# A Review on Depression Detection[[1]](#footnote-1)

## Abstract

**Purpose of Review**: Given advancements in hardware and software, along with the explosion of smartphone use, the forms of potential health care solutions have begun to change. Among them, automatic detection of depression has attracted increasing attention from researchers in psychology, computer science, linguistics, and related disciplines. Promising depression detection systems based on different data sources have been reported recent years. In this paper, we survey these efforts by presenting the review of representative researches and discusses the most promising approaches.

**Recent Findings**: We reviewed over fifty studies of detection of depression that used electronic health records (EHRs), sensor data (e.g. electroencephalogram(EEG), kinematic skeletal data), text information ( extracted from clinical interview or collected from social media platform), image data (brain image or facial expression) and audio to predict or classify depression. Depression detection systems across modalities are also included. Collectively, these studies provided excellent examples of using machine learning algorithms to address mental health questions.

**Summary**: Depression detection is inherently a multimodal problem. We hope such efforts may give an inspiration for further research.

## Introduction

According to WHO[1], it is estimated that over 300 million people suffer from depression, which corresponds to the 4.4% of the world’s population. Major depression significantly affects a person's family, personal relationships and other general health aspects. WHO also states that over 800,000 suicide deaths are reported each year due to depression, while for 15-29-year-old people, it is the leading factor of death. Depression has become a main health burden in the world. Therefore, the research on reliable depression detection system is of great value.

Through computer technology, some existing research works aim to build classification models to perform early detection of depression. Some of them use wearable devices to collect physiological signals of the patients such as Electroencephalogram (EEG)[2][3][4][5][6][7], eye tracking data[8], kinematic skeletal data[9] etc. for depression detection. We put these data under the same category because they are objective bio-signals. Brain imaging is another objective data source we contained. Some others use data at the form of text, audio, or image which may bring subjective biases. Study based on text extracted from clinical interview is mentioned separately. Other studies which utilized text, audio, image are surveyed from The Audio/Visual Emotion Challenge and Workshop (AVEC)[10], which is a competition event aimed at comparison of multimedia processing and machine learning methods for automatic audio, visual and physiological depression and emotion analysis, with all participants competing under strictly the same conditions. The competition promoted research, and built connections across the research community. Details are talked as below.

## Research Overview

### Electronic Health Records

Normally electronic health records(EHR) maintains personal information of each patient(such as birth date, gender, weight or something) and their medical data. The key point is how to utilize all these features.

The depression analysis using EHR can be classified into two categories. In the first category, the researchers focus on specific diseases or symptoms, analyze their association with depression. For example, Shen et al.[15] study the risk of depression among patients with uterine leiomyoma. In the second category, no specific diseases or symptoms are specified. Tai et al.[13] use the ratio information to find and validate the results of feature selection methods, and the prediction results of 30 days in advance reached F1 = 0.75 and ACC =0.75 with the classifier of SMO, SimpleLogistic, and LMT. Arun et al.[14] applied XGBoost to MYNAH cohort for depression detection and achieved accuracy of 97.70%.

### Sensor Data

#### EEG Based Depression Recognition

EEG is an electrophysiological technology reflecting the functional state of the brain. It is produced by the internal nervous system of brain function and is often used in auxiliary diagnosis of mental illnesses such as depression. Lots of existing researches showed that EEG could be used as a reliable source in depression detection. Erguzel et al. used the EEG data collected from 147 major depressed patients for classification with back-propagation neural network (BPNN) and got an accuracy of 89.12%[6]. Hosseinifard used a combination of genetic algorithm (GA) and support vector machine (SVM) classifier for classification achieving an overall accuracy of 88.6%[7]. Betul et al.[4] employed convolutional neural network (CNN) and long-short term memory (LSTM) architectures to build a deep hybrid model. They provided 99.12% and 97.66% classification accuracies for the right and left hemisphere EEG signals respectively. Li et al.[2] used convolutional neural network to detect depression using EEG signals. Spyrou et al.[5] made a comparison of classification methodologies based on selected features and demonstrated the high accuracy (95.5%) of Random Forest. Zhu et al.[8] proposed a content-based ensemble method (CBEM) to promote the depression detection accuracy and the model achieved accuracies of 82.5% and 92.65% respectively on the selected two datasets.

The EEG data used by researchers mentioned above were obtained in individual modality. Cai et al.[3] constructed a novel multimodal model by fusing different electroencephalogram (EEG) data sources and achieved a high classification accuracy of 86.98% with KNN classifier.

#### Body Gestures Based Depression Recognition

Human action has been found to reflect patients’ mental status, including the state of depressive disorders[11]. Taking walking as a striking example, some researchers found that depressed patients showed reduced walking speed, abnormal hand movements, larger lateral swaying movements of the upper body and larger reaction time variability[12]. We only picked one latest related work as a representative. Li et al.[9] using Kinect V2 device to record participants kinematic skeleton data of the participant’s 25 body joints, and with original Kinect skeleton data extracting and preprocessing, proposed experiment demonstrated four strong machine learning tools: Support Vector Machine, Logistic Regression, Random Forest and Gradient Boosting. Utilizing the best model Gradient Boosting for age-based classification, prediction accuracy got 76.92% in the older group(age > 40) and 53.85% accuracy in the younger group (age <= 40).

### Text

Psychological studies[16] have observed differences in the use of language between depressed and non-depressed individuals. For example, depressed individuals tend to use more absolutist words[19] , negatively valenced-words and the pronoun “I”[20] and mention pharmaceutical treatment for depressive disorder[21][22]. Thus the work which aims to automatically identify depressed and non-depressed participants through the analysis of linguistic information extracted from transcribed clinical interviews are included. It can be seen as a natural language processing(NLP) problem. Here we only pick some representative works published recent years. Xezonaki et al.[18] used a Hierarchical Attention Network that encodes words and dialogue turns in different levels of the architecture in 2020. The best performance achieved by the HAN+L model on DAIC-WoZ dataset[23] can be 0.69 (F1-macro). And in 2019 Mallol-Ragolta et al.[17] had already employed hierarchical attention-based networks. Trotzek et al.[25] trained an own fastText model on a dataset containing all reddit comments and used convolutional neural network(CNN) for classification. Their work obtained the best performance on the eRisk 2017[27] task for comparison to previously published results. Resnik et al.[28] proved that the LDA model can uncover a meaningful and potentially useful latent structure for the automatic identification of important topics for depression detection. Tadesse et al.[26] characterized a closer connection between depression and a language usage by applying NLP and text classification techniques. In their study, the strength and effectiveness of combined features (LIWC+LDA+bigram) are demonstrated with the MLP classifier reaching 91% accuracy and 0.93 F1-score.

For more related work, [57] can be a good reference.

### Image

In this part, we discuss two kinds of image based depression detection.

The first type is brain image. Objective brain imaging measurement is an effective diagnostic tool for depression detection. We surveyed several machine-learning-based classification and prediction studies of depression which utilize features derived from magnetic resonance imaging (MRI) data. We find that support vector machines (SVM) is the most prevalent method choice, for example, Rubin-Falcone et al.[29] applied SVM to distinguish Major Depressive Disorder (MDD) from Bipolar Disorder(BD) and achieved the accuracy of 75%. Other ML methods have also been applied to depression detection such as gaussian process classifier (GPC)[30], linear discriminant analysis (LDA)[31], decision tree (DT)[32], and partial least squares regression(PLS)[34], as well as more recent deep learning models. For example, Pominova et al.[33] tested various modifications of VoxCNN and VoxResNet models to solve epilepsy recognition task using structural MRI scans and voxelwise recurrent-convolutional networks for classification of patients with and without depression based on their functional MRI series. They achieved 0:73 ROC-AUC score when classifying depression versus a control group. For more related work, [35] is a good reference.

The second type is the facial expression of patients. Usually researchers use facial expression dynamics(video) for prediction, combined with vocal features(audio)[36]. Thus we put related work into the ‘Multimodal’ category.

### Audio

The emotional state of a person suffering from a depressive disorder affects the acoustic qualities of his/her speech[37]. Therefore, depression could be detected through an analysis of perceived changes in the acoustical properties of speech. The link between acoustic parameters in speech signals and depression has been researched extensively.

A wide range of features have been explored for automatic depressed speech classification. Moore et al.[39], Low et al.[40], and Ooi et al.[41][42] investigated the suitability of forming a classification system from combinations of prosodic, spectral, and glottal features. Alghowinem et al.[43] , Valstar et al.[44] , and Low et al.[45] summarized and compared Low-Level descriptors and statistical features of depression classification. Investigation of mel-frequency cepstrum coefficients (MFCC) by Cummins et al.[46][47] and Joshi et al.[48] found that the classification results were statistically significant for detecting depression. Ozdas et al.[49] found that depressed patients exhibited higher energy in the upper frequency bands of the glottal spectrum. The popular modeling and classification techniques used in the literature include SVM, GMM, KNN[38]. Ma et al.[50] proposed a deep model DepAudioNet which consists of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). They carried out evaluations on DAIC-WOZ which was used for the AVEC 2016 competition and demonstrate the effectiveness of the proposed method. Ma’s work inspired us to pay attention to AVEC for more creative and effective works, which are mentioned in next part.

### Multimodal

First we need to introduce the dataset provided by AVEC. The Distress Analysis Interview Corpus - Wizard of Oz (DAIC-WOZ) database is part of a larger corpus, the Distress Analysis Interview Corpus (DAIC) [23], that contains clinical interviews designed to support the diagnosis of psychological distress conditions such as anxiety, depression, and post-traumatic stress disorder. The database includes data from 189 subjects. For each subject, the database includes the audio/video features as well as the transcript of an interview ranging between 7-33 minutes, which is conducted by an animated virtual interviewer called Ellie, controlled by a human interviewer in another room.

Several classical baselines are then introduced. In Williamson et al.[53].’s work, biomarkers are derived from all of provided modalities. Semantic features are analyzed for subject/avatar dialogue content using a sparse coded lexical embedding space, and for contextual clues related to the subject’s present or past depression status. The results on the dataset arrives F1=0.81. Gong et al.[56] proposed a novel topic modeling based approach to perform context-aware analysis which has a performance of F1=0.70. Haque et al[52] proposed sentence-based model for measuring depressive symptom severity from de-identified multi-modal data. The method leverages a causal convolutional network (C-CNN) to “summarize" sentences into a single embedding and got the performance of F1=0.769. Alhanai et al.[54] modeled the interactions with audio and text features in a Long-Short Term Memory (LSTM) neural network model to detect depression and got the performance of F1=0.77. Lam et al.[55] proposed a method that incorporates a data augmentation procedure based on topic modelling using transformer and deep 1D convolutional neural network (CNN) for acoustic feature modeling and got the performance of F1=0.87. Zheng et al.[51] proposed a graph attention model embedded with multi-modal knowledge for depression detection in 2020 and the model improves the classification and prediction performance compared with the major state-of-the-art approaches mentioned above.

## Discussion

This review only overviews the tools and a few representative methods which help to detect depression. More work such as comparing advantages and disadvantages of each model can be done for improvement. We hope our work can give a reference for further research.

## References

1. W. H. Organization, Depression and Other Common Mental Disorders: Global Health Estimates., 2017
2. Li X, La R, Wang Y, et al. A deep learning approach for mild depression recognition based on functional connectivity using electroencephalography[J]. Frontiers in neuroscience, 2020, 14: 192.
3. Cai H, Qu Z, Li Z, et al. Feature-level fusion approaches based on multimodal EEG data for depression recognition[J]. Information Fusion, 2020, 59: 127-138.
4. Ay B, Yildirim O, Talo M, et al. Automated depression detection using deep representation and sequence learning with EEG signals[J]. Journal of medical systems, 2019, 43(7): 1-12.
5. Spyrou I M, Frantzidis C, Bratsas C, et al. Geriatric depression symptoms coexisting with cognitive decline: a comparison of classification methodologies[J]. Biomedical Signal Processing and Control, 2016, 25: 118-129.
6. Erguzel T T, Ozekes S, Tan O, et al. Feature selection and classification of electroencephalographic signals: an artificial neural network and genetic algorithm based approach[J]. Clinical EEG and neuroscience, 2015, 46(4): 321-326.
7. Hosseinifard B, Moradi M H, Rostami R. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal[J]. Computer methods and programs in biomedicine, 2013, 109(3): 339-345.
8. Zhu J, Wang Z, Gong T, et al. An improved classification model for depression detection using EEG and eye tracking data[J]. IEEE transactions on nanobioscience, 2020, 19(3): 527-537.
9. Li W, Wang Q, Liu X, et al. Simple action for depression detection: using kinect-recorded human kinematic skeletal data[J]. BMC psychiatry, 2021, 21(1): 1-11.
10. Valstar M, Gratch J, Schuller B, et al. Avec 2016: Depression, mood, and emotion recognition workshop and challenge[C]//Proceedings of the 6th international workshop on audio/visual emotion challenge. 2016: 3-10.
11. Kleinsmith A, Bianchi-Berthouze N. Affective body expression perception and recognition: A survey. IEEE Trans Affect Comput. 2012;4(1):15–33.
12. Kim J-Y, Liu N, Tan H-X, Chu C-H. Unobtrusive monitoring to detect depression for elderly with chronic illnesses. IEEE Sensors J. 2017;17(17):5694–704
13. Tai L K, Setyonugroho W, Chen A L P. Finding discriminatory features from electronic health records for depression prediction[J]. Journal of Intelligent Information Systems, 2020, 55(2): 371-396.
14. Arun V, Prajwal V, Krishna M, et al. A boosted machine learning approach for detection of depression[C]//2018 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE, 2018: 41-47.
15. Shen, T.-C., Yang, C.-Y., Huang, Y.-J., Lin, C.-L., Sung, F.-C. (2017). Risk of depression in patients with uterine leiomyoma: a nationwide population-based cohort study. Journal of Affective Disorders, 213,126–130.
16. J. D. Bernard, J. L. Baddeley, B. F. Rodriguez, and P. A. Burke,“Depression, Language, and Affect: An Examination of the Influence of Baseline Depression and Affect Induction on Language,”Journal of Language and Social Psychology, vol. 35, no. 3, pp.317–326, 2016
17. Mallol-Ragolta A, Zhao Z, Stappen L, et al. A hierarchical attention network-based approach for depression detection from transcribed clinical interviews[J]. 2019.
18. Xezonaki D, Paraskevopoulos G, Potamianos A, et al. Affective Conditioning on Hierarchical Networks applied to Depression Detection from Transcribed Clinical Interviews[J]. arXiv preprint arXiv:2006.08336, 2020.
19. Al-Mosaiwi M, Johnstone T. In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation[J]. Clinical Psychological Science, 2018, 6(4): 529-542.
20. Rude S, Gortner E M, Pennebaker J. Language use of depressed and depression-vulnerable college students[J]. Cognition & Emotion, 2004, 18(8): 1121-1133.
21. De Choudhury M, Gamon M, Counts S, et al. Predicting depression via social media[C]//Seventh international AAAI conference on weblogs and social media. 2013.
22. Ramirez-Esparza N, Chung C K, Kacewicz E, et al. The Psychology of Word Use in Depression Forums in English and in Spanish: Texting Two Text Analytic Approaches[C]//ICWSM. 2008.
23. Gratch J, Artstein R, Lucas G M, et al. The distress analysis interview corpus of human and computer interviews[C]//LREC. 2014: 3123-3128.
24. Orabi A H, Buddhitha P, Orabi M H, et al. Deep learning for depression detection of twitter users[C]//Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic. 2018: 88-97.
25. Trotzek M, Koitka S, Friedrich C M. Utilizing neural networks and linguistic metadata for early detection of depression indications in text sequences[J]. IEEE Transactions on Knowledge and Data Engineering, 2018, 32(3): 588-601.
26. M. M. Tadesse, H. Lin, B. Xu and L. Yang, "Detection of Depression-Related Posts in Reddit Social Media Forum," in IEEE Access, vol. 7, pp. 44883-44893, 2019, doi: 10.1109/ACCESS.2019.2909180.
27. Losada D, Crestani F, Parapar J. CLEF 2017 eRisk Overview:“Early Risk Predictionon the Internet: Experimental Foundations[C]//Working Notes of CLEF 2017-Conference and Labs of the Evaluation Forum. 2017.
28. P. Resnik, W. Armstrong, L. Claudino, T. Nguyen, V.-A. Nguyen and J. Boyd-Graber, "Beyond LDA: Exploring supervised topic modeling for depression-related language in twitter", Proc. 2nd Workshop Comput. Linguistics Clin. Psychol. Linguistic Signal Clin. Reality, pp. 99-107, 2015.
29. Rubin-Falcone H, Zanderigo F, Thapa-Chhetry B, et al. Pattern recognition of magnetic resonance imaging-based gray matter volume measurements classifies bipolar disorder and major depressive disorder[J]. Journal of affective disorders, 2018, 227: 498-505.
30. Frangou S, Dima D, Jogia J. Towards person-centered neuroimaging markers for resilience and vulnerability in Bipolar Disorder[J]. Neuroimage, 2017, 145: 230-237.
31. Sato J R, Moll J, Green S, et al. Machine learning algorithm accurately detects fMRI signature of vulnerability to major depression[J]. Psychiatry Research: Neuroimaging, 2015, 233(2): 289-291.
32. Patel M J, Andreescu C, Price J C, et al. Machine learning approaches for integrating clinical and imaging features in late‐life depression classification and response prediction[J]. International journal of geriatric psychiatry, 2015, 30(10): 1056-1067.
33. Pominova M, Artemov A, Sharaev M, et al. Voxelwise 3d convolutional and recurrent neural networks for epilepsy and depression diagnostics from structural and functional mri data[C]//2018 IEEE International Conference on Data Mining Workshops (ICDMW). IEEE, 2018: 299-307.
34. Yoshida K, Shimizu Y, Yoshimoto J, et al. Prediction of clinical depression scores and detection of changes in whole-brain using resting-state functional MRI data with partial least squares regression[J]. PloS one, 2017, 12(7): e0179638.
35. Gao S, Calhoun V D, Sui J. Machine learning in major depression: From classification to treatment outcome prediction[J]. CNS neuroscience & therapeutics, 2018, 24(11): 1037-1052.
36. Jan A, Meng H, Gaus Y F A, et al. Automatic depression scale prediction using facial expression dynamics and regression[C]//Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge. 2014: 73-80.
37. Cummins N, Scherer S, Krajewski J, et al. A review of depression and suicide risk assessment using speech analysis[J]. Speech Communication, 2015, 71: 10-49.
38. Jiang H, Hu B, Liu Z, et al. Investigation of different speech types and emotions for detecting depression using different classifiers[J]. Speech Communication, 2017, 90: 39-46.
39. Moore II E, Clements M A, Peifer J W, et al. Critical analysis of the impact of glottal features in the classification of clinical depression in speech[J]. IEEE transactions on biomedical engineering, 2007, 55(1): 96-107.
40. Low L S A, Maddage N C, Lech M, et al. Detection of clinical depression in adolescents’ speech during family interactions[J]. IEEE Transactions on Biomedical Engineering, 2010, 58(3): 574-586.
41. Ooi K E B, Lech M, Allen N B. Multichannel weighted speech classification system for prediction of major depression in adolescents[J]. IEEE Transactions on Biomedical Engineering, 2012, 60(2): 497-506.
42. Ooi K E B, Lech M, Allen N B. Prediction of major depression in adolescents using an optimized multi-channel weighted speech classification system[J]. Biomedical Signal Processing and Control, 2014, 14: 228-239.
43. Alghowinem S, Goecke R, Wagner M, et al. Detecting depression: a comparison between spontaneous and read speech[C]//2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013: 7547-7551.
44. Valstar M, Schuller B, Smith K. 3D dimensional affect and depression recognition challenge[C]//Proceedings of 4th ACM International Workshop on Audio/Visual Emotion Challenge (AVEC’14). 3-10.
45. Low L S A, Maddage N C, Lech M, et al. Influence of acoustic low-level descriptors in the detection of clinical depression in adolescents[C]//2010 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2010: 5154-5157.
46. Cummins N, Epps J, Breakspear M, et al. An investigation of depressed speech detection: Features and normalization[C]//Twelfth Annual Conference of the International Speech Communication Association. 2011.
47. Cummins N, Epps J, Sethu V, et al. Variability compensation in small data: Oversampled extraction of i-vectors for the classification of depressed speech[C]//2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2014: 970-974.
48. Joshi J, Goecke R, Alghowinem S, et al. Multimodal assistive technologies for depression diagnosis and monitoring[J]. Journal on Multimodal User Interfaces, 2013, 7(3): 217-228.
49. Ozdas A, Shiavi R G, Silverman S E, et al. Investigation of vocal jitter and glottal flow spectrum as possible cues for depression and near-term suicidal risk[J]. Ieee transactions on Biomedical engineering, 2004, 51(9): 1530-1540.
50. Ma X, Yang H, Chen Q, et al. Depaudionet: An efficient deep model for audio based depression classification[C]//Proceedings of the 6th international workshop on audio/visual emotion challenge. 2016: 35-42.
51. Zheng W, Yan L, Gou C, et al. Graph attention model embedded with multi-modal knowledge for depression detection[C]//2020 IEEE International Conference on Multimedia and Expo (ICME). IEEE, 2020: 1-6.
52. Haque A, Guo M, Miner A S, et al. Measuring depression symptom severity from spoken language and 3D facial expressions[J]. arXiv preprint arXiv:1811.08592, 2018.
53. Williamson J R, Godoy E, Cha M, et al. Detecting depression using vocal, facial and semantic communication cues[C]//Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge. 2016: 11-18.
54. Al Hanai T, Ghassemi M M, Glass J R. Detecting Depression with Audio/Text Sequence Modeling of Interviews[C]//Interspeech. 2018: 1716-1720.
55. Lam G, Dongyan H, Lin W. Context-aware deep learning for multi-modal depression detection[C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019: 3946-3950.
56. Gong Y, Poellabauer C. Topic modeling based multi-modal depression detection[C]//Proceedings of the 7th Annual Workshop on Audio/Visual Emotion Challenge. 2017: 69-76.
57. William D, Suhartono D. Text-based Depression Detection on Social Media Posts: A Systematic Literature Review[J]. Procedia Computer Science, 2021, 179: 582-589.
58. Rodrigues Makiuchi M, Warnita T, Uto K, et al. Multimodal fusion of bert-cnn and gated cnn representations for depression detection[C]//Proceedings of the 9th International on Audio/Visual Emotion Challenge and Workshop. 2019: 55-63.
59. Yin S, Liang C, Ding H, et al. A multi-modal hierarchical recurrent neural network for depression detection[C]//Proceedings of the 9th International on Audio/Visual Emotion Challenge and Workshop. 2019: 65-71.
60. Ray A, Kumar S, Reddy R, et al. Multi-level attention network using text, audio and video for depression prediction[C]//Proceedings of the 9th International on Audio/Visual Emotion Challenge and Workshop. 2019: 81-88.
61. He L, Niu M, Tiwari P, et al. Deep Learning for Depression Recognition with Audiovisual Cues: A Review[J]. arXiv preprint arXiv:2106.00610, 2021.

1. Yuyan Bu finished this work with the guidance of Beiyu Lin [↑](#footnote-ref-1)