International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

Methodologies for Depression Detection using Smart Wearables

Ritom Gupta[1], Pranav Bakre[2], Pratik Gorade[3], Vignesh Iyer[4], Shamla Mantri[5]

[1-4]Final Year, B.Tech CSE, Dr. Vishwanath Karad MIT World Peace University, Pune, India
[5]Professor, Dr. Vishwanath Karad MIT World Peace University, Pune, India

---

Abstract - *Detecting depression using wearable smart devices is a rapidly emerging method in the detection and diagnosis of depressive disorders. The previous in this domain have used behavioural techniques, machine learning and correlation analysis to detect/diagnose depression. In this work, we recapitulate the methods of data collection, self-reporting mechanisms, decision methods and parameter correlation with depression tendencies. This work aims to give a direction for future research in this domain, by identifying the needs and concerns to be worked upon. We have identified the need for greater emphasis on intraindividual variability in mood, multimodal approaches, and general solutions.*

Keywords: Machine Learning, Correlation, Ecological Momentary Assessment, Depression Scale, DSM, HAMD, Multimodal Techniques

1. INTRODUCTION

This work is an in-depth review of depression detection techniques using smart wearables. Prominence is given to various approaches which establish correlation of parameters with depression. This work aims to throw some light on the existing works with the main questions being:

1) What are the parameters closely related to depression?
2) What are the most used techniques?
3) Which methods have been used to collect data?
4) What are the areas to further examine?

A thorough Internet search was done using key words like depression, fitness data, smart wearables. This work encapsulates the works based on quantitative analysis, advantages and limitations. Data collection protocols and the corresponding methods applied to analyse presence/severity of depression have been described.

2. RELATED WORK

Numerous studies have been carried out for predicting depression and measuring its severity using smartphones and wearables. Most of the studies have found a relation between predicted severity of and actual severity. The various mechanisms used so far have ranged from simple questionnaires and passive sensing to machine learning algorithms, EEG-based measurement and some combinations of these methods allied together in developing a system.

In person submission of recorded data in the organizer's lab after a period of having the wearable collecting their data has been a common method for collecting data. In most cases, data has been collected from specific groups like children, university students, unemployed men, women, etc. Number of people whose data was collected is between 10-100, possibly, due to the human effort involved in collection, limited funding and the time required to get enough data for thorough scanning for one person. The sensitivity of clinical data being high, prevents data from being made public. With the known mentioned practices in collection, we go through the related works in this field.

After gathering the answers to 11 question survey from the stakeholders (healthcare providers, people having or previously diagnosed with depression and their caregivers, healthcare insurance companies, etc), the results showed that majority of the stakeholders somewhat or strongly agree that the music streaming service can be used as an addition in therapy for depression. Music streaming accentuated as an additive in digital therapy for depression [1].

In another approach, 40 participants were categorized as depressed, their GPS data and phone usage data were monitored for 2 weeks which was used to derive 10 different parameters, namely: Location variance, Number of location clusters (K), Entropy, normalized entropy, home stay, circadian movement, total distance, phone usage frequency, phone usage duration. Linear regression was used to estimate the PHQ-9 score and actual PHQ-9 scores were used as a test. Logistic regressor classification was used to classify participants as depressed/not depressed. Significant correlation of variables was discovered between simplified entropy, varying location, staying home, and the PHQ-9 scores. Mobile usage data suggests correlation with parameters like usage duration, and usage frequency (Correlation Coefficient =.54, P-value=.011, and Correlation coefficient=.52, P-value=.015). An accuracy of 86.5% was achieved using the normalized entropy feature and a classifier that distinguished between participants having and not having depressive symptoms. It is noteworthy, that participants with PHQ-9 scores >=5 were considered having symptoms of disorder. A regression model that used the exact feature to find an estimate of the participants' PHQ-9 scores saw an average error of 23.5%. GPS data has also suggested a correlation with depressive symptom and its severity, naturally recurring movement or circadian movement (Correlation Coefficient =-0.63, P-value=0.005), mobility between favorite places (Correlation Coefficient=-0.58, P-value=0.012) and measure of location variability (Correlation Coefficient =-0.58, P-value=.012). [2]

Table 1: Correlation between data collected from phone and depressive tendencies [2]

Parameter	Correlation Value	P-value
Circadian Movement	-0.63	0.005
Mobility between favourite places	-0.58	0.012
Location Variance	-0.58	0.012
Usage Duration	0.54	0.011
Usage Frequency	0.52	0.015

Some works extended this exploratory study into development of context-aware systems. These systems detect when the patients need assistance. [4, 10-11] They involve using machine learning models (i.e., learners) which seek to predict patients' moods, emotions, cognitive/motivational states, activities, environmental context, and social context based on at least 38 concurrent phone sensor values (including GPS, ambient light & recent calls). A mobile phone application served as an interface for data collection for 38 sensor values. 8 adults with major depressive disorder were enrolled to receive the application and complete clinical assessments for 8 weeks. This is, in general considered as an easy to scale, reasonable intervention with preliminary evidence of efficacy. [4]

In a comprehensive approach to develop automatic depression assessment, P. Anastasia, P.G. Panagiotis, et al. have summarized visual manifestations of depression, existing datasets, and procedures from an ocean of image processing and machine learning based studies on depression. They summed up issues in Automatic Depression Detection, which are closely related to the context. The most important issue with data being the heavy regulatory overhead it bears. In imagery analysis, it is almost impossible to maintain anonymity. Licensing is restricted. There is a strong need for cooperation between institutions to share data under regulated conditions. This exhaustive analysis has yet again upheld the multimodal approaches and indicates deep learning being useful in depression detection purposes. [5]

An approach to evaluate depression using wearables was able to predict HAMD scores with a correlation coefficient of 0.61 and an accuracy of 76% for symptomatic patients. A patient's mean heart rate was 30- 200 beats per minute, it was included in the measurement. Heart rate data every 5 minutes, that is 12 times in an hour. To properly represent heart rate, data which had been collected for six or less times were discarded. 20+ hours of data in a day was later classified as daily data, one that had less than 20 hours was discarded. There was emphasis on using data which fell in acceptable biologically defined ranges. 63 parameters were extracted for each person's assessment 3-7 days before clinical assessment. A wearable device was used to calculate steps, movement, sleep time, heart rate, body temperature, heart rate and exposure to Ultraviolet rays. Every hour these parameters for patients and healthy volunteers were compared. The XG-Boost algorithm was used to build ML models. XG Boost and SVM was tried on small datasets, but XG-Boost gave better results, hence, XG Boost was used on the final dataset. 10-fold cross validation was applied to find the concreteness of the relation between the given parameter and depression. [6]

IoT based system with wearables and sensors have been proposed aiming to detect depressive disorders in an individual. A proposed system comprises of a wearable synced with an Arduino microcontroller. The system detects pulse rate, beats-per minute, oxygen saturation levels, skin temperature and uses an SVM classifier to the obtained data to detect depression. The input in the form of signal data is obtained from the wearable, it is then sent to the smartphone using the HC-05 Bluetooth module. The user can view the obtained data through the system and visit a professional if needed. [7]

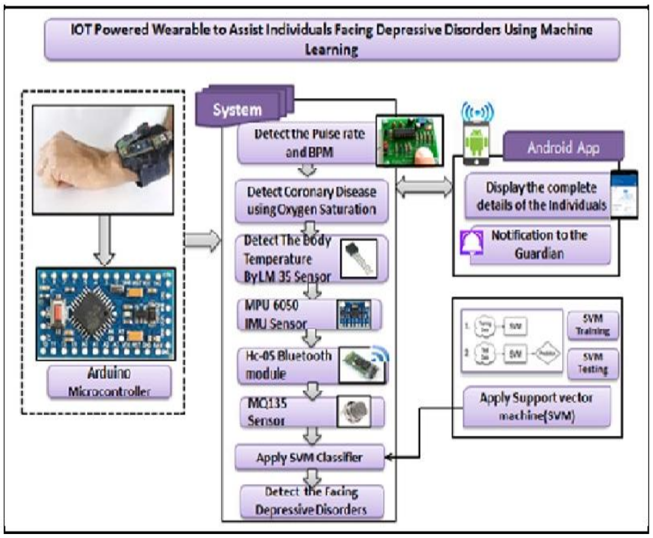

Figure 1: Proposed IoT based system [7]

In an approach to detect and diagnose people using computer programs, 60 participants at risk of depression (unemployed men) were divided into 3 groups and randomly put into one of the groups: 1) intervention program (22 of them), 2) intervention program plus sensors (19 of them) and 3) control group (19). The participants were measured for anxiety, depression, positive and negative effects and observable stress measures. Unemployed men suffering stress due to financial constraints and commitments, aged between 18 and 65, who gave their consent to take part in the study. Access to computers and internet was another important condition for inclusion in this study. EEG and ECG sensors were used to analyze physiological conditions and the strength of their cognitive abilities, and an ACT test to capture all kinds of movement. One group of participants had access to sensors (IP+S), one didn't have access to sensors (IP) and the control group answered the pre and post treatment questions at the end of assessment period. Feedback was taken from participants of all these groups. This method emphasizes on therapy using computers and tests a Computerized Cognitive Behavioural Therapy (CCBT) system for prevention and treatment of depression. It helps in learning different ways to manage everyday problems and symptoms. A combination of sensors and self-reporting, the former aided by ECG, ACT and EEG sensors and the latter by questionnaires on activity, feeling and mood. *C. Botella et al.* [8]

In a recent study that aimed to study the depressed mood variation with time, machine learning models that were nomothetic and ideographically-weighted were used. Outcomes of this study are significant because of the reduction in time intervals, which is a step in the direction to track depression in real time. 31 participants whose mood scores to reflect a significant variation over time were recruited for the study. They were given a wearable device and had to download an Android application. The number of hours they would be awake in the next 7 days was asked. Following this, the participants had to answer the questionnaire once an hour for ecological momentary assessment and heart rate assessment for the times they were awake. Data from passive sensors was collected the study period. The data was downloaded from the participants' app around a week later. Time to time depression and timely mood variation detection is a much closer to real-time depression detection. The long-time normal behavior reflected by a piece of data can outweigh the short-term data which can help in detecting depression at a nascent stage. Some systems aim at collection a large amount of self-reporting data at a very short period of time and use this data to predict the mood of an individual in the next time period. The comparison of predicted and observed mood indicate help us in evaluating the performance of such a system. Mood analysis can be done by assessing self-reporting data for dysphoria, hopelessness, devaluation, lack of interest/involvement or the inability to experience pleasure from usually pleasurable activities which are said to be core symptoms of MDD. The instantaneous or dynamic mode depends on the individual's loneliness and sadness. Higher scores, as expected indicate more depressed mood. Self-reported loneliness and sadness have found strong relations with depression. These self-reported parameters are further validated by passive sensor data. Number of features like location information, location typebased information, heart rate, outgoing calls frequency can be collected passively. Using Machine learning models, the last 24 hours data can be used to predict the next hour's mood. Using the nomothetic modelling outputs for idiographic modelling, an individual's idiographic patterns can be given a higher weightage. Random Forest Models have been found to be better than XG Boost models in terms of computational efficiency. This approach also helps in estimating hourly mood changes accurately. The correlation coefficient came out to be 0.587 across persons and 0.376 within persons. *N.C. Jacobson and Y.J Chung* [9]

Practical approaches with the aid of behavioral tasks, questionnaires, Smart Wearables, and statistical models have been applied to extracted features to identify children with incorporation disorders using a small behavioral task. Movement during the task was observed using a consumer-level wearable sensor. Children in the age group of 3 to 8 years who could speak English and whose parents/guardians were older than 18 years. Multimodal tests, clinical interviews were conducted for all but one child. Parents/Guardians answered a set of questions and the interview for the child's psychiatric assessment while the child went through a series of tasks in another room. Behavioral tasks, Questionnaires, Clinical interviews, Feature Extraction from Wearable Devices and Statistical Models for identifying internalizing disorders. Ryan McGinnis et al. [10]

Research works have pointed to the possibility of associating various speech patterns and their features to determine the depression state of the patient. [3] An extended version of the same adds facial images and EEG to the speech patterns and works on a multimodal system that puts together the output of all the three features and yields an HAMD score. Decision Fusion techniques have been used to blend the output of all these modalities to produce a single output. Multimodal systems have proved out to be more accurate compared to their unimodal subsets. Speech features like pitch, intensity and loudness were captured. [12]

Multimodal approaches have been different sets of parameters like talking practices, ocular activity and head posture. A clinically examined and validated dataset of 30 healthy and 30 severely depressed patients was obtained. Feature selection was done based on T-tests. Individual modality-based classification results were obtained and compared with multimodal systems after fusing all the unimodal components together. In this particular work, we can see the usage of multiple fusion techniques like feature fusion, decision fusion, score fusion and hybrid fusion techniques were tried. The best accuracy (88%) was produced by feature Fusion technique. Feature fusion was done using Principal Component Analysis, producing fused features. [13]

The individuals were filtered by using PHQ-9 and further DSM-5 was used to classify individuals into three categories of depressed and one category of control. A total of 17 features were retrieved for each axis of motion from both, the mobile and the smartwatch, along with step counter and significant motion sensor for a fixed window of 50 samples(1s). Heart Rate Monitor was also used as a feature. Incoming and Outgoing calls, Social Media app usages and call duration were also recorded from the phone. EMA reports were taken 6 times in the day starting from 7:00 a.m. Food Intake was recorded from the EMA reports. However, weight changes were not recorded in the study. Total data was collected over 30 days for 20 patients. The data and features were categorized into Action, Mood, Social Activity, Sleep and Food Intake. Two classification blocks were developed. A personalized model was created for each user for observational purposes whether there was a possibility of indicator level detection using passive sensing. Sensor data was trained with EMA data using Support Vector Machines (SVM) and Random Forests (RF) classifiers. 10-fold cross-validation was applied to the model and repeated 10 times using the final dataset, which was split into a training and testing ratio of 70 to 30. Finally, after comparison between the models, Random Forest Classifier was preferred which achieved an accuracy of more than 90%. Sleep classifier implementation was based on following states: when phone was being or not being used, and the light intensity of environment in extracted from a light sensor and measured in lux. The state was tracked during the potential hours of sleep. Longest unused period was classified as sleep time. To maintain scaling, baseline was set as 7 hours. The Second block was a common model for all participants to classify the depression category from the data collected from the devices. The model used the classified levels outcomes for each symptom cluster from personalized models as input. This was because the amount of data collected was not enough to build a model to classify directly using the features. The final feature vector for depression classification consisted of five elements, which were the levels for action, interaction, mood, sleep, and food intake. Total samples were supposed to be 3600 from 6 EMA per day with 30 days for 20 participants. However, only 2046 samples were used because of candidates missing EMAs. The data was once again trained with SVM and RF algorithms. For RF, the number of trees could go up to 100. Sleep was found to be

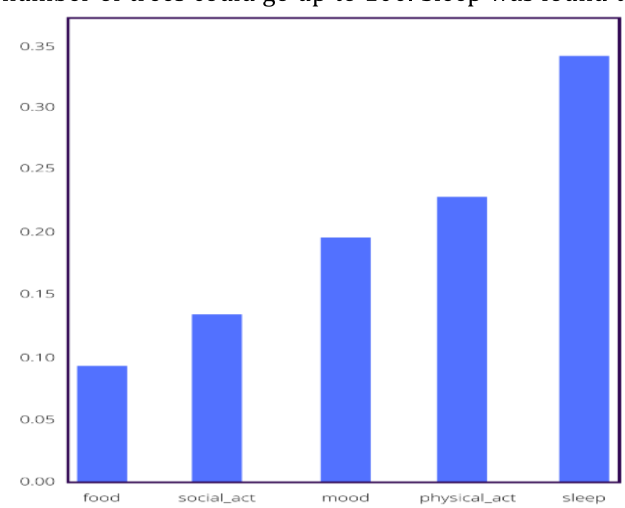

Figure 2: DSM-5 Feature Importance for depression classification model [14]

one of the highest indicators of depression with it contributing 34% to the model performance. It was found that for action level classification, notable motion and step features were highly important. For classifying mood, heart rate sensor data was more presiding. The severe group showed results contrary to assumption. [14]

In a Dartmouth University research, 83 students were analyzed across two terms of winter and spring. PHQ-8 test was conducted at the start and the conclusion of the observation period. The application inferred the student's activity (halted, moving, cycling, running), sleep (duration, bedtime, risetime) and sociability (i.e., the number of distinct conversations a participant is involved in and the timespan of the conversation). It also collected audio amplitude, location co-ordinates and lock/unlock events. Additionally, it administered PHQ-4 weekly to capture parameters related to depression. A smart band was used to record heart rate, galvanic skin response (GSR), skin temperature, and activity data, however, GSR and skin temperature were not used due to poor quality of the data. DSM-5 defines nine common symptoms connected with MDD: gloomy mood, loss of interest, sleep changes, weight loss or gain, feeling low energy, restlessness, diminished interest, reduced cognitive and concentration abilities. Set of passive sensor-based symptom features was derived from phones and wearables. Analysis of Variance was performed between the non-depressed and depressed groups. Features related to PHQ-4 depression were identified using regression analysis. Students who reported higher PHQ-4 scores have been found to be having less sleep, less conversations, go to bed late, wake up late and visit fewer places. Significant correlations have been found between derived from passive sensing which is collected from phone and the PHQ-9 depression. Features from GPS data, including circadian movement, normalized entropy, location variance, phone usage features, including usage duration and usage frequency are associated with depression symptom severity. The mobile app senses data and PHQ-4 EMA component in the app was built for the purposes of self-reporting. PHQ-8 and PHQ-4 was used as direct observation. PHQ-8 test was taken at the beginning and the end of the observation period while PHQ-4 was used to capture data once in a week. Also, the changes in dynamics were also recorded through the PHQ-4 test. Suicidal thoughts and suicide symptoms were not considered in this study. PHQ-4 helps in measuring diminished interest and depressed mood. [15]

Table 2: DSM Symptoms and related Features [15]

DSM Symptom	Features
Sleep Changes	Sleep duration, Sleep start and Sleep End
Reduced ability to concentrate	Unlock duration, unlock duration at dorm, unlock duration at study places
Taking less interest in things	Stationary, less conversations, time at dorm, time at study places,
Depressed Mood, fatigue, loss of energy	Heart Rate

The correlation between PHQ-8 scores and the above symptoms was studied.

Table 3: Correlations between symptoms and pre-post PHQ-8 scores [15]

Symptom measures	PHQ-8 pre		PHQ-8 post	
	r	P	r	P
Sleep Duration		>0.05		>0.05
Sleep start	0.236	0.059	0.301	0.024
Sleep End	0.183	0.145	0.271	0.043
Unlock Duration	0.282	0.010	0.268	0.024
Unlock Duration at dorm(mean)	0.245	0.027	0.206	0.085
Unlock duration at dorm (std)	0.270	0.014	0.222	0.062
Unlock duration at study places (mean)	0.391	<0.001	0.322	0.006
Unlock duration at study places (std)	0.260	0.018	0.120	0.319
Stationary Time	0.256	0.040	0.347	0.009
Conversation duration	0.467	<0.001	0.223	0.062
Number of places visited	-0.066	0.556	-0.269	0.023
Time at on-campus health facilities	0.210	0.059	0.029	0.812
Time at dorm		>0.05		>0.05
Time at study place		>0.05		>0.05
Heart rate		>0.05		>0.05

Table 4: Correlations between symptoms and pre-post PHQ-8 scores [15]

Symptom features	PHQ-8 pre		PHQ-8 post	
	r	P	r	P
PHQ-4 depression subscale (mean)	0.743	<0.001	0.849	<0.001

PHQ-4 depression subscale (std)	0.328	0.003	0.521	<0.001
PHQ-4 depression subscale (slope)	0.045	0.688	0.438	<0.001

PHQ-4 depression scale helped in tracking how depression dynamics change over the term. PHQ-4 responses correlated well with both pre PHQ-8 and post PHQ-8 results. Pre-PHQ-8 and Post-PHQ-8 scores could be predicted by lasso regularization by selecting a few parameters with a mean absolute error of 3.44 and 3.60. There is a possibility of heart rate being combined with some other parameters that could enhance the accuracy of PHQ-8 predictions. Correlation was tested between PHQ-4 and 2-week symptom features. The data suggests students who have scarce or negligible number of conversations, less sleep and visited fewer places are more likely to be depressed. Regularization of PHQ-4 results also pointed to heart rate being a key parameter in predicting and detecting depressed mood. 9 features were lasso selected for further analysis and ROC curve was obtained from a logistic regression model after applying 10-fold cross validation. The area under the ROC curve was found out to be 0.809. The model achieved 81.5% recall and 69.1% precision. [15]

3. CONCLUSION AND FUTURE WORK

Research on detecting depression using smart wearables is an evolving topic for research which is powered by new methodologies, more powerful algorithms and evolution of the consumer level technologies in offering more features and precision. Considering the noticeable development in the field, reducing human efforts in collecting data is likely to be helpful for research and diagnosis.

Generally, the hypothesis has been in sync with the clinical outcomes. But the scope of this is limited to a specific piece of research. There is a need for commensuration which establishes the benchmarks for comparison of results. Given the current scenario, machine learning algorithms have been particularly dominant in the classification and regression process.

Clinical research on depression also finds several queries which need to be answered. Variability in an individual due variety of possible reasons has not been given emphasis, with the exception of Nicholas C. Jacobson and Yeon Joo Chung [9], who examined intraindividual mood changes by collecting EMA data more frequently. Multimodal approaches have yielded better accuracy, few research works have seemed to have used multimodal techniques, there is a scope for multimodal analysis with more parameters. A large chunk of research has seen a homogeneous group of people being under observation. There is a need for a more generalized study which uses customization in algorithm to make up for the difference in social groupings.

In conclusion, this review sums up the present techniques and highlights the considerations which shall aid future research for producing more accurate clinical results.

ACKNOWLEDGEMENT

We wish to express our sincere gratitude to Prof. Dr. Shamla Mantri as well as Prof. Dr. Vrushali Kulkarni, H.O.S, School of CET, MIT World Peace University, Pune for guiding us in this survey. We also thank our friends and other MIT World peace University staff members for guidance and encouragement in carrying out this work. We thank them for their valuable guidance and support.

REFERENCES

[1] S. Gadd, C. Tak, and G. Bulaj, “Developing music streaming as an adjunct digital therapy for depression: A survey study to assess support from key stakeholders,” Journal of Affective Disorders Report, vol. 2, Dec. 2020.

[2] S. Saeb, M. Zhang, C. J. Karr, S. Schueller, M. Corden, K. P. Kording, and D. C. Mohr, “Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study,” Journal of Medical Internet Research vol. 17, no. 7, Jul. 2015.

[3] D. Patil, S. Mantri, R. Agrawal, S. Bhattad, A. Padiya, and R. Rakshit, “A Survey: Pre-processing and Feature Extraction Techniques for Depression Analysis Using Speech Signal,” International Journal of Computer Science Trends and Technology, vol. 2, no. 2, pp. 83-89, Mar-Apr. 2014.

[4] M. N. Burns, M. Begale, J. Duffecy, D. Gergle, C. J. Karr, E. Giangrande, and D. C. Mohr, “Harnessing Context Sensing to Develop a Mobile Intervention for Depression,” Journal of Medical Internet Research, vol. 13, no. 3, Jul-Sep. 2011.

[5] P. Anastasia, Simos P.G, M. Kostas, M. Fabrice, Yang F, Pediaditis M., and T. Manolis, “Automatic Assessment of Depression Based on Visual Cues: A Systematic Review”,

IEEE Transactions on Affective Computing, Vol. 10, No. 4, Oct-Dec 2019

[6] Y. Tazawa, K. Liang, M. Yoshimura, M. Kitazawa, Y. Kaise, A. Takamiya, A. Kishi, T. Horigome, Y. Mitsukura, M. Mimura, and T. Kishimoto, "Evaluating depression with multimodal wristband-type wearable device: screening and assessing patient severity utilizing machine-learning," Heliyon vol. 6, no. 2, Feb. 2020.

[7] V. V. Waykule, D. Mahale, M. Kaledhonkar, N. Habib, and V. Surwase, "IOT Powered Wearable to Assist Individuals Facing Depressive Disorders using Machine Learning," International Journal for Research in Applied Science & Engineering Technology, vol. 8, issue IX, Sep. 2020

[8] C. Botella, A. Mira, I. Moragrega, A. Garcia-Palacios, J. Breton-Lopez, D. Castilla, A. Riera Lopez del Amo, C. Soler, G. Molinari, S. Quero, V. Guillen-Botella, I. Miralles, S. Nebot, B. Serrano, D. Majoe, M. Alcaniz, and R. Maria Banos, "An Internet Based Program for depression using activity and physiology sensors: efficacy, expectations, satisfaction and ease of use," Neuropsychiatric Disease and Treatment, vol. 12, pp. 393–406, Feb. 2016.

[9] N. C. Jacobson, and Y. J. Chung, "Passive Sensing of Prediction of Moment-To-Moment Depressed Mood among Undergraduates with Clinical Levels of Depression Sample Using Smartphones," Sensors, vol. 20, no. 12, Jun. 2020.

[10] R. S. McGinnis, E. W. McGinnis, J. Hruschak, N. L. Lopez-Duran, K. Fitzgerald, K. L. Rosenblum, and M. Muzik, "Rapid detection of internalizing diagnosis in young children enabled by wearable sensors and machine learning," PLoS ONE, vol. 14, no. 1, Jan. 2019.

[11] R. S. McGinnis, E. W. McGinnis, J. Hruschak, N. L. Lopez-Duran, K. Fitzgerald, K. L. Rosenblum, and M. Muzik "Rapid Anxiety and Depression Diagnosis in Young Children Enabled by Wearable Sensors and Machine Learning," Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3983-3986, Jul. 2018.

[12] D. Patil, S. Mantri, V. Wadhai, and P. Agrawal, "Real time Multimodal Depression Analysis," International Journal of Innovative Technology and Exploring Engineering, vol. 8, issue 9, Jul. 2019.

[13] S. Alghowinem *et al.*, "Multimodal Depression Detection: Fusion Analysis of Paralinguistic, Head Pose and Eye Gaze Behaviors," in *IEEE Transactions on Affective Computing*, vol. 9, no. 4, pp. 478-490, 1 Oct.-Dec. 2018, doi: 10.1109/TAFFC.2016.2634527.

[14] N. Narziev, H. Goh, K. Toshnazarov, S. Ah Lee, K. Chung, and Y. Noh, "STDD: Short-Term Depression Detection with Passive Sensing," Sensors, vol. 20, no. 5, Mar. 2020.

[15] R. Wang, W. Wang, A. daSilva, J. F. Huckins, W. M. Kelley, T. F. Heatherton, and A. T. Campbell. "Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing." Proceedings of the ACM on interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 2, no. 1, article 43, Mar. 2018.