

In this paper, image/video depression classification methods are studied and analyzed based on the deep learning algorithms. We have found that in order to enhance the accuracy of the depression classification method the pre-processing the dataset using the filtering, scaling, enhancement algorithms before train the deep learning algorithm. Followed by, several authors are utilized several deep learning algorithms for depression classification and combined their output using various approaches such as fusion and score level. Further, open research challenges are defined in which we have described how depression classification method is enhanced by utilizing the metaheuristic algorithm in the pre-processing and deep learning algorithm to hyper-tune their parameters.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

None.

REFERENCES

S. Aleem, N. U. Huda, R. Amin, S. Khalid, S. S. Alshamrani, and A. Alshehri, "Machine Learning Algorithms for Depression: diagnosis, insights, and research Directions," Electronics, vol. 11, no. 7, p. 1111, Mar. 2022, doi: 10.3390/electronics11071111. Available: https://doi.org/10.3390/electronics11071111

Z. Zhang et al., "Multimodal sensing for Depression risk Detection: integrating audio, video, and text data," Sensors, vol. 24, no. 12, p. 3714, Jun. 2024, doi: 10.3390/s24123714. Available: https://doi.org/10.3390/s24123714

M. Kang, H. Kwon, J.-H. Park, S. Kang, and Y. Lee, "Deep-Asymmetry: Asymmetry Matrix Image for Deep Learning Method in Pre-Screening Depression," Sensors, vol. 20, no. 22, p. 6526, Nov. 2020, doi: https://doi.org/10.3390/s20226526.

P. Wang, E. Fan, and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," Pattern Recognition Letters, vol. 141, pp. 61–67, Aug. 2020, doi: 10.1016/j.patrec.2020.07.042. Available: https://doi.org/10.1016/j.patrec.2020.07.042

P. Meshram and R. K. Rambola, "Diagnosis of depression level using multimodal approaches using deep learning techniques with multiple selective features," Expert Systems, vol. 40, no. 4, Jan. 2022, doi: 10.1111/exsy.12933. Available: https://doi.org/10.1111/exsy.12933

P. Deshmukh and H. Patil, "Depression Prediction Model based on Ensemble Learning Classifier," Indian Journal of Science and Technology, vol. 17, no. 39, pp. 4084-4093, 2024.

W. C. de Melo, E. Granger, and A. Hadid, "Depression Detection Based on Deep Distribution Learning," IEEE Xplore, Sep. 01, 2019. doi: https://ieeexplore.ieee.org/abstract/document/8803467

A. Rustagi, C. Manchanda, N. Sharma, and I. Kaushik, "Depression Anatomy Using Combinational Deep Neural Network," Advances in Intelligent Systems and Computing, vol. 2, pp. 19–33, Jul. 2020, doi: https://doi.org/10.1007/978-981-15-5148-2_3.

X. Kong, Y. Yao, C. Wang, Y. Wang, J. Teng, and X. Qi, "Automatic Identification of Depression Using Facial Images with Deep Convolutional Neural Network," Medical Science Monitor, vol. 28, Jun. 2022, doi: https://doi.org/10.12659/msm.936409.

Yan Cheng, Tengwei Liao, and Nailong Jia, "Classification Algorithms for Brain Magnetic Resonance Imaging Images of Patients with End-Stage Renal Disease and Depression," Contrast Media & Molecular Imaging, pp.1-9, 2022. doi: https://doi.org/10.1155/2022/4795307

Anand Singh Rajawat, Pradeep Bedi, S B Goyal, Pawan Bhaladhare, Alok Aggarwal, Ravi Shankar Singhal, "Fusion Fuzzy Logic and Deep Learning for Depression Detection Using Facial Expressions," Procedia Computer Science, vol. 218, pp. 2795–2805, 2023.

W. Yang, L. Xue, J. Chen, Y. Wang, S. Ding, and H. Zhang, "Classification of recurrent depression using brain CT images through feature fusion," Journal of Radiation Research and Applied Sciences, vol. 17, no. 3, pp. 100986–100986, Jun. 2024, doi: https://doi.org/10.1016/j.jrras.2024.100986

V. Tomar, M. Bansal, and P. Singh, P., "Metaheuristic Algorithms for Optimization: A Brief Review," Engineering Proceedings, vol. 59, no. 1, p.238, 2024.