IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Survey On Automatic Depression Level Detection.

[1]Shilpa S B,
Assistant Professor,
Cambridge Institute of Technology,
Bangalore, India.

[2]Dr Shashikumar D R,
Prof & Head, Dept of CSE,
Cambridge Institute of Technology,
Bangalore, India.

ABSTRACT

The depression is a medical condition and it is among the major mental diseases that affects numerous people worldwide. For this work, the Distress Analysis Interview Corpus-Wizard of Oz interviews (DAJC-WOZ) dataset will be utilized which contains 189 sessions in zip file format and responses from 59 depressed and BO non-depressed patients[1]. This dataset has video, audio and text clips so individual preprocessing techniques are utilized which is given as a source for depression detection level. The inputs are in the form of video, audio and text, so individual preprocessing methods for feature extraction are used with separate techniques. Then, the extracted features from separate techniques are concatenated and given as input for the classification of depression level. The Long-Short Term Memory (LSTM) can be utilized in this research for classifying the depression levels[3].

Key Words: Depression level detection, Feature Fusion, Grey-Level Co-occurrence Matrix, Long-Short Term Memory.

1. INTRODUCTION

In the present years, depression is one of the major fundamental social issues because of its powerful relationship with health and well-being [1]. It is a hazardous infection that not just effects mental condition of the person and leads physical injury to patient. The severity is forecasted based on the patient's mental health conditions [2]. The restlessness, anxiety, sleeping, addictive, stress and eating disorders are the most recurrent cases of mental health disorder [3]. It is a sort of mental disease in that a patient regularly feels demoralization, hopeless, loss of interest in mental, physical and other activities that cause physical and an emotional damage [4] [5]. It affects the learning ability which leads mood variations and decreases the work effectiveness of a patient [6]. According to World Health Organization **(WHO)** 280 million peoples are suffered from depression and 800000 peoples of suicide approximately due to depression in globally [7]. The depression has few stages of incidences like dysthymia, bipolar, clinical and seasonal affective disorder [8].

The prevention and detection of depression is most important and various advanced and traditional methods had proposed in recent years [9]. In the high severity, the brain activities are slow and provides the hormone cortisol that affects neuron development in brain [10]. It is a fourth leading disease in worldwide and affects the all-age peoples like children, adults and old age [11]. Due to the unavailability of early and accurate treatment, more than 80% of depressed peoples not get proper treatment [12]. Recently, the depression measurements are commonly utilized by self-reports, but self-reports have some limitations in practical like the reports of symptoms can be exaggerated or the report depends upon the patients mindset [13] [14]. In the recent researches in wearable technology, the probability of simultaneously monitoring the depression by utilizing the many physiological factors has huge attention [15].

2. LITERATURE REVIEW

Balint Hajduska-Der[4] created a sensitive model that uses speech processing to identify depression. Additionally, models for classification and regression were created using a machine learning technique. Each subject's BDI score and the HAMD score of 20% of the participants are both comprised in the database. The BDI and the HAMD were contrasted in terms of their usefulness by the authors in the direction to train speech signal processing models for automatic depression recognition. By fitting the endpoints of the two scale categories, HAMD scores were translated to the BDI scale for comparison. Within each category, linear scaling was carried out to get the converted HAMD score: H2B.

Kailai Yang[6] presented mental condition Contrastive and Knowledge-aware Network (KC-Net). Initially, KC-Net presents mental state information derived from the generative knowledge base COMET, which specifically models speakers' mental states. The mentalization method then makes use of GRU models and knowledge-aware dot-product attention, which helps the model choose more pertinent knowledge elements. In order to completely utilize label information for capturing class-specific features, a supervised contrastive learning module is utilized. Additionally, it should improve the direction of the mentalization process' knowledge selection process.

Ravi Prasad Thati[7] suggested a unique method for identifying depression that blends task-based processes with real-time MCS. Numerous features from multimodalities have been studied separately and in combination. These features include selection methods based on PCA and Pearson's correlation analysis, as well as various classifiers of machine learning for classification, including Support Vector Machines, Naive Bayes, Decision Trees, and Logistic Regression. SVM produced the best accuracy of 86% when compared to different machine learning classifiers.

Masakazu Higuchi[8] created a composite index that can accurately distinguish major depressive disorder sufferers from healthy persons based on vocal acoustic parameters. OpenSMILE was used to process the speech data in the training set and extract 6552 vocal acoustic features. With ~90% sensitivity, specificity, and accuracy in the training set and ~80% sensitivity, specificity, and accuracy in the test set, the suggested criterion—MDDI—distinguished between depressed and normal participants. Further evidence of MDDI's possible effectiveness in practice comes from the almost complete lack of gender-related variances in the instrument.

Minghao Du[9] proposed a novel machine speech chain model for depression recognition (MSCDR) that can capture text-independent depressive speech representation from the speaker's mouth to the listener's ear to improve recognition performance. In the proposed MSCDR, linear predictive coding (LPC) and Mel-frequency epstral coefficients (MFCC) features are extracted to describe the processes of speech generation and of speech perception, respectively. Then, a one-dimensional convolutional neural network and a long short-term memory network sequentially capture intra- and inter-segment dynamic depressive features for classification.

Vandana *et al.* [I6] introduced a hybrid deep learning for identify depression that integrated text and audio features. This developed model contains three components, initially a text CNN was trained through text features and then the audio CNN was trained through audio features only. Next, an integration of text and audio named as hybrid model in which Long-Short Term Memory (LSTM) was utilized. The improved LSTM like Bidirectional LSTM is utilized for identifying the depression level. This model receives data from both forward and backward directions that enables to capture context features in an input. However, it enhances the computational sources and high training time.

Kirill Milintsevich *et al.* [17] developed an automatic text-based estimation for depression through predicting the symptoms. This paper-trained multi-target hierarchical regression technique for predicting the depression level of individuals from the patient-psychiatrist on DAIC-WOZ dataset. The developed model produced a fine-grained overview of individual symptoms for every person. This model automatically learns meaningful features from text data that was utilized to recognize the patterns and depressions. However, it cannot capture the full spectrum of individual patient health.

Alice Othmani *et al.* [18] implemented a multimodel computer-aided diagnosis scheme for predicting the depression level through audiovisual. The correlation-based anomaly detection and similarity measure of depression was detected when deep audiovisual pattern of a depression was free. The correlation among audiovisual encoding of test subject and deep audiovisual depression illustration was estimated and utilized for monitoring the depressed subjects. However, the developed model requires vast amount of captured data which is complex to obtain.

Clinton Lau *et al.* [I 9] presented a parameter-efficient tuning for an automatic depression severity with deep learning model. This model influences the pretrained large language models and parameter effective tuning techniques. It is created based on adopting a minimum tunable parameter sets which is known as prefix vector. This vector was employed to guide a pretrained model to predict the patient depression Levels. However, this model was complex and overfit the training data which leads the class imbalance issue.

Sandeep Kumar Pandey *et al.* [20] suggested a profound tensor-based technique for classifying depression in speech patterns. The developed model needs significantly simple implementation and extracts traits that aid in the identification of depression. The pooling of mindful outline was achieved to extricate highlights at the sack level that are classified using fully connected layers. This developed model has computationally expensive which required substantial computational resources.

3. PROPOSED WORK

In this section, the novel deep learning based automatic depression level detection is proposed. The proposed work involves dataset of DAIC-WOZ which has videos, audios and texts. The preprocessing and feature extraction are performed individually with separate techniques. The extracted features are fused and then utilized for the classification. Figure I represents the overall performance of proposed work.

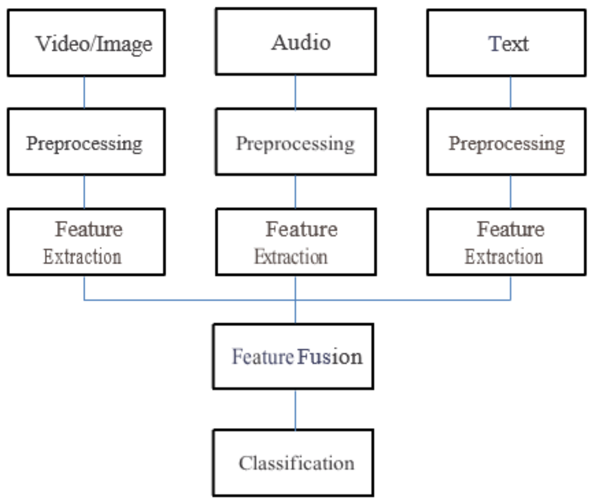

Figure 1. Overall process of proposed work

Preprocessing

The inputs are in the form of video, audio and text, so individual preprocessing techniques will be utilized.

Feature extraction

The preprocessed data are given into feature extractor for extracting the features. For three various pre-processed data, the extraction of features is also done separately.

Feature fusion

The Several characteristics are taken out of individual feature extraction methods that are concatenated in this stage. The concatenated features are given as input to Long-Short Term Memory (LSTM) classifier.

Classification

In this research, for the detection of depression levels, the Long-Short Term Memory(LSTM) neural network can be utilized. The LSTM is developed from RNN for resolving gradient descent problems and capable to learn long time dependence. Before processing the model, the similar pad sequence of data is created so that it can be put to the model with the same length. The LSTM is suitable for this Natural Language Processing (NLP) task due to the existance of three gates such as input, output and forget gates which manages the flow of data over the network.

LSTM performs pre-measurements for the input before producing an output to the final layer in the network. The LSTM have different memory cells in the hidden layer which facilitated by three gates such as input, output and forget gate and these three gates will determine the data, it is need to be store in memory. The cell state which passes information from one layer to another. The forget gate allows only the necessary data to pass through the cell state, the input gate is used for processing the data and the output gate determines the true values of the hidden layers.

4. CONCLUSION

The depression is a medical condition and it is among the major mental diseases that affects numerous people worldwide. For this work, the Distress Analysis Interview Corpus-Wizard of Oz Interviews (DAJC-WOZ) dataset is utilized which contains 189 sessions in zip file format and responses from 59 depressed and BO non-depressed patients. This dataset has video, audio and text clips so individual preprocessing techniques are utilized which is given as input for the detection of depression level. The data sources are as video, sound and text, so individual preprocessing furthermore, include extraction methods are utilized with independent procedures. Then, the extracted features from independent procedures are combined and given as input for the classification of depression level. The Long-Short Term Memory (LSTM) is utilized in this research for classifying the depression levels.

REFERENCES

[1] Ramin Safa, Peyman Bayat and Leila Moghtader,. "Automatic detection of depression symptoms in twitter using multimodal analysis". The Journal of Supercomputing (2022),2021.

[2] Anxo Pérez ∗,Javier Parapar and Álvaro Barreiro. "Automatic depression score estimation with word embedding models". Automatic depression score estimation with word embedding models(www.elsevier.com/locate/artmed), 2022.

[3] Mingyue Niu, Ziping Zhao, Jianhua Tao, Ya Li and Bjorn W. Schuller. " Dual Attention and Element Recalibration Networks for Automatic Depression Level Prediction", IEEE, 2023.

[4] Bálint Hajduska-Dér*, Gábor Kiss, Dávid Sztahó, Klára Vicsi and Lajos Simon. " The applicability of the Beck Depression Inventory and Hamilton Depression Scale in the automatic recognition of depression based on speech signal processing" Frontiers in Psychiatry, 2022.

[5] Usman Naseem, Adam G. Dunn, Jinman Kim and Matloob Khushi. "Early Identification of Depression Severity Levels on Reddit Using Ordinal Classification". Trovato and Tobin, et al., 2022.

[6] Kailai Yang, Tianlin Zhang and Sophia Ananiadou. "A mental state Knowledge–aware and Contrastive Network for early stress and depression detection on social media". Information Processing and Management(www.elsevier.com/locate/ipm),2016.

[7] Ravi Prasad Thati, Abhishek Singh Dhadwal, · Praveen Kumar and Sainaba P. "A novel multi-modal depression detection approach based on mobile crowd sensing and task-based mechanisms". Springer.

[8] Masakazu Higuchi, Mitsuteru Nakamura , Shuji Shinohara, Yasuhiro Omiya , Takeshi Takano, Daisuke Mizuguchi, Noriaki Sonota, Hiroyuki Toda, Taku Saito, Mirai So, Eiji Takayama, Hiroo Terashi, Shunji Mitsuyoshi and Shinichi Tokuno. "Detection of Major Depressive Disorder Based on a Combination of Voice Features: An Exploratory Approach". International Journal of Environmental Research and Public Health, 2022.

[9] Minghao Du, Shuang Liu, Tao Wang, Wenquan Zhang, Yufeng Ke, Long Chen and Dong Ming. "Depression recognition using a proposed speech chain model fusing speech production and perception features ". ELSEVIER, 2022.

[10] Hamad Zogan, Imran Razzak, Xianzhi Wang, Shoaib Jameel and Guandong Xu. "Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media". Springer 2022.

[11] Mingyue Niu, Ziping Zhao, Jianhua Tao and Ya Li. "Selective Element and Two Orders Vectorization Networks for Automatic Depression Severity Diagnosis via Facial Changes" IEEE, 2022.

[12] Bochao Zou, Jiali Han, Yingxue Wang, Rui Liu, Shenghui Zhao, Lei Feng, Xiangwen Lyu and Huimin Ma. "Semi-structural interview-based Chinese multimodal depression corpus towards automatic preliminary screening of depressive disorders". 2022 IEEE.

[13] *Emna Rejaibi, Ali Komaty, Fabrice Meriaudeau, Said Agrebi and Alice Othmani*." MFCC-based Recurrent Neural Network for Automatic Clinical Depression Recognition and Assessment from Speech". Elsevier, 2021.

[14] David Lin, Tahmida Nazreen, Tomasz Rutowski, Yang Lu, Amir Harati, Elizabeth Shriberg, Piotr Chlebek and Michael Aratow. "Feasibility of a Machine Learning-Based Smartphone Application in Detecting Depression and Anxiety in a Generally Senior Population". Frontiers in Psychology, 2022.

[15] Xiaohan Zang, Baimin Li, Lulu Zhao, Dandan Yan and Licai Yang. "End-to-End Depression Recognition Based on a One-Dimensional Convolution Neural Network Model Using Two-Lead ECG Signal". Journal of Medical and Biological Engineering.

[16] Vandana a, Nikhil Marriwala and Deepti Chaudhary. "A hybrid model for depression detection using deep learning", Elsevier 2022.

[17] Kirill Milintsevich, Kairit Sirts and Gaël Dias. "Towards automatic text-based estimation of depression through symptom prediction". Springer Open, 2023.

[18] Alice Othmani and Assaad Oussama Zeghina. "A multimodal computer-aided diagnostic system for depression relapse prediction using audiovisual cues: A proof of concept". Elsevier, 2022.

[19] Clinton Lau, Xiaodan Zhu and Wai-Yip Chan. "*Automatic depression severity assessment with deep learning using parameter-e cient tuning"*. Frontiers in Psychiatry, 2023.

[20] Sandeep Kumar Pandey, Hanumant Singh Shekhawat, S. R. M. Prasanna, Shalendar Bhasin and Ravi Jasuja. "A deep tensor-based approach for automatic depression recognition from speech utterances".

Table 1. Experimental result of existing methods

Author	Methodology	Evaluation Metrics
Vandana *et al.* [16] (2023)	Hybrid Bi-LSTM	Accuracy-0.88 Precision-0.75 Recall-0.73 Fl-score-0.74
Kirill Milintsevich *et al.* [17] (2023)	Symptom prediction	Macro-F1-0.739 Micro-H-0.766 MAE-3.78
Alice Othmani *et al.* [18] (2022)	Similarity Measure	Accuracy-82.55% F1-score-79.9%
Clinton Lau *et al.* [19] (2023)	Dual encoder model	MAE-3.80 RMSE-4.87
Sandeep Kumar Pandey *et al.* [20] (2022)	3DTF	Accuracy-0.7245 Fl-score-0.705