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

DEPRESSION DETECTION FROM FACIAL BEHAVIOUR THROUGH DEEP LEARNING USING MTCNN ALGORITHM

Dr.B. Gopinathan*, Sneha p, Swapna v**

(*Associate Professor, **Student)

Department of Computer Science Engineering

Adhiyamaan college of engineering (autonomous) Hosur

Abstract- This study examines related literature to propose a model based on machine learning (ML) that can assist in the diagnosis of depressive disorder. Depressive disorder can be diagnosed through a self-report questionnaire, but it is necessary to check the mood and confirm the consistency of subjective and objective descriptions. Millions of people worldwide suffer from depression. There are some differences in condition of mental health between two people who have the same disorder. The degree of depression is analyzed through video-recorded clinical meetings. In worldwide there are 350 million people suffering from depression. Depression patients find hard to concentrate on their software work field. The camera-based assistance in diagnosing depressive disorders can quickly lead to their identification and provide data for intervention provision. Through Multi task cascaded convolution networks (MTCNN), a deep learning method that recognizes vector-based information, a model to assist in the diagnosis of depressive disorder can be devised by checking the position change of the eyes and lips and guessing face emotions based on accumulated photos of the participants who will repeatedly participate in the diagnosis of depressive disorder.

Keywords: **Fast MTCNN, depressive disorder, deep learning, diagnosis, facial expression**.

Domain: **Deep learning**.

I.INTRODUCTION

Emotions play a vital role in our existence. While most of us experience the emotions of happiness, surprise, and fear in reasonable measure, we also tend to experience feelings of sadness, loneliness, or depression at times. These are some of the common reactions to the loss of possessions, life struggles, or wounded self-esteem. Prolonged periods of these feelings can be attributed to physical symptoms that prevent individuals from leading a normal and active life and at the same time indicate depression. This project is mainly aimed at professionals working in IT and university students. A camera installed in the systems continuously monitors the user. Their facial expressions are recorded based on their eye and lip movements. Through multitask convolutional networks (MTCNN), a deep learning method that recognizes vector-based information, a model is achieved that helps in the diagnosis of depressive disorder. It is then compared with a pre-trained emotion detection model. If depression is recorded, it is stored in the database.

II. OBJECTIVE

1. The aim of this project is to analyse and predict depression in its early stages.

2. This project is to help find people in the depression phase and their data will be stored in an excel file.

3. Suicidal thoughts or actions or thinking a lot about death and dying can be easily detected through this project.

4. We can find the behaviour thorough employs work effort on the project. This project will help to find the answer of employees for specific tasks and work space.

III. LITERATURE SURVEY

In this project, a deep learning emotion detection algorithm combining image preprocessing techniques, MTCNN leading to high accuracy of application in depression state detection with complex features and balanced data set was developed. Model trained on Kara′s and Tensor Flow. We can deploy this software application to server end application. [1]. Emotional analysis is essential for affective computing. Emotions can be described as "affecting" and the verb "count" means to calculate or quantify such feelings. To understand humanmachine interactions, we need to create devices or systems that can process and identify, interpret, and replicate human emotions. Text, speech, facial expressions, etc. are examples of this data. [2] People these days produce large amounts of data in the form of unstructured text due to the 2 proliferation of the Internet. Using bad syntax, misspelled words and emerging slang are just some of the problems that can arise when using social networking sites. The presence of these obstacles makes it difficult for bots to perform sentiment and emotion analysis. There are cases when people have difficulty expressing their feelings directly.[3] Classification in deep architectures is made possible by using supervised or unsupervised learning methods in combination with deep architectures. Neural networks can be used to rapidly separate different explanatory components in data to find more abstract qualities at higher levels of representation. Its superior performance in many disciplines, including object perception, speech recognition, computer vision, collaborative filtering and natural language processing, has earned it the title of "Best in Class" (NLP), has recently gained much attention. It is based on the idea that the human brain has multiple representations, including lower-level characteristics and higher-level abstractions. In our minds, concepts are organized hierarchically. Learn basic ideas before building on them to create more abstract ones There are many layers of neurons in the brain that act as feature detectors that become more abstract as the levels increase. Machines can more easily generalize this more abstract style of information representation.[4] With the growth of social media, investors can now communicate with each other more quickly and easily. Because of this, the attitude of investors has an impact on their investment choices, which can quickly spread and amplify throughout the network and affect the stock market. Sentiment and emotion analysis has had a profound impact on the way we do business. [5]

IV.EXISTING SYSTEM

- In the existing method, patient images are saved in JPEG format and these images are used for depression analysis.
- In the existing image processing system, emotion detection based on edge detection. • To do this, they created a model that tested an image to see if it was a picture of a depressed person or a normal person. It ran without neural network algorithms and worked on the stages of the test image in a cascade classifier. The prediction accuracy is poor for the cascade classifier.

• First, they have a trained model. Model training means that the system is able to predict, adapt and judge something. To do this, the system or machine must train.

• They used some deep learning approaches in this work. Here they used CNN (Convolutional Neural Networks) for training data. They used the data set to make the network able to predict the outcome.

A. DISADVANTAGES

• This method used only the images that could be provided and gives an adequate result, but live depression detection is not achieved.

• The accuracy achieved by this system is 74 percent (approximately).

V. PROPOSED SYSTEM

The proposed system focuses on detecting a person's emotions and if the emotion is found to be sad, continuous evaluation is followed to classify sadness and depression. Develop this project to detect employee response in specific situations and workspace environment. It will help increase the level of production and also increase the profitability of the industries.

• Through Multi-Task Cascade Convolutional Networks (MTCNN), a deep learning method that recognizes vector information.

• A model to aid in the diagnosis of depressive disorder can be devised by checking the change of eye and lip position and estimating facial emotions based on the accumulated photographs of participants who will repeatedly participate in the diagnosis of depressive disorder.

• Thus, earlier detection of depression on live web cam is achieved.

A. ADVANTAGE

1. Earlier detection of depression.

2. The results are obtained from the live feed of the camera.

3. Greater accuracy is achieved by 95 percent.

VI.ARCHITECTURE DESIGN

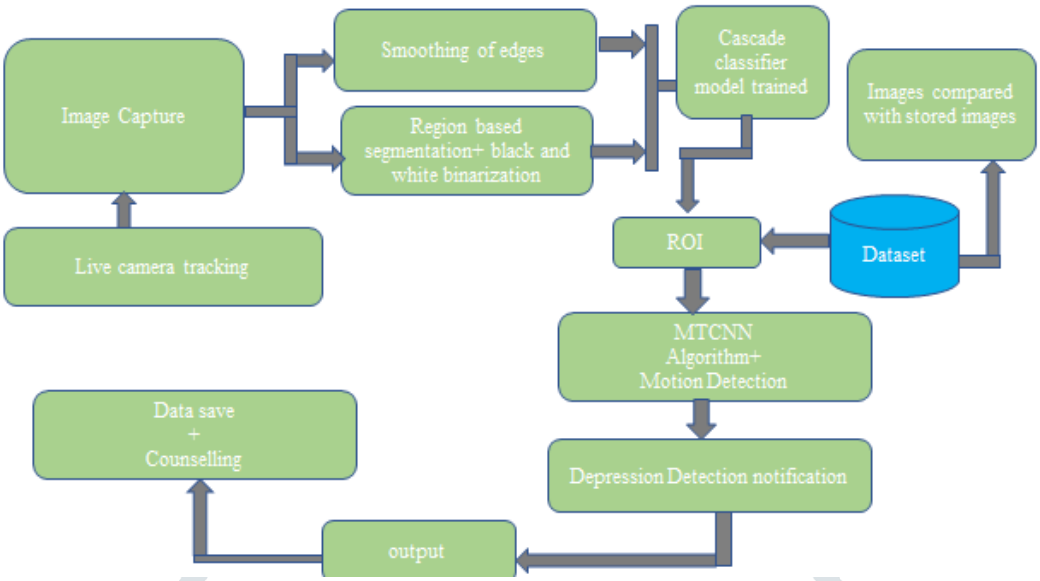

Fig no: **1 Architecture design**

VII.MODULE

There are 5 modules used in this project

- Image acquisition
- Pre-processing
- Cascade Classifier
- Classification
- Feature extraction (ROI)
- MTCNN neural network

MODULE DESCRIPTION

Image acquisition

Frontend for admin login page Admin can see the user's face expression. The expression data will be stored in an Excel sheet and the administrator can view the graph.

Preprocessing

Image pre-processing is the basic processor to satisfy the neural network. We need to resize the image and separate the image from the video. The video is 24 frames per second for individual frames per second, so that the segmentation resembles a binary format for finding edges.

Cascade classifier

Cascade Classifier is used to match the image in the sequence, to train the face model. We need to take 100 positive face related images which we have in a separate folder to implement the training. We need to take 1000 negative images of face related images, there should not be face related images. Create a more accuracy-based epoch (will train the data count algorithm).

Classification

Based on symptom duration, timing, and presumed ethology, depressive disorders are classified into 7 subtypes: major depressive disorder, persistent depressive disorder (or dysthymia), premenstrual dysphoric disorder, substance/medication-induced depressive disorder, depressive disorder due to other medical conditions, other specified depressive disorder and unspecified depressive disorder.

Feature extraction module

A face will be extracted from the ROI. It is used to analyse the religion of the index (reaction, facial expression) until deployment in the neural network. To remove unwanted items.

VIII.SYSTEM FUNCTION

1. Image acquisition

Image acquisition can be defined as the act of retrieving an image from sources. This can be done using hardware systems such as cameras, encoders, sensors, etc. It is undoubtedly the most important step in the MV workflow, as an inaccurate image will render the entire workflow useless. Since machine vision systems do not analyse the acquired digital image of the object and not the object itself, it is paramount to obtain an image with the correct brightness and contrast. In the image acquisition step, the incoming light wave from the object is converted into an electrical signal by a combination of photosensitive sensors. These small subsystems perform the task of providing your machine vision algorithms with an accurate description of the object.

Fig no: **2 Image acquisition**

2. Pre-processing

Data preprocessing is the process of preparing raw data and fitting it for a machine learning model. It is the first and crucial step in building a machine learning model. It's about cleaning, transforming and integrating data to make it ready for analysis. For example, you can resize the input image to match the size of the input image layer. You can also pre-process the data to enhance 4 desired features or reduce artifacts that may affect the network.

Fig no:**3 preprocessing**

3. Classification

Based on symptom duration, timing, and presumed etiology, depressive disorders are classified into 7 subtypes: major depressive disorder, persistent depressive disorder (or dysthymia), premenstrual dysphoric disorder, substance/medication-induced depressive disorder, depressive disorder due to other medical conditions, other specified depressive disorder and unspecified depressive disorder.

Fig no: **4 classifications**

4. Feature extraction (ROI)

Feature extraction refers to the process of transforming raw data into numerical features that can be processed while preserving the information in the original dataset. It delivers better results than applying machine learning directly to raw data. Feature extraction identifies the most discriminating characteristics in signals that can be more easily consumed by a machine learning or deep learning algorithm. Training machine learning or deep learning directly with raw signals often produces poor results due to high data rates and information redundancy.

Fig no:5 Feature extraction (ROI)

5. MTCNN neural network graph

MTCNN or Multi-Task Cascaded Convolutional Neural Networks neural network that detects faces and landmarks in images the process consists of three stages of convolutional networks that are able to recognize faces and the location of landmarks such as eyes, nose and mouth. In the first stage, it uses a shallow CNN to quickly generate candidate windows. The second stage CNN consists of four convolution layers, four pooling layers, and three fully connected layers, which are used to classify the low- and high-risk classes of hypertensive multileads. MTCNN has the best and incredibly high accuracy. Although training takes some time, we can save time by using a pre-trained model and maintain a relatively high accuracy.

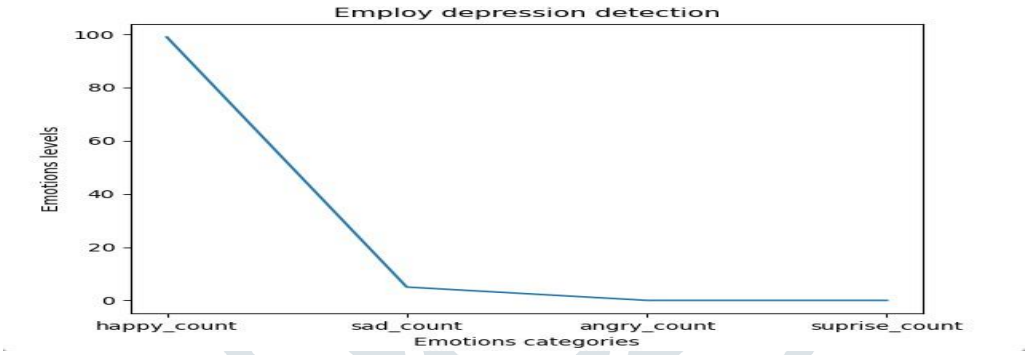

Fig no: 6 MTCNN neural network graph

XI.RESULT

The first model is trained using a cropped face region from video images extracted from the dataset and transfer learning with three different pre-trained 2DMTCNN architectures. The other two models are similarly trained using the cropped eye and mouth regions of the same video image, which are independently fed to the three architectures above. A new algorithm is proposed that optimally combines the three models to ensure the best performance in terms of depression detection accuracy. In addition, it was also observed that the merged models performed satisfactorily in the presence of overfitting problems of the trained models to detect potentially depressed clinical patients.

XII.CONCLUSION

The emotion on the face expresses the intention of the person's work, the employee lags behind with interest after finding out, and after increasing the production process, we can provide guidance to the employee. Based on the fact that we can improve their involvement, this process can be done using the MTCNN algorithm. MTCNN has denser layers before the kernel module is connected by SoftMax and RELUs are converted to a single flat layer to deploy the neural network by MTCNN to improve the accuracy of the test results.

REFERENCE

[1] N.M. Hakak, M. Mohd, M. Kirmani, and M. Mohd, “Emotion Analysis: A Survey”, 2017, 2020.

[2] E. Batbaatar, M. Li, and K. H. Ryu, “SemanticEmotion Neural Network for Emotion Recognition from Text,”2021.

[3] V. Pream Sudha and R. Kowsalya, “a Survey on Deep Learning Techniques, Applications and Challenges,” 2022,

[4] A. Bhardwaj, Y. Narayan, Vanraj, Pawan, and M. Dutta, “Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty,” 2022.

[5] David Sussillo, L. F. Abbott , “Random Walk Initialization for Training Very Deep Feedforward Networks," arXiv: Neural and Evolutionary Computing, 2019.

[6] Hasbi Ash Shiddieqy, FarkhadIhsanHariadi, Trio Adiono ,"Implementation of deep-learning based image classification on single board computer," International Symposium on Electronics and Smart Devices (ISESD),17-19 Oct. 2021.

[7]Ochin Sharma "A New Activation Function for Deep Neural Network," International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon) ,14-16 Feb. 2021.

[8] J. Thevenot, M.B. Lopez, and A. Hadid, “A survey on computer ´vision for assistive medical diagnosis from faces,” IEEE J. Biomed. Health Inform., vol. 22, pp. 1497–1511, 2021.

[9] M. Bishay, P. Palasek, S. Priebe, and I. Patras, “SchiNet: Automatic estimation of symptoms of schizophrenia from facial behaviour analysis,” IEEE Trans. Affective Computing, 2019.

[10] A. Bandini, J.R. Green, B. Taati, S. Orlandi, L. Zinman, and Y. Yunusova, “Automatic detection of amyotrophic lateral sclerosis (ALS) from videobased analysis of facial movements: Speech and nonspeech tasks,” in Proc. IEEE Int. Conf. Automatic Face Gesture Recog., 2020, pp. 150–157.

[11] J.F. Cohn, T.S. Kruez, I. Matthews, Y. Yang, M.H. Nguyen, M.T. Padilla, F. Zhou, and F.D. la Torre, “Detecting depression from facial actions and vocal prosody,” in Proc. Int. Conf. Affective Computing Intel. Interaction W., 2019.

[12] A. Jan, H. Meng, Y.F.B.A. Gaus, and F. Zhang, “Artificial intelligent system for automatic depression level analysis through visual and vocal expressions,” IEEE Trans. on Cognitive and Developmental Systems, vol. 10, pp. 668–680, 2020.

[13] X. Zhou, K. Jin, Y. Shang, and G. Guo, “Visually interpretable repre- sentation learning for depression recognition from facial images,” IEEE Trans. on Affective Computing, pp. 1–12, 2021.

[14] X. Zhou, P. Huang, H. Liu, and S. Niu, “Learning content-adaptive feature pooling for facial depression recognition in videos,” Electronics Letters, vol. 55, no. 11, pp. 648–650, 2019.

[15] W. C. de Melo, E. Granger, and A. Hadid, “Combining global and local convolutional 3D networks for detecting depression from facial expressions,” in Proc. IEEE Int. Conf. Automat. Face Gesture Recog., Lille, France, May 2019.

[16] M. A. Jazaery and G. Guo, “Video-based depression level analysis by encoding deep spatiotemporal features,” IEEE Trans. Affect. Comp., Early Acess, pp. 1–8, Sep. 2019.

[17] W. C. de Melo, E. Granger, and M. B. Lopez, “Encoding temporal information for automatic depression recognition from facial analysis,” in Proc. IEEE Int. Conf. Acoust. Speech and Sig. Process., Barcelona, Spain, May 2020.

[18] X. Zhou, Z. Wei, M. Xu, S. Qu, and G. Guo, “Facial depression recognition by deep joint label distribution and metric learning,” IEEE Tran. Affect. Comput., Early Access, pp.1–14, Sep. 2020. 6

[19] X. Zhou, P. Huang, H. Liu, and S. Niu, “Learning content-adaptive feature pooling for facial depression recognition in videos,” Electronics Letters, vol. 55, no. 11, pp.648–650, Jun. 2019.

[20] W. C. de Melo, E. Granger, and A. Hadid, “A deep multiscale spatiotem-poral network for assessing depression from facial dynamics,” IEEE Tran.Affect. Comput., Early Access, pp.1–12, Aug. 2019.