International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Machine Learning and Deep Learning Multi-Modal Approaches in Mental Health Diagnosis: A Comprehensive Review

Apurv Panchal[1], Karan Bhatt[2], Rahul Shah[3], Ashok Rathva[4], Jwalant Baria[5], Yagnik Rathod[6]

Submitted: 21/05/2023 **Revised:** 11/07/2023 **Accepted :** 30/07/2023

Abstract: This paper receives recent advancements in Machine learning (ML) and Deep Learning (DL) approaches applied to mental health diagnosis over the past years. We examine a range of techniques, including Supervised and Unsupervised learning, Natural Language Processing (NLP), and image recognition in neuroimaging. The review encompasses various data sources, such as electronic health records, social media and wearable sensor outputs, targeting diagnoses of depression, anxiety, bipolar disorder, schizophrenia, PTSD, anorexia nervosa and ADHD. The analysis explores the use of various machine learning (ML) models such as support vector machines, decision trees, random forests and ensemble methods. Findings also includes deep learning models such as CNNs, RNNs, transformer-based models, have been applied to multi-modal data such as texts, speech and image data to extract meaningful features and patterns that are indicative of mental health conditions. This review concludes by discussing implications of integrating ML and DL approaches into clinical practice, emphasizing the importance of interdisciplinary collaboration among data scientists and mental health professionals.

Keywords: *Deep learning, Healthcare, Machine learning, Mental disorders, Mental health diagnosis*

1. Introduction

Mental health disorders represent a significant global health burden, affecting individuals across all age groups and societies. According to the WHO, approximately 1 in 8 people worldwide live with a mental disorder, with depression and anxiety disorders being the most prevalent. Traditional diagnostic methods often rely on clinical interviews and subjective assessment tools, which can be time-consuming and prone to human error. In response to these challenges, there has been a surge in research focusing on the application of machine learning (ML) and deep learning (DL) to improve the accuracy and efficiency of mental health diagnoses.

Recent advancements in deep learning have improved the classification of Autism Spectrum Disorder (ASD) using neuroimaging data. A new deep learning architecture in [1], tested on the ABIDE I dataset with rs-fMRI scans, achieved a 72.46% classification accuracy by leveraging functional connectivity (FC) features. This method outperformed existing approaches, highlighting its potential to enhance ASD diagnosis through advanced neuroimaging and deep learning techniques.

This article [2] provides a detailed comparative analysis of current wearable and handheld safety devices, evaluating them based on factors like power efficiency, weight, affordability, and user experience. In the context of developing wearables for mental health diagnosis, ensuring user comfort, affordability, and seamless integration into daily life is crucial. While smart wearables have advanced in worker safety, similar challenges exist in mental health applications, such as the need for non-intrusive, cost-effective devices. The insights from wearable technology in industrial settings highlight the importance of designing mental health wearables that are power-efficient, lightweight, and user-friendly, ensuring they can be effectively incorporated into everyday use for accurate and continuous mental health monitoring.

Magnetic Resonance Imaging (MRI) is vital for disease diagnosis and treatment, but artefacts and noise can degrade image quality. To address this, a new hybrid noise reduction framework is proposed in [3], combining wavelet transform, exponential function thresholding optimized by a Genetic algorithm, and the Wiener filter. This method enhances robustness against various noise types, such as Gaussian and Rician noise. Performance measures like Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) indicate that the proposed approach outperforms existing methods, providing clearer MR images with better noise reduction and less blurring.

[1] *UG Student, Department of Computer Science & Engineering (Data Science), Vishwakarma Government Engineering College, Chandkheda*
Email: apurv4605@gmail.com
[2]*Assistant Professor, Department of Computer Engineering, Vishwakarma Government Engineering College, Chandkheda*
Email: kpbhatt@vgecg.ac.in
[3] *Assistant Professor, Department of Computer Engineering, Vishwakarma Government Engineering College, Chandkheda*
Email: rkshah@vgecg.ac.in
[4] *Assistant Professor, Department of Computer Engineering, Government Engineering College, Dahod*
Email: akr.coed@gecdahod.ac.in
[5] *Assistant Professor, Department of Computer Engineering, Government Engineering College, Dahod*
Email: jwalant.baria@gmail.com
[6] *Assistant Professor, Department of Computer Engineering, Government Engineering College, Dahod*
Email: yar.coed@gecdahod.ac.in

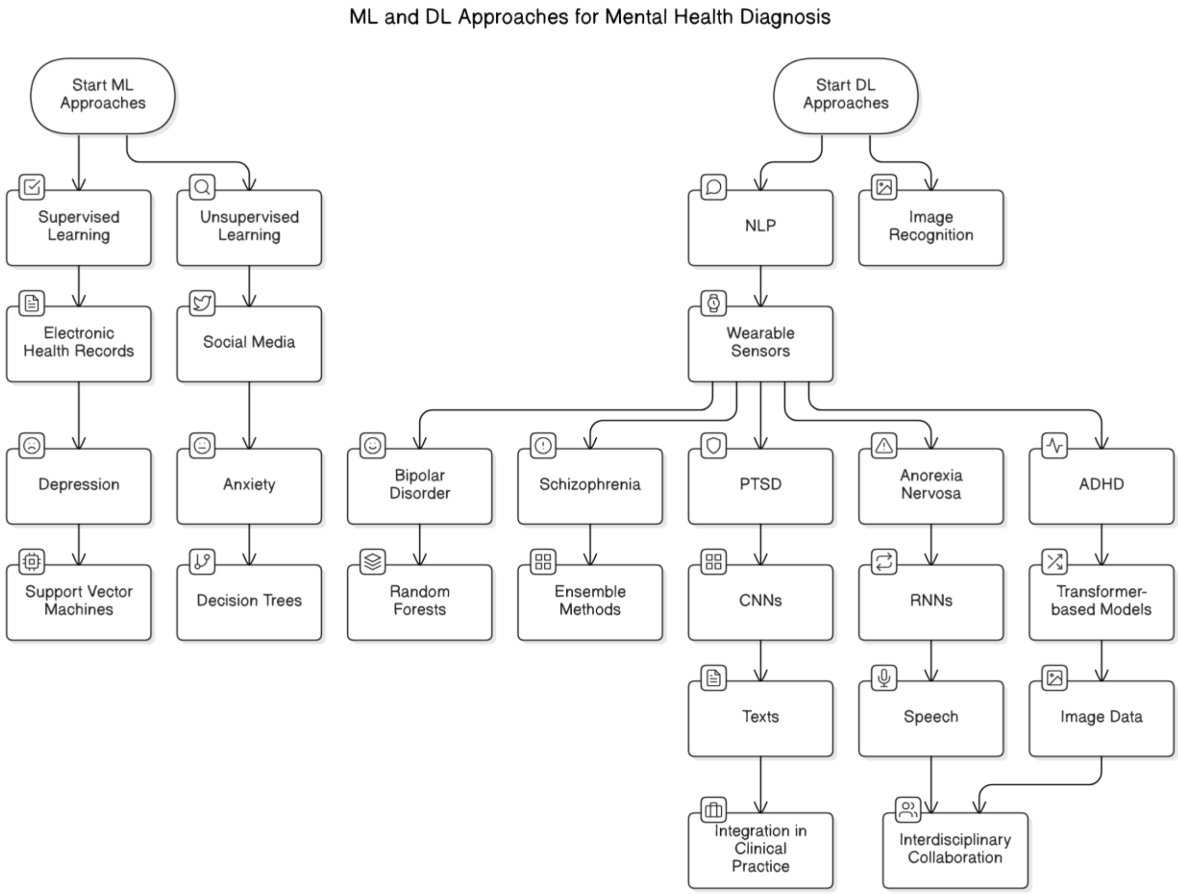

Fig.1. Flowchart of Methodology

2. Literature Studies of methodologies

2.1. Methods for Schizophrenia Prediction

A Spatial Sequence Attention Network (SSAN), a sequence-based spatial attention mechanism, that captures fine-grained patterns in brain structures from structural MRI data that are indicative of schizophrenia [4]. By leveraging transfer learning with pre-trained DenseNet to extract initial feature maps, the SSA enhances these features to capture detailed spatial interactions within the brain. Experimental results on a clinical dataset show that this approach surpasses existing methods, such as SqueezeNet, includes an accuracy of 87%, an increase of 10% over baseline methods, with an area under the receiver operating characteristic curve (AUROC) reaching 0.85.

A different approach by [5] explores using temporal patterns in motor activity data. From 22 schizophrenia patients and 32 controls over an average of 12.7 days, the research evaluates the impact of various temporal segmentations on classification performance. Sixteen statistical features are extracted and tested across seven ML models. It achieves an AUC-ROC of 0.98 and an F1 score of 0.93 when using segmentation. The findings suggest that a day-night partitioning approach is effective for schizophrenia classification.

[6] focuses on improving schizophrenia diagnosis using deep learning and EEG brain recordings. Spectrograms were extracted from EEG signals, and a convolutional neural network (CNN) was used for initial diagnosis. Two synthetic datasets were generated using Wasserstein GAN with Gradient Penalty (WGAN-GP) and Variational Autoencoder (VAE) to enhance the original dataset and mitigate overfitting. The VAE-augmented dataset improved accuracy to 99.0% and demonstrated faster convergence. The study also utilized Local Interpretable Model-agnostic Explanations (LIME) to improve trust in the model by identifying
key features in the diagnosis process.

[7] focuses on enhancing brain disorder prediction accuracy by examining disconnected subnetworks and graph structures related to schizophrenia. It introduces the structural connectivity-deep graph neural network (sc-DGNN) model, which outperforms traditional machine learning (ML) and deep learning (DL) methods. Utilizing diffusion magnetic resonance imaging (dMRI) data from eighty-eight subjects, the sc-DGNN model achieved a 93% accuracy rate in classifying schizophrenia, compared to 72% accuracy with linear discriminant analysis (LDA) in ML models.

2.2. Methods for Depression and Anxiety Detection

[8] addresses the challenge of detecting depression and anxiety in Twitter users through textual data. It proposes using Bidirectional Long Short-Term Memory (BiLSTM) networks,

which improve upon traditional Long Short-Term Memory (LSTM) networks by reading text in both directions, capturing more context and meaning. The BiLSTM model outperforms standard LSTMs and traditional machine learning models, achieving a high accuracy of 94.12% in identifying users with potential depression and anxiety.

[9] introduces the Depression and Anxiety Predictor (DAP), a multi-layer perceptron model designed to track and predict anxiety and depression trends during the emergency. Utilizing a dataset that tracked mental health symptoms weekly over ten weeks during the initial wave of the pandemic, the model analyzes patterns in a diverse cohort of U.S. adults. DAP provides insights into how demographic factors, behavioural changes, and social determinants interact with mental health during crises.

[10] explores using deep learning (DL) models to identify geriatric depression and anxiety by analyzing time-series data from consumer-grade wrist-worn activity trackers. It aims to provide a more precise and cost-effective alternative to traditional self-reported assessments and costly diagnostic methods. The DL model processes both time-series data (step counts and sleep stages) and minimal assessment scores for multi-label classification of depression and anxiety. Various models, including Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Residual Networks (ResNet), were tested. The ResNet model achieved notable results with a Hamming loss score of 0.0946, demonstrating the effectiveness of the approach. This study is the first to develop a mixed-input DL model for identifying late-life depression and anxiety using activity tracking data, highlighting the potential for improving mental health diagnosis in older adults.

[11] introduces PupilSense, a mobile system that uses deep learning to track pupillary responses through smartphone interactions to detect depressive episodes. They developed three models, with the top-performing one achieving an AUROC of 0.71. This model outperformed individual sensor-based models (Bluetooth, Calls, GPS, Steps) and surpassed a smartphone facial image-based depression detection system with a balanced accuracy of 60%. Additionally, it performed better than a recent wearable data-based depression model with an F1 score of 0.68.

2.3. Methods for Bipolar Disorder Detection

[12] introduces a novel multimodal fusion approach that combines brain structural magnetic resonance imaging (sMRI) scans with DNA whole-exome sequencing (WES) data to improve detection rates. The research analyzed data from 321 East Asian participants: 147 with MDD, 78 with BD, and 96 healthy controls. Six fusion models were developed and tested using popular deep learning and machine learning frameworks like Vision Transformer (ViT), Inception-V3, ResNet50, XGBoost, and LightGBM, with a 10-fold cross-validation for model validation. The most effective model, ViT $\oplus$ XGBoost, utilized MRI scans, genomic Single Nucleotide Polymorphism (SNP) data, and unweighted polygenic risk scores (PRS). This model showed a significant improvement in performance, with a 32.03% increase in the area under the curve (AUC) and a 25.14% increase in accuracy over SNP-only models, and an 8.19% AUC and 13.28% accuracy increase compared to image-only models.

[13] aims to analyse the unique linguistic characteristics of posts in mental disorder subreddits and compare them with posts in subreddits unrelated to mental illness to validate their distinctiveness. The research collected 3,133,509 Reddit posts from 919,722 users. Linguistic Inquiry and Word Count (LIWC) software, along with 1-way ANOVA and post hoc tests, were used for statistical analysis to identify sentiment differences in various lexical features. Textual features were extracted using Bidirectional Encoder Representations from Transformers (BERT) to perform supervised and unsupervised clustering. Clustering methods using BERT embeddings highlighted distinctive features for each subreddit group, with Davies-Bouldin scores ranging from 0.222 to 0.397 and silhouette scores from 0.639 to 0.803 for supervised clustering, and scores of 1.638 and 0.729, respectively, for unsupervised clustering.

2.4. Methods for Post-Traumatic Stress Disorder Detection

Post-traumatic stress disorder (PTSD) is a mental disorder that can develop after experiencing or witnessing extremely traumatic events, affecting individuals of all ethnicities and cultures. Traditional diagnostic methods, such as the Clinician-Administered PTSD Scale (CAPS) and the PTSD Check List for Civilians (PCL-C), rely on questionnaires that can be manipulated by the respondents' answers.

[14] introduces a deep learning-based approach to detect PTSD using audio recordings from clinical interviews. By extracting low-level Mel-frequency cepstral coefficients (MFCC) features from the audio and applying high-level learning with a Stochastic Transformer, the proposed method achieves state-of-the-art performance. Specifically, the approach attained a root mean square error (RMSE) of 2.92 on the eDAIC dataset, leveraging stochastic depth, stochastic deep learning layers, and a stochastic activation function to enhance accuracy.

[15] explores the ability of the Med-PaLM 2 large language model (LLM), which is trained on extensive medical knowledge, to predict psychiatric functioning from patient interviews and clinical descriptions without specific training for these tasks. The research examined 145 depression assessments, 115 PTSD assessments, and 46 clinical case studies of various high-prevalence disorders, including depressive, anxiety, psychotic, trauma and stress-related, and addictive disorders. By utilizing prompts to extract estimated clinical scores and diagnoses, it was found that Med-PaLM 2 can effectively assess psychiatric functioning across multiple conditions. The model demonstrated its strongest performance in predicting depression scores based on standardized assessments, achieving an accuracy range of 0.80 to 0.84, which is statistically comparable to human clinical ratters. The findings suggest that general clinical language models hold promise for predicting psychiatric risk through descriptions of functioning provided by both patients and clinicians.

2.5. Methods for Anorexia Nervosa Detection

[16] developed a therapeutic diagnostic tool based on sentiment analysis specifically for diagnosing anorexia nervosa.

They used a dataset consisting of 44 anorexic and 52 healthy girls aged 12 to 18. Their approach involved analyzing patients' statements about their bodies using sentiment analysis methods based on Recurrent Neural Networks (RNN) and a dictionary-based sentiment analysis. Their study found that the RNN method achieved a 72% effectiveness rate in diagnosing anorexia, outperforming other methods.

[17] utilized data from the CLEF eRisk 2018 challenge (Task 2) to detect signs of anorexia nervosa from textual data. They applied various classification methods, including AdaBoost, Logistic Regression (LR), Random Forest (RF), and Support Vector Machines (SVM), to validate and combine Bag of Words (BOW) and Unified Medical Language System (UMLS) features. Among these, SVM showed superior performance on BOW features with a precision of 0.97, Recall of 0.98, and an F-measure of 0.98. For UMLS features, SVM also performed well but with a lower F-measure of 0.55. When combining BOW and UMLS features, the AdaBoost classifier outperformed others with an F-measure of 0.47.

[18] conducted a study using genome genotyping data from 390 anorexia patients and 9,266 non-anorexic patients to predict the risk of anorexia nervosa. They split their dataset randomly into training and test sets and compared the performance of Logistic Regression (LR), SVM, and Gradient Boosting Trees. This comparison aimed to assess the effectiveness of these models in predicting anorexia risk.

Transfer learning was applied to diagnose anorexia nervosa using Deep Neural Networks (DNNs) and transformer-based models with Spanish tweets. The Spanish Anorexia Detection (SAD) dataset was utilized for model evaluation. While no significant difference was noted in the DNN model, Convolutional Neural Networks (CNN) slightly outperformed Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM). Among the transformer-based models, the BETO model achieved the best performance with an F1-Score of 94.1% [19].

2.6. Approaches for Attention Deficit Hyperactivity Disorder (ADHD) Detection

Neurodevelopmental disorders (NDD), including Intellectual Disability, ADHD, and Autism Spectrum Disorder, often involve attention deficits due to sensory system dysfunctions. These deficits negatively impact the quality of life and the ability to apply knowledge in different situations. Given the interactive nature of virtual reality (VR), it has emerged as a beneficial tool for learning and rehabilitation in NDD individuals. [20] studied, a VR application named Wildcard, previously used to improve attention skills in NDD participants, was redesigned to incorporate eye-tracking technology for more precise interactions. A four-week experiment with 38 NDD participants was conducted to assess the application's usability and effectiveness in enhancing visual attention skills. The results demonstrated the application's effectiveness, supporting the potential of VR and eye-tracking technologies in NDD interventions.

Diagnosing Autism Spectrum Disorder (ASD) can be challenging due to its strong comorbid similarities with other neurodevelopmental disorders (NDDs), like ADHD. To address this issue, a novel framework called OpenNDD was developed by, [21] for open set recognition to aid in ASD diagnosis. This framework combines autoencoder and adversarial reciprocal points learning to differentiate between known and unknown classes, enhancing the accuracy of ASD identification. Additionally, a joint scaling method using MinMax scaling and Standardization (MMS) was introduced to improve class differentiation. Experiments conducted on hybrid datasets from the Autism Brain Imaging Data Exchange I (ABIDE I) and ADHD-200 samples, totalling 791 samples from four sites, demonstrated the effectiveness of OpenNDD. The framework achieved an accuracy of 77.38%, an AUROC of 75.53%, and an open set classification rate of 59.43%, showcasing its promising performance in distinguishing ASD from other NDDs.

ADHD is common in younger populations but can be difficult to diagnose due to symptom overlap with other disorders like depression and conduct disorder. [22] explores advanced machine learning techniques to improve ADHD diagnosis using the ADHD200 dataset. The research involves applying various machine learning methods, including neural networks and support vector machines (SVMs), to classify ADHD based on phenotypic data and functional MRI connectivity. The SVM model, used for multiclass classification on phenotypic data, outperformed other supervised learning methods such as Logistic Regression and KNN. Neural networks were employed on MRI data from 40 subjects to achieve high accuracy without prior neuroscience knowledge. An ensemble technique combining the phenotypic classifier with the neural network model was developed and tested on 400 subjects from the ADHD200 dataset. This ensemble approach achieved a binary classification accuracy of 92.5%, with training and testing accuracies reaching up to 99%.

3. Results and Discussions

The advancements in ML and DL approaches have significantly improved the accuracy and efficiency of mental health diagnosis across various disorders. These methods provide a promising alternative to traditional diagnostic techniques, offering non-invasive, cost-effective, and scalable solutions. However, challenges remain, including the need for large, diverse datasets, addressing biases in models, and ensuring the ethical use of data. Future research should focus on integrating multimodal data sources, developing interpretable models, and validating these approaches in real-world clinical settings to enhance their applicability and reliability.

4. Conclusion

This review emphasizes the advancements in machine learning (ML) and deep learning (DL) for diagnosing and predicting mental health disorders such as schizophrenia, depression, anxiety, bipolar disorder, post-traumatic stress disorder (PTSD), anorexia nervosa, and attention deficit hyperactivity disorder (ADHD). These studies illustrate the potential of AI-driven diagnostics to enhance the accuracy and accessibility of mental health assessments. Future research should aim to refine these models, explore new data sources, and ensure the ethical use of AI in mental health care, ultimately improving outcomes for patients and healthcare providers.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Khosla M, Jamison K, Ngo GH, Kuceyeski A, Sabuncu MR. Machine learning in resting-state fMRI analysis. Magnetic resonance imaging. 2019;64:101-21. 10.1016/j.mri.2019.05.031

[2] Fadaei S. New dominant color descriptor features based on weighting of more informative pixels using suitable masks for content-based image retrieval. International Journal of Engineering, Transaction B: Applications. 2022;35(8):1457-67. 10.5829/IJE.2022.35.08B.01

[3] Hasan C, Jailani R, Tahir N, Desa H. Vertical ground reaction force gait patterns during walking in children with autism spectrum disorders. International Journal of Engineering, Transaction B: Applications. 2018;31(5):705-11. 10.5829/ije.2018.31.05b.04

[4] Wang L, Li K, Hu XP. Graph convolutional network for fMRI analysis based on connectivity neighborhood. Network Neuroscience. 2021;5(1):83-95. 10.1162/netn_a_00171

[5] Yang T, Al-Duailij MA, Bozdag S, Saeed F, editors. Classification of autism spectrum disorder using rs-fmri data and graph convolutional networks. 2022 IEEE International Conference on Big Data (Big Data); 2022: IEEE. 10.1109/BigData55660.2022.10021070

[6] Hamidi H, Vafaei A, Monadjemi SAH. Analysis and evaluation of a new algorithm based fault tolerance for computing systems. International Journal of Grid and High Performance Computing (IJGHPC). 2012;4(1):37-51. 10.4018/jghpc.2012010103

[7] Eslami T, Saeed F, editors. Auto-ASD-network: a technique based on deep learning and support vector machines for diagnosing autism spectrum disorder using fMRI data. Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics; 2019. 10.1145/3307339.3343482

[8] K. Setyo Nugroho, I. Akbar, and A. Nizar Suksmawati, "Seminar Nasional Hasil Riset Prefix-RTR DETEKSI DEPRESI DAN KECEMASAN PENGGUNA TWITTER MENGGUNAKAN BIDIRECTIONAL LSTM," 2021.

[9] Shao L, Fu C, You Y, Fu D. Classification of ASD based on fMRI data with deep learning. Cognitive Neurodynamics. 2021;15(6):961-74. 10.1007/s11571-021-09683-0

[10] Karampasi A, Kakkos I, Miloulis S-T, Zorzos I, Dimitrakopoulos GN, Gkiatis K, et al., editors. A machine learning fMRI approach in the diagnosis of autism. 2020 IEEE international conference on big data (big data); 2020: IEEE. 10.1109/BigData50022.2020.9378453

[11] Dvornek NC, Ventola P, Pelphrey KA, Duncan JS, editors. Identifying autism from resting-state fMRI using long short-term memory networks. Machine Learning in Medical Imaging: 8th International Workshop, MLMI 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 10, 2017, Proceedings 8; 2017: Springer. 10.1007/978-3-319-67389-9_42

[12] Zuo X-N, Xu T, Jiang L, Yang Z, Cao X-Y, He Y, et al. Toward reliable characterization of functional homogeneity in the human brain: preprocessing, scan duration, imaging resolution and computational space. Neuroimage. 2013; 65:374-86. 10.1016/j.neuroimage.2012.10.017

[13] Craddock RC, James GA, Holtzheimer III PE, Hu XP, Mayberg HS. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Human brain mapping. 2012;33(8):1914-28. 10.1002/hbm.21333

[14] Rashno E, Rashno A, Fadaei S, editors. Fluid segmentation in neutrosophic domain. 2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS); 2019: IEEE. 10.1109/ICSPIS48872.2019.9066085

[15] S. C. Guntuku, D. B. Yaden, M. L. Kern, L. H. Ungar, and J. C. Eichstaedt, Detecting depression and mental illness on social media: an integrative review, Current Opinion in Behavioral Sciences, vol. 18, pp. 43-49, 2017.

[16] D. Spinczyk, M. Bas, M. Dzieciątko, M. Maćkowski, K. Rojewska, and S. Maćkowska, "Computer-aided therapeutic diagnosis for anorexia," BioMedical Engineering OnLine, vol. 19, no. 1, pp. 53–53, Dec. 2020, doi: 10.1186/s12938-020-00798-9.

[17] S. Paul, J. Kalyani, and T. Basu, Early Detection of Signs of Anorexia and Depression Over Social Media using Effective Machine Learning Frameworks. 2018.

[18] J. A. Cully, D. E. Jimenez, T. A. Ledoux, and A. Deswal, Recognition and treat ment of depression and anxiety symptoms in heart failure, Primary Care Com panion to the Journal of Clinical Psychiatry, vol. 11, no. 3, pp. 103109, 2009.

[19] P. López-Úbeda, F. M. Plaza-del-Arco, M. C. Díaz-Galiano, and M.-T. Martín-Valdivia, "How Successful Is Transfer Learning for Detecting Anorexia on Social Media?," Applied Sciences, vol. 11, no. 4, pp. 1838–1838, Feb. 2021, doi: 10.3390/app11041838.

[20] Smith SM. The future of FMRI connectivity. Neuroimage. 2012;62(2):1257-66. 10.1016/j.neuroimage.2012.01.022

[21] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, predicting depression via social media, in Proceedings of ICWSM, 2013, pp. 1-10.

[22] A. Sharma, A. Jain, S. Sharma, A. Gupta, P. Jain, and S. P. Mohanty, "iPAL: A Machine Learning Based Smart Healthcare Framework for Automatic Diagnosis of Attention Deficit/Hyperactivity Disorder (ADHD)," Feb. 2023, [Online]. Available: http://arxiv.org/abs/2302.00332