

Exploring MODMA Database for Major Depressive Disorder Diagnostic

Ysabelle Sousa, Jefferson Martins, Kenji Nose-Filho and Aline Neves

Abstract—Depression is a prevalent psychiatric disorder associated with diagnostic challenges, as conventional assessment tools such as the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) and Hamilton Depression Rating Scale (HAMD) may produce inconsistent results. This study investigates the potential of electroencephalogram (EEG) and audio biomarkers to enhance diagnostic accuracy using the Multi-modal Open Dataset for Mental-disorder Analysis (MODMA) dataset. We employed a comprehensive pipeline encompassing preprocessing, feature extraction, feature selection using Minimum Redundancy Maximum Relevance (mRMR), and classification using Support Vector Machine (SVM) for both EEG and audio signals. The EEG-based classification achieved an accuracy of 76%, while the audio-based models reached 80.90% accuracy. These results suggest that speech features may offer a slight advantage in discriminating individuals with major depressive disorder (MDD).

Keywords—Major Depressive Disorder, Electroencephalogram, Audio signal processing, Biomarkers, Machine Learning.

I. INTRODUCTION

Considered a global issue, Major Depressive Disorder (MDD), or Depression, is a distinct neurological condition that goes beyond typical episodes of sadness. According to the World Health Organization [1], depression affects approximately 332 million individuals globally, with a 50% higher occurrence rate among women. Currently, diagnosis relies on the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) criteria and the Hamilton Depression Rating Scale (HAMD) [2], which are commonly used to assess symptom severity. The HAMD categorizes severity based on symptoms such as sadness, loss of interest, sleep disturbances, and suicidal thoughts. This classification may have lack of accuracy, since there is no tangible evidence to describe the severity of symptoms reliably.

Recent technological advancements have introduced new innovative solutions in healthcare. Nowadays, Artificial Intelligence (AI), particularly Machine Learning (ML), enables computers to learn from large datasets and previous experiences. In healthcare, ML supports diagnosis and treatment, especially for cancer, nervous system disorders, and cardiovascular diseases [3]. For mental and neurological disorders, ML has shown promising results in identifying brain-related features that help distinguish patients from healthy individuals, thereby contributing to more accurate diagnoses [1].

Ysabelle Sousa, Jefferson Martins, Kenji Nose-Filho and Aline Neves are in the Center of Engineering, Modeling, and Social Sciences Applied, Federal University of ABC, Santo André - SP, e-mail: {ysabelle.sousa, jefferson.martins, kenji.nose, aline.neves}@ufabc.edu.br. This work was partially supported by CAPES – Finance Code 001 and by FAPESP the São Paulo Research Foundation (FAPESP), grant #2020/09838-0.

Researchers have significantly investigated the feasibility of diagnosing depression and other neurological disorders using biomarkers obtained from examinations such as Functional Magnetic Resonance Imaging (fMRI), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Electroencephalography (EEG) [4], and Magnetoencephalography (MEG) [5].

Among these techniques, EEG is a cost-effective, non-invasive, and portable tool that captures rapid electrical brain activity through scalp electrodes. Although EEG provides high temporal resolution, its spatial resolution remains limited. Despite this limitation, EEG-based support for diagnosis approaches have shown promising results in the assessment of depression [6]. In addition to EEG, the present study also investigates audio signals as a separate diagnostic modality due to their ease of acquisition, non-invasive nature, and minimal equipment requirements, making them a practical and accessible tool for depression assessment.

The primary objective of this research is to investigate and compare the effectiveness of EEG and audio signals in supporting depression diagnosis, focusing on identifying relevant biomarkers extracted from each modality. Using a dataset of EEG and audio recordings from depressive and non-depressive patients, this study applies preprocessing, feature extraction, feature selection, and classification techniques to assess the discriminative potential of each signal independently.

This work is organized as follows. Section II presents the literature review on EEG analysis for MDD. Section III describes the proposed methodology, including the dataset description, preprocessing techniques, feature extraction, feature selection, and classification procedures applied in both EEG and audio signals. Section IV presents the experimental results, like the hyperparameters selected for each modality and the performance evaluation of the proposed models. Finally, Section V concludes the paper and discusses future research directions.

II. RELATED WORK

EEG signal processing has been widely explored for MDD detection, particularly through linear and nonlinear feature extraction followed by classification. Linear methods capture frequency and time-domain properties, while nonlinear approaches reveal signal complexity [7], [8]. Techniques such as Welch's method, Relative Wavelet Energy, and various entropy measures have been employed with high performance using classifiers like SVM and neural networks [9], [10]. Studies frequently highlight the relevance of Alpha, Beta, and Theta

bands, though alpha asymmetry remains in debate due to gender and comorbidity variability [4].

Network-level features—such as Coherence, Functional Connectivity, and Small-Worldness—have also been proposed to characterize depressive brain activity patterns [11]. Although resting-state EEG is predominant, some works analyze Event-Related Potential (ERP) components with mixed results regarding latency and signal complexity [12]. Nonetheless, EEG-based biomarkers show strong potential in clinical applications despite variability in protocols and sample sizes.

In parallel, audio-based approaches have focused on prosodic and spectral descriptors, including pitch, jitter, Mel-Frequency Cepstral Coefficients (MFCCs), and log-Mel features. Studies using datasets like MODMA reported promising classification performance with SVMs and deep learning models [13]. Audio features often capture vocal instability and reduced expressiveness—characteristics of depressive speech—making them suitable for automated screening.

Lexical and linguistic features have also been used to enhance audio-based models [14], while gender-related differences in speech suggest that personalized models may offer performance benefits [15]. Despite challenges related to speech variability and recording conditions, audio remains a non-invasive and accessible modality for MDD assessment.

Overall, both EEG and audio literature present robust methodologies and high classification performance for MDD detection. However, most approaches remain limited by heterogeneous protocols, small sample sizes, and lack of standardization. These gaps reinforce the importance of systematically comparing different signal modalities, such as EEG and audio, to identify which offer greater discriminative power and reliability in supporting clinical depression diagnosis.

III. METHODOLOGY

This section presents the methodological pipeline adopted in this study for the classification of Major Depressive Disorder (MDD) using EEG and audio signals. It describes the preprocessing procedures applied to each modality, the strategies employed for feature extraction and selection, and the construction of classification pipelines for EEG and audio separately. Finally, the evaluation pipeline used to assess model performance are detailed, allowing a direct comparison of the discriminative capabilities of both signal types.

A. Dataset Description

The Multi-modal Open Dataset for Mental-disorder Analysis (MODMA) [16] is a publicly available dataset from Lanzhou University, China, focused on Major Depressive Disorder (MDD). Approved by the Ethics Committee of the Second Affiliated Hospital of Lanzhou University, it includes resting-state EEG, Event-related Potential (ERP), and audio data. This study primarily utilizes resting-state EEG (128 channels) and audio recordings.

The EEG dataset comprises 53 participants from Lanzhou University Second Hospital (24 diagnosed with depression: 13 males, 11 females; and 29 healthy controls: 20 males, 9 females), aged 16 to 56. The audio dataset includes 52

participants (23 with depression: 16 males, 7 females; and 29 healthy controls), aged 18 to 55. A total of 38 subjects (17 depressive, 21 controls) are common to both modalities.

All MDD participants underwent the Mini-International Neuropsychiatric Interview (MINI) [17], based on DSM-IV [18]. Inclusion required a prior diagnosis of depression. Exclusion criteria included other mental or physical disorders, severe suicidal tendencies, recent substance abuse, and—for women—pregnancy, lactation, or contraceptive use. Healthy controls were excluded if they or close relatives had a history of mental illness.

EEG signals were recorded during a 5-minute resting-state session, with participants instructed to keep their eyes closed and remain motionless. Following MODMA's baseline protocol, a subset of 16 electrodes (Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6) was adopted for analysis, as commonly applied in depression studies [19].

In the audio experiment, three speaking tasks were performed:

- **Interview:** questions from DSM-IV [18] and HAMD;
- **Short Story:** reading affective terms from the ontology corpus and sentimental words extremum table [20];
- **Picture and TAT:** description of images from CFAPS and the Thematic Apperception Test (TAT) [21].

Emotional valence (positive, neutral, negative) was assessed. Each session lasted 25 minutes in a controlled acoustic environment with ambient noise maintained below 60 dB to avoid interference.

B. Pre-processing

The pre-processing stage aimed to enhance the quality of EEG and audio signals by removing artifacts and standardizing the data for further analysis. For EEG, all steps were conducted using the EEGLAB toolbox in MATLAB [22]. Channel location correction was first applied to standardize electrode positions from the 10-10 to the 10-20 international system, ensuring consistency across participants. From the original 128 electrodes, a subset of 17 channels commonly associated with depression research—Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6—was selected for computational efficiency [19], with Cz used as the reference channel. Baseline correction was performed using the `pop_rmbase()` function to remove slow drifts by centering the signal around zero. Next, a bandpass filter between 1Hz and 40Hz was applied, eliminating low-frequency components and high-frequency noise, including 50 Hz powerline interference, as the data were recorded in China [16].

Three patients were excluded due to low-quality EEG recordings. To reduce physiological and non-physiological artifacts, Independent Component Analysis (ICA) was performed using the InfoMax algorithm [23], implemented through the EEGLAB `runica()` function. Artifact-related components, including eye blinks and muscle activity, were identified by visual inspection and manually removed. Finally, each 5-minute EEG recording was segmented into five non-overlapping 1-minute windows, increasing the dataset from 50 to 150 samples.

A similar pre-processing strategy was applied to the audio signals using Python's `librosa` and `numpy` libraries. First, a pre-emphasis filter with $\beta=0.97$ was applied to enhance high-frequency components typically diminished during recording [24]. Then, peak normalization was performed to scale all signal amplitudes to the range of -1 to 1, ensuring consistency across samples. To facilitate time-frequency analysis, a windowing procedure was applied using a Hamming window function [24] with 1-minute segments and overlapping transitions. Fourier transforms were computed for each segment to convert signals into the frequency domain while minimizing edge discontinuities. These pre-processing steps ensured that both EEG and audio signals were properly cleaned and formatted for robust feature extraction and classification in subsequent stages.

C. Feature Extraction

After signal pre-processing, features were extracted from EEG and audio signals to serve as input for classification models.

For EEG signals, linear and nonlinear features were extracted to capture statistical, spectral, and complexity-based characteristics. Linear features included Variance, Root Mean Square (RMS), Peak-to-Peak amplitude, and Hjorth parameters (Mobility and Complexity), along with Power Spectral Density (PSD) metrics—maximum, peak power, and spectral integral—calculated using Welch's method with a Hanning window [25], [26]. Nonlinear features included Singular Value Decomposition Entropy (SVDen), Spectral Entropy, and Rényi Entropy of order 3, capturing the irregularity and complexity of non-stationary EEG signals [27], [19]. After extraction, all features were standardized using z-score normalization [9] to ensure a consistent scale throughout the dataset.

For audio signals, the INTERSPEECH 2010 (IS10) feature set was extracted using the openSMILE toolkit [28]. This set includes 1,582 low-level descriptors (LLDs) including MFCCs, Loudness, Jitter, Shimmer, Band Power, Voicing Probability, and Fundamental Frequency, along with statistical functionals such as percentiles, skewness, and quartile ranges [29]. Redundant or constant valued features were removed and additional attributes, such as onsets and segment duration, were added to improve the audio representation for classification.

D. Feature Selection and Classification

In high-dimensional datasets such as EEG and audio signals, redundant or irrelevant features can degrade model performance and generalization. Feature selection is thus essential to reduce dimensionality, improve accuracy, and enhance interpretability. In this work, the Maximum Relevance - Minimum Redundancy (mRMR) method [30] was employed. This supervised technique selects a subset of features that are highly relevant to the target variable while minimizing redundancy among themselves. Based on mutual information, mRMR evaluates each feature P_i through the following objective function:

$$f^{mRMR}(P_i) = I(T, P_i) - \frac{1}{|S|} \sum_{P_s \in S} I(P_s, P_i) \quad (1)$$

where T is the target class, S is the set of already selected features, and $I(\cdot, \cdot)$ is the mutual information function. The first term maximizes relevance with the target, while the second penalizes redundancy with features already selected. Mutual information is calculated as:

$$I(T, P_i) = \sum_{t \in \Omega T} \sum_{f \in \Omega P} p(f, t) \log \left(\frac{p(f, t)}{p(f)p(t)} \right) \quad (2)$$

where ΩT and ΩP represent the sample spaces of the target and feature, respectively, and $p(\cdot)$ denotes the joint and marginal probability distributions. The algorithm iteratively selects the feature that maximizes $f^{mRMR}(P_i)$ until the desired number is reached.

For classification, the Support Vector Machine (SVM) [31] was adopted due to its wide use and robust performance in biosignal applications [9], [32]. SVM aims to find an optimal hyperplane separating two classes with the maximum margin, defined as:

$$f(x) = \alpha \cdot x + b \quad (3)$$

where α is the weight vector and b the bias. For nonlinear separability, kernel functions such as the Radial Basis Function (RBF) can be applied:

$$K(x, x') = \exp \left(-\frac{\|x - x'\|^2}{2\sigma^2} \right) \quad (4)$$

Here, σ defines the kernel's spread. Given its generalization ability, effectiveness with high-dimensional data, and resistance to overfitting, SVM was selected as the main classifier in this study.

E. Model Evaluation Pipeline

The evaluation process, illustrated in Figure 1, follows a systematic approach involving Monte Carlo simulations, Leave-One-Out Cross-Validation (LOOCV), and hyperparameter optimization to ensure robust model performance for both EEG and audio modalities. Starting with the MODMA dataset, 100 iterations of Monte Carlo simulations with a 1:100 train-test split were executed, where 90% of the data was allocated for training and the remaining 10% for testing. Within each training subset, hyperparameter optimization for the SVM classifier was performed using LOOCV to select the best configuration. This iterative process enabled the identification of optimal features and hyperparameters independently for EEG and audio signals, allowing a direct comparison of their classification performance.

IV. RESULTS AND DISCUSSION

This section presents the results obtained for MDD classification using EEG and audio signals based on the MODMA dataset.

All classification models were trained using Support Vector Machines (SVM) with rigorous hyperparameter tuning. For each modality, 100 iterations of Monte Carlo simulations were conducted with a stratified 90/10 train-test split—45 subjects used for training and 5 for testing in each iteration. Within each training subset, optimal SVM parameters (C , γ) were selected using Leave-One-Out Cross-Validation (LOOCV). The best

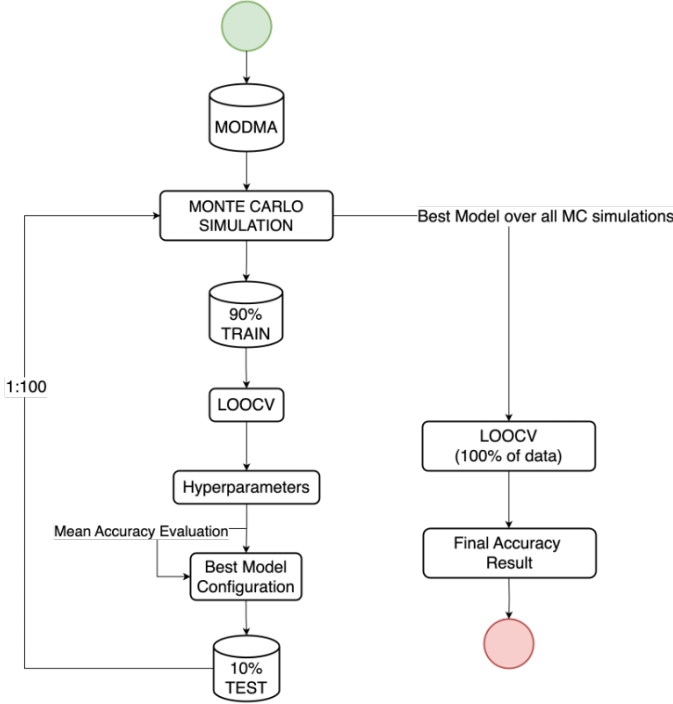


Fig. 1. Evaluation pipeline for EEG and audio modalities. The best model is selected from 100 Monte Carlo simulations using LOOCV and validated on the complete dataset.

model was chosen within all the MC simulations and a final LOOCV was performed with the entire dataset to find the final accuracy. Implementation was carried out using *Scikit-learn*, *Pandas*, and *NumPy* Python libraries.

The results are structured into two main sections: EEG and audio. Each includes classification accuracy, selected feature characteristics, and methodological insights. Tables summarize simulation results across signal types and configurations, with columns indicating:

- **Method:** The feature selection method: mRMR or None;
- **Signal:** The characteristic of the signal processed in the simulation. It can be "All samples" or "windowed samples";
- **Activity:** Activity type of audio processing: Interview, Short Story, Picture or TAT. Only applied to audio results;
- **Best $|S|$:** refers to the number of features selected by mRMR;
- **Best C :** the best regularization parameter value of the SVM classifier;
- **Best γ :** the best polynomial kernel coefficient of the SVM classifier;
- **Final Accuracy:** the accuracy percentage obtained as a mean of 100 MC simulations.

It is important to highlight that, in this study, we performed simulations across EEG and audio modalities, using the complete datasets including all patients who participated in the respective experiments. This comprehensive approach allowed us to assess the overall performance of various feature selection methods on the entire experiments population.

A. EEG

The analysis of EEG signals showed the efficacy of feature selection techniques in improving classification accuracy. The highest classification accuracy of 76% was achieved using mRMR with only 10 selected features ($\approx 2.2\%$ of the total 448 features). This performance was obtained using optimal hyperparameters ($C = 0.0001$, $\gamma = 10$) with a polynomial kernel. In contrast, when using all features without selection, the accuracy reached only 46%, emphasizing the value of mRMR.

When using 1-minute windowed EEG segments, accuracy decreased to 68.60%, suggesting that segmentation did not improve model performance. Table I summarizes the results for all EEG configurations. The results were unable to exceed the benchmark of 82,31% obtained in [19].

TABLE I
EEG INDIVIDUAL ACCURACY RESULTS WITH mRMR

Method	Signal	Best $ S $	Best C	Best γ	Final Accuracy
mRMR	All samples	10	0.0001	10	76.00%
None	All samples	10	1	1	46.00%
mRMR	Windowed samples	10	10	5	68.60%

The features selected most frequently by mRMR included beta-band metrics such as band power and peak power in the frontal (Fp1, Fp2), occipital (O1), and temporal (T5) regions, reflecting their relevance to emotional and cognitive processing. The inclusion of entropy-based features also highlighted the importance of neural complexity in depression classification. These findings are consistent with prior literature on EEG biomarkers in mood disorders [7], [33].

B. Audio

Audio classification using mRMR achieved a maximum accuracy of 80.90% in the "Short Story" task with 30 selected features, surpassing the previous benchmark obtained in [34], of 76.8%. We were able to increase performance in 4.1%, a very interesting result. The "Picture" task followed with 80.86%, while "TAT" reached 72.09%. The "Interview" task yielded the lowest accuracy with 62.27%.

Without feature selection, performance dropped notably to 55.00% accuracy when considering "All Activities", while, with mRMR, a result of 61.68% was achieved. Full results are shown in Table II. Here, all results were obtained applying mRMR for feature selection.

TABLE II
AUDIO INDIVIDUAL ACCURACY RESULTS BY FEATURE SELECTION METHODS

Activity	Best $ S $	Best C	Best γ	Final Accuracy
Interview	50	0.0001	5	62.27%
TAT	50	0.0001	1	72.09%
Picture	60	0.0001	5	80.86%
Short Story	30	0.0001	5	80.90%
All Activities	30	0.0001	1	61.68%

In the "Short Story" task, the most relevant audio features included MFCCs, log Mel-frequency bands, and prosodic

markers such as jitter, shimmer, and voicing-related descriptors. These features capture vocal instability and reduced expressiveness—hallmarks of depressive speech. Statistical functionals such as percentiles and skewness further enhanced discrimination by summarizing variability patterns across speech.

V. CONCLUSIONS

This study evaluated the effectiveness of EEG and audio signals in the classification of Major Depressive Disorder (MDD) using Support Vector Machines and the MODMA dataset. Two unimodal approaches were considered: EEG-only and audio-only. The EEG-based model achieved a maximum accuracy of 76% using the mRMR feature selection technique and a polynomial kernel, highlighting the relevance of frontal and occipital brain regions in depression detection. The audio-based approach reached an accuracy of 80.90% during the “Short Story” task, with selected features including MFCCs, log Mel-frequency bands, and prosodic descriptors such as jitter and shimmer. Even though the EEG result was not able to attain the literature benchmark, the audio result surpassed it in 4.1%. These findings suggest that both neurophysiological and vocal biomarkers are informative for depression classification, with audio-based features showing slightly greater discriminative performance under the conditions evaluated. Future work may explore a multimodal strategy that merges both types of signals, to further enhance diagnostic accuracy and generalization.

REFERENCES

- [1] World Health Organization, “Depressive disorder (depression),” <https://www.who.int/news-room/fact-sheets/detail/depression>, 2023.
- [2] J. C. Tolentino and S. L. Schmidt, “DSM-5 Criteria and Depression Severity: Implications for Clinical Practice,” *Frontiers in Psychiatry*, vol. 9, p. 534, October 2018.
- [3] F. Jiang, Y. Jiang, H. Zhi *et al.*, “Artificial intelligence in healthcare: past, present and future,” *Stroke and Vascular Neurology*, vol. 2, no. 4, pp. 230–243, June 2017.
- [4] G. E. Bruder, J. W. Stewart, and P. J. McGrath, “Right brain, left brain in depressive disorders: Clinical and theoretical implications of behavioral, electrophysiological and neuroimaging findings,” *Neuroscience Biobehavioral Reviews*, vol. 78, pp. 178–191, 2017.
- [5] Q. Lu, K. Bi, C. Liu, G. Luo, H. Tang, and Z. Yao, “Predicting depression based on dynamic regional connectivity: a windowed granger causality analysis of meg recordings,” *Brain research*, vol. 1535, pp. 52–60, 2013.
- [6] F. S. De Aguiar Neto and J. L. G. Rosa, “Depression biomarkers using non-invasive EEG: A review,” *Neuroscience Biobehavioral Reviews*, vol. 105, pp. 83–93, 2019.
- [7] S. Mahato and S. Paul, “Electroencephalogram (EEG) signal analysis for diagnosis of major depressive disorder (MDD): A review,” in *Nanoelectronics, Circuits and Communication Systems*. Singapore: Springer Singapore, 2018, pp. 323–335.
- [8] A. Dev, N. Roy, M. K. Islam, C. Biswas, H. U. Ahmed, M. A. Amin, F. Sarker, R. Vaidyanathan, and K. A. Mamun, “Exploration of EEG-based depression biomarkers identification techniques and their applications: A systematic review,” *IEEE Access*, vol. 10, pp. 16 756–16 781, 2022.
- [9] W. Mumtaz, L. Xia, S. S. A. Ali, M. A. M. Yasin, M. Hussain, and A. S. Malik, “Electroencephalogram (EEG)-based computer-aided technique to diagnose major depressive disorder (MDD),” *Biomedical Signal Processing and Control*, vol. 31, pp. 108–115, 2017.
- [10] B. Hosseinifard, M. H. Moradi, and R. Rostami, “Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal,” *Computer Methods and Programs in Biomedicine*, vol. 109, pp. 339–345, 2013.
- [11] L. Orgo, M. Bachmann, K. Kalev, M. Järvelaid, J. Raik, and H. Hinrikus, “Resting EEG functional connectivity and graph theoretical measures for discrimination of depression,” in *IEEE EMBS International Conference on Biomedical and Health Informatics*, 2017, pp. 389–392.
- [12] C. Greco, O. Matarazzo, G. Cordasco, A. Vinciarelli, Z. Callejas, and A. Esposito, “Discriminative power of EEG-Based biomarkers in major depressive disorder: A systematic review,” *IEEE Access*, vol. 9, pp. 112 850–112 870, 2021.
- [13] M. Gheorghe, S. Mihalache, and D. Burileanu, “Using deep neural networks for detecting depression from speech,” in *2023 31st European Signal Processing Conference (EUSIPCO)*, 2023, pp. 411–415.
- [14] F. Menne, F. Dörr, J. Schröder *et al.*, “The voice of depression: speech features as biomarkers for major depressive disorder,” *BMC Psychiatry*, vol. 24, no. 794, 2024.
- [15] S. Lee, S. W. Suh, T. Kim, K. Kim, K. H. Lee, J. R. Lee, G. Han, J. W. Hong, J. W. Han, K. Lee *et al.*, “Screening major depressive disorder using vocal acoustic features in the elderly by sex,” *Journal of Affective Disorders*, vol. 291, pp. 15–23, 2021.
- [16] H. Cai, Z. Yuan, Y. Gao, S. Sun, N. Li, F. Tian, H. Xiao, J. Li, Z. Yang, X. Li *et al.*, “A multi-modal open dataset for mental-disorder analysis,” *Scientific Data*, vol. 9, 2022.
- [17] D. V. Sheehan, Y. Lecrubier, K. H. Sheehan, P. Amorim, J. Janavs, E. Weiller, T. Hergueta, R. Baker, G. C. Dunbar *et al.*, “The mini-international neuropsychiatric interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10,” *The Journal of Clinical Psychiatry*, pp. 22–33, 1998.
- [18] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed. Washington, DC: Amer Psychiatric Pub Inc, 1994.
- [19] S. Sun, J. Li, H. Chen, T. Gong, X. Li, and B. Hu, “A study of resting-state EEG biomarkers for depression recognition,” 2020, <https://arxiv.org/pdf/2002.11039>.
- [20] C. Quan and F. Ren, “Construction of a blog emotion corpus for chinese emotional expression analysis,” in *Proceedings of the 2009 conference on empirical methods in natural language processing*, Singapore, 2009.
- [21] L. Bellak, “The thematic apperception test in clinical use,” in *Projective Psychology: Clinical Approaches To The Total Personality*, L. E. Abt and L. Bellak, Eds. Alfred A. Knopf, 1950, pp. 185–229.
- [22] The Mathworks, Inc, *Statistics and Machine Learning Toolbox Documentation*, <https://www.mathworks.com/help/stats/index.html>, February 2023, available at <https://www.mathworks.com/help/stats/index.html>.
- [23] T.-P. Jung, S. Makeig, A. J. Bell, and T. J. Sejnowski, “Independent component analysis of electroencephalographic and event-related potential data,” in *Central Auditory Processing and Neural Modeling*. Boston, MA: Springer US, 1998, pp. 145–151.
- [24] T. Bäckström, *Speech Coding: with Code-Excited Linear Prediction*. Springer, 2017.
- [25] B. Hjorth, “EEG analysis based on time domain properties,” *Electroencephalography and Clinical Neurophysiology*, vol. 29, pp. 306–310, 1970.
- [26] P. Welch, “The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms,” *IEEE Transactions on Audio and Electroacoustics*, vol. 15, pp. 70–73, 1967.
- [27] Z. Liang, Y. Wang, X. Sun, D. Li, L. J. Voss, J. W. Sleight, S. Hagihira, and X. Li, “EEG entropy measures in anesthesia,” *Frontiers in Computational Neuroscience*, vol. 9, 2015.
- [28] F. Eyben, M. Wöllmer, and B. Schuller, “OpenSMILE: the munich versatile and fast open-source audio feature extractor.” New York, NY, USA: Association for Computing Machinery, 2010. [Online]. Available: <https://doi.org/10.1145/1873951.1874246>
- [29] F. Eyben, *Real-time Speech and Music Classification by Large Audio Feature Space Extraction*. Cham, Switzerland: Springer, 2016.
- [30] H. Peng, F. Long, and C. Ding, “Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 1226–1238, 2005.
- [31] V. Vapnik, *The Nature of Statistical Learning Theory*. Springer Science & Business Media, 2013.
- [32] Y. Mohammadi, M. Hajian, and M. H. Moradi, “Discrimination of depression levels using machine learning methods on EEG signals,” in *IEEE 27th Iranian Conference on Electrical Engineering (ICEE)*, 2019, pp. 1765–1769.
- [33] M. Čukić, M. Stokić, S. Simić, and D. Pokrajac, “The successful discrimination of depression from EEG could be attributed to proper feature extraction and not to a particular classification method,” *Cognitive Neurodynamics*, vol. 14, pp. 443–455, 2020.
- [34] X. Chen and Z. Pan, “A convenient and low-cost model of depression screening and early warning based on voice data using for public mental health,” *International Journal of Environmental Research and Public Health*, vol. 18, no. 12, 2021.