Children ADHD Disease Detection using Pose Estimation Technique

[1]Faria anjum [2]Dr. E. Seshatheri [3]Dr.B.Sasi Kumar
[1]MTech Student in CSE Department in Dr. V.R.K. Women's College of Engineering and Technology
[2]Professor Department of CSE in Dr. V.R.K. Women's College of Engineering and Technology
[3]Principal and Professor Department of CSE in Dr. V.R.K. Women's College of Engineering and Technology

Abstract—There are a multitude of mental health conditions that can affect individuals, with various explanations accounting for their occurrence. This paper explores the current machine learningbased methods used to identify Attention Deficit Hyperactivity Disorder (ADHD) and depression in humans.

Prevalence of mental ADHD and depression is increasing worldwide, partly due to the devastating impact of the COVID-19 pandemic for the latter but also because of the increasing demand placed on the mental health services. It is known that depression is the most common mental health condition, affecting an estimated 19.7% of people aged over16. ADHD is also a very prevalent mental health condition, affecting approximately 7.2% of all age groups, with this being conceived as a conservative estimate. We explore the use of machine learning to identify ADHD and depression using different wearable and non-wearable sensors/ modalities for training and testing. These modalities include functional Magnetic Resonance Imagery (MRI), Electroencephalography (EEG),Medical Notes, Video and Speech. With mental health awareness on the rise, it is necessary to survey the existing literature on ADHD and depression for a machine learning based reliable Artificial Intelligence (Al). With access to in-person clinics limited and a paradigm shift to remote consultations, there is a need for Al-based technology to support the healthcare bodies, particularly in developed countries.

Index Terms—Artificial Intelligence, ADHD, attention deficit hyperactivity disorder, machine learning, metal health

I. INTRODUCTION

There are a multitude of mental health conditions that can affect individuals, with various explanations accounting their occurrence. There is no single definitive answer that has been identified. Conditions like depression and schizophrenia have been associated with hereditary factors and chemical imbalances in the human body [L]. However, this research mainly focuses on ADHD and depression, the two most prevalent mental disorders in humans. Both conditions often co-occur, with people diagnosed with one being more likely to be diagnosed with the other. In fact, adults with ADHD are three times more likely to have depression, and individ-uals with depression have a 30-40% prevalence of ADHD. There are also links between ADHD and increased suicidal ideation.

Distinguishing between the two can be challenging,due to overlapping symptoms and the potential side effects of ADHD medications. Saying this, differences exist in mood, motivation and sleep patterns between the two conditions [21, [3]. Both ADHD and depression are very broad topics, so to specialise our paper we focus only on wearable/non-wearable sensing and machine learning. Due to the link in symptoms, if one machine learning model can accurately detect one of the disorders, there is a chance that the model can be generalised in identifying the other. These connections and the high prevalence rate is what motivated this paper.ADHD among children aged 18 and under to be 7.2% with a 95% confidence level [4]. Notably, cases of persistent ADHD, where symptoms that begin in childhood continue into adult- hood, have a lower prevalence of 2.58% [5]. This discrepancy is believed to stem from limited access to diagnosis during youth.

II. LITERATURE SELECTION CRITERIA

Before the paper compilation, research questions were pro-posed to allow for concise conclusions and efficient searches.

A. Research questions(RQs)

The following research questions were finalized to focus the scope of the survey:

• What wearable and non-wearable sensing have been used in datasets for mental health ML-based ADHD and depression detection research?

• What are the advantages and disadvantages with individual modalities when trying to diagnose ADHD or depression using machine learning?

• What is the most popular classification algorithm applied?

• What is the standard of the classification method used?

B. Search strategy

Compiling the papers was achieved through a keyword string query over several literature databases. The keywords in the string were chosen to produce results that fit the RQs. The keyword string query is as follows: ("Classifica- tion" OR "Neural Networks" OR "Machine Learning" OR "Deep Learning" OR
"Supervised Learning" OR "Unsu- pervised Learning") AND
("Depression" OR "ADHD") AND ("Diagnosis"). The following query was used on the following literature databases:

TEEE Xplore, Science Direct, PubMed and Web of Science.

C. Criteria for classification of studies

The inclusion criteria:

• Publication in English.

• Inclusion of data containing an individual with a formal diagnosis of depression or ADHD.

• Articles involving the diagnosis of a mental health con- dition by using Machine Learning.

• Investigating the diagnosis of ADHD or depression in humans using Machine Learning.

• Publication in a peer-reviewed journal.

• Publication within the last 11 years (2011-2022).

• Publication in a non-peer-reviewed journal.

• Publication in conference proceedings,

III. TESTING FOR A MENTAL HEALTH CONDITION

The DSM-V includes descriptions, symptoms, and other relevant criteria for specific mental health disorders to aid in diagnosis. Moreover, it provides diagnostic criteria for both children and adults. As a result, the majority of studies referenced in this paper employ the DSM-V to accurately identify individuals with ADHD or depression.

The International Classification of Diseases (ICD-11) was created by the World Health Organisation at a similar time to the DSM-V [33]. Similarly, it provides a broad range of knowledge on the extent, causes and consequences of human diseases (both medical and mental). The ICD-11 allows for systematic recording, interpretation and therefore analysis of mortality and morbidity data that is collected globally.

As both the DSM-V and ICD-11 are very similar in nature, there is a push to harmonise both together. To make this happen, in new iterations, the main focus will be to have the greatest clinical impact. Achieving this means increasing their international uniformity, with the enhancement of cul- tural compatibility being the primary goal.

ADHD

Regarding ADHD specifically, the DSM-V asserts that to be diagnosed with Attention Deficit Disorder (ADD), an individual must exhibit five or more symptoms of inattention persisting for over six months. Additionally, to be diagnosed with ADHD, five or more symptoms of both hyperactivity and impulsivity, along with inattention symptoms, must be present for more than six months.

Symptoms are classified into three major components:

Inattention, Hyperactivity and Impulsivity [13], [34]:

• Inattention:

1) Forgetfulness in daily tasks/work.
2) Making careless mistakes in work or tasks.
3) Difficulty sustaining attention in tasks.
4) Fails to complete tasks.
5) Doesn't listen when spoken directly to.
6) Reluctance in joining tasks that require sustained attention.
7) Often loses things necessary for tasks.
8) Easily distracted by external stimuli.
9) Often forgetful in daily tasks.

• Hyperactivity:

1) An individual constantly moving around, even during inappropriate times such as in a cinema

• Impulsivity:

1) Interrupting conversations or answering before the question has been asked in full.

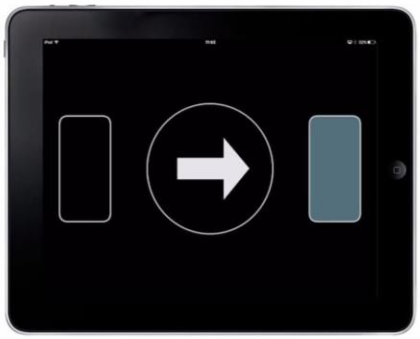

FIGURE 1. An example of the SST. It can be done on laptop

2) Making a decision in the short term without considering the effects of the long term.

3) Difficulty with self control.

1) Questionnaires

Numerous questionnaires, including Conners-3 [14], cal id in the diagnosis of ADHD. These questionnaires cai be completed by clinicians, patients, primary caregivers, or sec- ondary caregivers. When completed by someone other than the clinician, the questionnaire offers valuable insight into the individual's behavioural history.

However, the subjective nature of the responses may lead to inaccuracies and false positives in individuals pursuing a diagnosis, even if they do not genuinely have ADHD.

2) Stop signal tasks (SST)

Fig. 1 illustrates the screen a participant would encounter when taking the SST. This test represents a unique version of a classic method for measuring response inhibition (i.e., impulse control). Participants respond to an arrow stimulus by selecting one of two options based on the arrow's direction. The test comprises two parts:

• First, the participant is introduced to the test and instructed to press the left-hand button when they see a left-pointing arrow and the right-hand button when they see a right-pointing arrow. The participant practices this task in 16 trials.

• Next, the participant is asked to continue selecting but- tons corresponding to the arrow directions.

However, if an auditory signal (such as a beep) occurs, they should refrain from responding and not press the button.

3) Continues performance test(CPT)

The CPT is a task-oriented, computerized assessment, that evaluates attention-related issues in individuals aged 8 years and older. It measures the participant's performance in areas such as attentiveness, impulsivity, sustained attention, and vigilance. The CPT supplements[14], offering insights into an individual's performance in attention tasks.

Depression

The DSM-V presents depression as persistent feelings of sadness and hopelessness while showing lack of interest in activities that were once enjoyed [13]. It mentions that individuals could experience additional physical symptoms sucl as chronic pain or digestive issues. The DSM-V states that th subject must be experiencing five or more of the following symptoms during the same 2 week period:

1) Depressed mood most of the day, experienced nearly every day.

2) Noticeable diminished interest or pleasure in all (or almost all) activities.

3) Experiencing significant weight loss or gain with a decrease or increase in appetite.

4) A reduction of physical movement and thoughts slowing down.

5) Fatigue or loss of energy nearly every day.

6) Feelings of worthlessness nearly every day.

7) Diminished ability to think or concentrate nearly every day.

8) Recurrent thoughts of death.

At least one of the symptoms should be either a depressed mood or loss of interest or please. It should be noted that to receive a diagnosis of depression, the symptoms must cause the subject clinically significant distress and impairment to everyday life.

1) Patients health questionnaires -9 (PHQ-9)

The PHQ-9 is a self-administered diagnostic tool used for criteria-based diagnosis of depression, as established by Kroenke, Spitzer and Williams [35]. The initial study involved 6,000 patients from various clinics. Criterion valid- ity, which is predictive of outcomes, and construct validity, which assesses how well a test measures its intended subject, were determined against a mental health professional-led interview and the 20-item Short-Form General Health Survey respectively [36].

The PHQ-9, comprising of only nine questions, is based on the actual nine criteria for DSM-V depressive disorders diagnosis and can also indicate depressive symptom severity.

Subjects respond based on their feelings and thoughts over the past 2 weeks as presented in Table 1. Clinicians interpret answers and scores to determine the presence and severity of depression (Table 2). Four or more ticks in the bold area suggest a depressive disorder.

2) Beck depression inventory (BDI-II)

The BDI-II is in its second iteration and is one of the most widely used instruments for detecting depression [37].rating scales like Conners-3 [14], offering insights into an individual's performance in attention tasks.

Depression : The DSM-V presents depression as persistent feelings of sadness and hopelessness while showing lack of interest in activities that were once enjoyed [13]. It mentions that individuals could experience additional physical symptoms such as chronic pain or digestive issues. The DSM-V states that the subject must be experiencing five or more of the following symptoms during the same 2 week period:

9) Depressed mood most of the day, experienced nearly every day.

10) Noticeable diminished interest or pleasure in all (or almost all) activities.

11) Experiencing significant weight loss or gain with a decrease or increase in appetite.

12) A reduction of physical movement and thoughts slowing down.

13) Fatigue or loss of energy nearly every day.

14) Feelings of worthlessness nearly every day.

15) Diminished ability to think or concentrate nearly every day.

16) Recurrent thoughts of death.

At least one of the symptoms should be either a depressed mood or loss of interest or please. It should be noted that to receive a diagnosis of depression, the symptoms must cause the subject clinically significant distress and impairment to everyday life.

FIGURE 2. A typical MRI Scanner [108]. MRI scanners are expensive pieces of equipment so they are constantly in use at hospitals for multiple needs. The bed moves in and out of the main scanner depending on what area of the human body is being e either a depressed should be noted that symptoms must cause and impairment to
E-9 (PHQ-9)
caustic tool used for as established by initial study involved ion valid- ity, which uct validity, which ended subject, were sional-led interview al Health

Survey
questions, is based on depressive disorders e symptom severity. nd thoughts over the Clinicians interpret ence and severity of ks in the bold area BDI-II
is one of the most depressing [37]. It is being a self-report asure the severity of estions where each respective question has a list of four statements that are arranged in increasing severity. Each question is focused on a particular symptom of depression. Its second revision aligned its questions with the DSM-IV criteria by having the answers focused on the last 2 weeks upon taking the test.

F. MAGNETIC RESONANCE IMAGING (MRI)

Magnetic Resonance Imaging (MRI) is a type of scan that uses powerful magnetic fields and radio waves to provide highly detailed images of the inside of the body. A scan can last between fifteen and ninety minutes, depending on the size of the area being scanned. The main advantage of MRI scan- ners is that they are harmless to the subject. A main downside to them is that they can be claustrophobic. There are several types of MRI scanner measurements with the focus of this survey being functional MRI (f-MRI) and resting-state MRI (rs-MRI). Initially developed to showcase regional/localized, time-varying changes in brain metabolism, f-MRI scanners have gained popularity due to their versatility with invasive and non-invasive techniques, good spatial resolution, and rel- atively low cost [38]. In the context of the studies mentioned throughout this paper, f-MRI's primary use is to observe increased neural activity by having a subject perform a task while in an MRI scanner, as shown in Fig. 2.

corrupted by background noise, depending on the recording environment. Therefore, it is recommended that a dual microphone con- figuration is implemented. Ideally, a microphone, such as a lavalier microphone, is attached to the participant being recorded. While a secondary microphone, or microphone array is placed in the room to record the environment noises. To achieve the best quality audio recording, it is recom- mended to record in the highest sampling rate that the chosen microphone has to of Pair the highest sampling rate with a 24-bit rate to increase the quality of the recordings while increasing the level of detail. When designing video data recording, it is vital to consider the stability of the cameras. Correct tripods and mounts are essential as you n't want any additional blurriness or motion capture pending on what is being captured, the resolution the camera and frame rate can differ due to there being a trade-off between resolution and performance with machine learning algorithms.
Provided the camera is stable, recording at a resolution of 1920 × 1080 (High-Definition) at a frame-per-second (fps) of 30 is suitable. If a budget allows for it, there is also the option to record in 4K (4096 × 2160) at 30 fps.

IV. DATA SET

There are not a lot of publicly available datasets for Mental Health challenges due to the highly sensitive nature of the data. The main concern is the protection of the participant's privacy i.e, identity and health

information. Therefore in some cases, it is safer to not release the data publicly. In the cases where data has been made public, mainly the video modality, it is processed into features that can not be reverse transformed into their original format. This pre-processing of the data can impact the algorithms used to experiment with such data.

It is known that a common problem with datasets involving medical information are usually small in size.

This is due to complications with preserving the participant's identity while also facing challenges in finding enough individuals with the condition being researched. If a researcher has access to the original video, the small dataset size could be increased using data augmentation techniques. The choice is with the researcher but as a few examples, blurring can be applied to the video or frame mirroring could be applied [40]. set consisting of 35 subjects where 12 are depressed and 23 are controls.

Use either SI (MKS) or CGS as primary units. (SI units are strongly encouraged.) English units may be used as secondary units (in parentheses). This applies to papers in data storage. For example, write ―15 Gb/cm^2 (100 Gb/in^2).‖ An exception is when English units are used as identifiers in trade, such as ―3½ in disk drive.‖ Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do

V. MACHINE LEARNING AND COMPUTER VISION IN ADHD DETECTION

A. Non-wearable techniques

1) Imaging

Table 6 shows that numerous studies have analyzed imag- ing data using various techniques. Exploiting an SVM is a popular approach for classification, applied to both imaging and EEG data. This popularity could be attributed to SVM's afor factors such as age, ADHD subtypes, and con- text in future research to ensure generalizability and clinical relevance of these models.

VI. MACHINE LEARNING AND COMPUTER VISION IN DEPRESSION DETECTION

Research by Zhu et al. [64] and Meshram and Rambola [65] use CNNs to extract facial features, either static or dynamic, to estimate the BDI-II depression severity score. The model by Zhu et al. notably improved upon results from the AVEC-2013 dataset, while Meshram and Rambola's model achieved a high classification average of 92.56% on the AVEC-2016 challenge datasetsuggesting the effectiveness of deep learn- ing techniques in detecting depression.

Similar trends can be observed in the work of He et al. [66], who also employed the extraction of dynamic facial features, specifically using the MRLBP-TOP tech- nique. The promising RMSE value of 8.90 achieved by this framework compared to the AVEC-2013 baseline (10.72) demonstrates the potential of advanced feature extraction techniques in this field.

Li et al. [67] took a slightly different approach by focus- ing on eye movement as a predictor of depression. Despite this novel approach, the model still achieved a classification accuracy of 80.1%, indicating that diverse biological signals can potentially be valuable in depression detection.

In the research conducted by Hong et al. [68], the focus was on distinguishing between bipolar and unipolar disor-ders, as well as identifying healthy controls. Their method- ology combined action unit descriptors and motion vectors processed by machine learning techniques, yielding a reasonable classification accuracy of 72.2%.

Zhou et al. [69] introduced the MR-DepressNet, a deep regression network that used visual features to estimate depression severity. This model improved on the AVEC-2013 baseline RMSE value (13.61) by achieving 8.28, demonstrating the potential of visual feature exploitation in improving depression estimation.

The research of Tadalagi and Joshi [70], Shang et al. [71], and Uddin, Joolee and Lee [72] all took innovative approaches by combining various techniques and method- ologies.

Tadalagi and Joshi implemented their model on a real-time system, while Shang et al. introduced a quaternion- based method. Uddin, Joolee and Lee used a two-stream deep spatiotemporal network. All of these methods were evaluated on AVEC datasets and achieved competitive results.

Song et al. [73] focused on the extraction of multi-scale video-level features, providing a novel perspective on depres-sion analysis. Using spectral representations processed by CNNs and Artificial Neural Networks (ANN), their model achieved a competitive MAE/RMSE score on the AVEC2013 test set. Yang et al. [74] introduced a multi-modal framework that exploited video, audio and text data to estimate PHQ-8 scores and infer the mental condition of the subject. The combi- nation of a deep CNN, DNN, Paragraph Vector, SVM, and random forest methods showed the value of integrating dif- ferent types of data in depression detection.Lastly, De Melo, Granger and López [75] tackled the cost-effectiveness of 3D-CNNs by proposing a deep learn- ing architecture that operates without 3D convolutions.Taking deep learning to the forefront, He and Cao [79] used CNNs to extract deep-learned features from raw speech waveforms and spectrograms. This shift towards automated depression analysis tools that could generate complex features signaled a significant step towards more sophisticated models.

Conversely, Jiang et al. [80] proposed an ensemble logis- tic regression model, offering separate models for males and females. This gender-specific approach, combined with diverse feature extraction, added a new dimension to the field, recognizing the potential differences in depression expression across genders.

Meanwhile, Li et al. [81] introduced the Multiscale Audio Data Normalization (MADN) algorithm, marking another significant advancement in feature extraction. Their approach further emphasized the potential for innovative methods to improve upon existing models.

Pushing the boundaries of feature analysis, Muzammel et al. [82]
focused on

the acoustic features of vowel and
consonant spaces. Their method of augmenting data and segmenting speech brought attention to the nuanced elements of speech and their potential role in depression detection.

A notable turning point occurs with Zhao et al. [83], who introduced a comprehensive approach combining unsupervised learning, hierarchical attention, and knowledge transfer.

Their impressive results underscored the potential of complex deep learning architectures in determining depression severity.

In summary, this collection of studies forms an engaging narrative that underscores the evolution of speech analysis in depression detection. From basic lexical analysis to deep learning, and from gender-specific models to nuanced vowel and consonant analysis, it shows a consistent progression towards more complex and sophisticated models. This narrative serves as a testament to the ongoing advancement of machine learning techniques in mental health research, particularly in understanding and addressing depression.

1) IMAGING

Depression detection research has expanded beyond speech analysis to leverage advanced neuroimaging technologies, and machine learning has remained a constant ally. Various studies have illustrated how different feature extraction and classification methods can yield significant results in identifying depression at an individual level.

In a novel approach, Cao et al. [84] made strides by focusing on the individual rather than group dynamics, using probability density functions (PDFs) to target functional con-nectivity. They integrated a t-test for primitive selection, followed by Kernel density estimation for PDFs, resulting in a significant classification accuracy of 84.21% using an SVM classifier. This shift towards an individual-oriented approach marked a significant milestone in depression detection.

Building upon functional connectivity, Guo et al.The effective use of topological metrics as inputs to classifiers, particularly the SVM-RBF, yielded an impres- sive classification accuracy of 83.0%. Guo et al. [86] further refined their approach by constructing an automatic classifier based on a high-order minimum spanning tree functional brain network. Their multi-kernel SVM, after intricate feature extraction and selection, achieved an exceptional accuracy of 97.54%.

In parallel, Li et al. [87] used voxel-based morphometry (VBM) and regional homogeneity (ReHo) analyses to extract key features.

By adopting the LASSO approach to isolate the most informative brain regions, they achieved a validated classification accuracy of 86.4% with an SVM classifier. Their emphasis on feature extraction highlighted the importance of selecting relevant regions in the brain for effective depression detection.

Additionally, Rosa et al. [88] proposed a sparse frame- work for depression classification. Their utilization of sparse inverse covariance models to estimate functional connectiv- ity, coupled with an L1-norm SVM, resulted in an accuracy of 85%. This approach once again demonstrated the importance of functional connectivity in the detection of depression.

Conversely Li et al. [89], who leveraged independent component analysis to define the triple network model. Their integration of effective connectivity features, dynamic functional connectivity features, and rigorous statistical testing led to an accuracy of 90.91% with an SVM classifier.

Simultaneously, Sen et al. [90] focused on dynamic and static connectivity measures, extracted from rs-fMRI data, as a basis for feature extraction. Their use of Pearson's cor- relation and entropy measures resulted in a combination of static and dynamic features, yielding a classification result of 82% with an RBF-SVM.

Finally, Wang et al. [91] distinguished themselves by using
functional near-infrared spectroscopy (NIR) instead of fMRI. By utilizing the unique properties of near-infrared light and the absorptive characteristics of blood, they were able to extract crucial features. With an AlexNet structured network, they achieved an impressive accuracy of 90%.

Together, these studies form a compelling narrative that showcases the interplay between neuroimaging technologies and machine learning in depression detection. From functional connectivity and brain network analysis to innovative uses of light in fNIR, the research direction showcases a continued progression toward individual-level analysis and a growing emphasis on sophisticated feature extraction and selection methods. The consistent use of SVM classifiers across most studies points to their effectiveness in this con- text, further highlighting the importance of machine learning in mental health research.laboratory, and sociodemographic data. The Random Forest (RF) algorithm emerged as the best performer, achieving a robust classification accuracy of 89% and an AUC of 0.87.

This study solidifies the foundational premise that machine learning can effectively discern depression from a combination of diverse data types.

Next, Liu et al. [93] introduced EarlyDetect (ED), a composite screening application utilizing machine learning to incorporate a wide spectrum of variables. From family history of mental illness to suicide ideation, ED exemplifies a comprehensive approach. Using the ElasticNet algorithm, it achieved a balanced accuracy of 72% with an AUC of 0.781.

This underlines the potential for machine learning to be effective in complex,real-world settings, synthesizing multiple factors into predictive models.

Adding to the narrative, Ma et al. [94] strived to create a machine learning framework that could expedite the Affective Disorder Evaluation scale. The resulting Bipolar Diagnosis Checklist in Chinese (BDCC), which used the RF algorithm to rank feature importance, achieved an outstanding classification accuracy of 99.6%. The success of BDCC emphasises the role machine learning can play in simplifying and accelerating mental health evaluations.

Simultaneously, Mato-Abad et al. [95] leveraged Arti- ficial Neural Networks (ANNs) to identify a subtype of mild cognitive impairment (MCI) associated with depression. The ANN's success, with an 86% accuracy, highlights the potential of machine learning in elucidating the nuanced intersections between different mental health conditions.

On a parallel track, Meng et al. [96] introduced a temporal deep learning model performing bi-directional representation learning on Electronic Health Record (EHR) sequences. The model's AUC ranged from 0.73 - 0.85, based on the prediction window timeframe. This exploration of temporal modelling in EHR data showcases machine learning's ability to draw insights from longitudinal health data.

In another initiative, Meng et al. [97] devised a model incorporating temporal Hierarchical Clinical Embeddings with opic Modelling (HCET), addressing data sparsity issues. The improvement in AUCs further emphasized the potential of machine learning in handling complex, sparse datasets.

Adding another dimension, Parker et al. 198] sought to discriminate between bipolar and unipolar subjects. They achieved a promising classification accuracy of 96%, demonstrating machine learning's capacity to distinguish between different mental health disorders, even within the challenging context of unbalanced datasets.

Sharma and Verbeke [99] employed the Extreme Gradi- ent Boosting algorithm on a biomarker dataset, achieving a balanced classification accuracy of 94.42% despite ini- tial dataset biases. This finding underscores machine learn-ing's robustness and adaptability in the face of imbalanced data.

Lastly, Zhou et al. [100] leveraged natural language processing in analysing discharge summaries of depressed patients. Their system, MTERMS, consistently outperformed standard classifiers, reaffirming the strength of machine learning

in interpreting unstructured text data.Together, these studies craft a compelling tale of how machine learning has been applied to diverse data types and challenges in depression detection, consistently achieving impressive results. It illustrates a trend towards increasingly complex and real-world applicable models, with promising indications for the future of machine learning in mental health diagnostics.

B. Wearable techniques

A vast amount of research has been conducted into detect- ing depression using a wearable sensors that produce EEG signals.

Table 7 is summarising the best performing studies with the most popular classification method being CNNs and SVMs.

Following on, different classification methods have been used with EEG signals. Firstly, Cai et al. [101] introduced a multimodal model fusing different EEG data sources, gath- ered under a range of emotional conditions. Feature weight- ing was performed through a genetic algorithm on linear and non-linear features, with this unique approach leading to a robust classification accuracy of 86.98%. This method emphasizes the potential of multimodal models in detecting depression and accentuates the strength of genetic algorithms in feature weighting.

Akbari et al. used k-Nearest Neighbours (k-NN) with geometric features extracted from the EEG signals'

Self- Organising Decision Process (SODP) [102]. The Binary Par- ticle Swarm Optimisation (BPSO) algorithm was utilised for feature selection, culminating in impressive results: 98.79% classification accuracy, 97.72% sensitivity, and 99.86% specificity. This work underscores the value of geometric fea- tures and the effectiveness of the BPSO algorithm in feature selection for depression detection.

Moreover, Li et al. [103] ventured to extract multiple lin- ear and non-linear features from EEG signals. A rigorous comparison of five different feature selection methods was carried out, with significant discriminant features being iden- tified using Bonferroni correction t-tests. The outcome was a commendable average classification accuracy of 95%, high- lighting the importance of meticulous feature selection in achieving high classification accuracy.

Simultaneously, Saeedi et al. applied sample and approx- imate entropy to wavelet packets, with significant features selected using a Genetic Algorithm (GA) [104]. This method achieved a classification accuracy, sensitivity, and specificity of 98.44%, 97.10%, and 100% respectively. The use of GA once again demonstrates its potency in feature selection, enhancing the classification performance Simultaneously.illuminates the potential of fractal dimensions and chaos theory in depression detection.In a parallel effort, Cukic et al. examined Higuchi's Fractal Dimension and Sample Entropy as non-linear mea- sures in discriminating between depressed patients and controls [106]. By leveraging Principal Component Analysis (PCA) for feature dimensionality reduction, they achieved an average classification accuracy of 97.56%. This reinforces the idea that non-linear measures can be highly discrimina- tive and the role of dimensionality reduction techniques in boosting classification

Overall, these studies show the exploration into the detec- tion of depression using EEG signals. They illustrate the evolution of methodologies, from the use of different feature extraction techniques to the application of various machine learning algorithms. The consistently high classification accuracies across studies reinforce the potential of these approaches in advancing depression detection.

VII. CONCLUSION AND FUTURE WORK

This survey has gone into detail about machine learning applications in mental health detection. It can be observed that the most popular methods for automatic detection of depression and ADHD is by exploiting imaging data and EEG data. The non-intrusive nature of the EEG provides an argument that it is the preferred choice. This is due to the vast amount of methods that can be applied to analysis, while causing no harm to the subject.

The biggest drawback about research involving mental health conditions is the size of the dataset. Due to the nature of the conditions, for both ADHD and depression it is dif- ficult to get enough subjects to participate in the research. Furthermore there are possible implications with protecting the privacy of all subjects due to it being very sensitive data. When subjects have agreed to have their data used, there is also the issue of whether the data can be publicly shared or whether it remains private. Lastly, with regards to ADHD and depression, the spectrum of behaviour is vast, meaning some behaviour is very rigid or too excitatory.

Therefore, training a classifier to detect these behaviours can be even harder as there is not enough data to cover such a vast

Following on, there is more research being conducted into depression.

This could be due to the awareness of the mental health condition being bigger or because of the available datasets.

We suggest that for both ADHD and Depression respectively, there is a collective movement for a joint database containing multimodal data for the respective mental health conditions.

Within these databases, there would be an established method for protecting the participants pri- vacy such as converting their identity to a number/letter and processing the video/image data using techniques such as the Histogram of Gradients. The file types would be made consistent so that all users would know what to expect and baseline scores would be achieved to provide state-of-the-art comparisons. Lastly, for use in research, an End user license (EULA) would have to be signed to protect the organisers and subjects' data that is involved within the dataset.

Machine learning is transforming the landscape of ADHD and depression detection and classification through innovative data collection and analysis methods. These encompass imaging techniques, processing of medical notes, and wearable technology, reflecting ADHD's complex nature and showcasing machine learnings potential in diagnosis and treatment.

ADHD diagnosis has seen successful employment of imaging techniques, leveraging SVM and deep learning models.

Despite needing large data sets and often dealing with unbalanced ADHD-200 datasets, these challenges are overcome using data augmentation and hypothesis testing frameworks.

High classification accuracies from multiple studies reinforce the value of imaging data in ADHD detection. ML has also proven successful in extracting rich clinical informa- tion from medical notes, with Decision Trees, SVMs, and hybrid Al models delivering impressive classification accu- racy.

While there are issues like overfitting and data het-erogeneity, these applications highlight the role of Al in clinical decision-making. Incorporating wearable technology provides a non-invasive means of collecting EEG signals for ADHD classification. Techniques such as CNNs and SVMs have been effective in analyzing this data. However, ensuring the models' applicability to new patients and real-world conditions remains a challenge.

For depression detection, machine learning has simi- larly demonstrated remarkable adaptability and effective- ness.

Brain imaging data, clinical notes, sociodemographic data, laboratory data, wearable sensor data, and electronic health records have all been effectively utilized. Algorithms such as SVMs, Random Forest, ElasticNet, Extreme Gradi- ent Boosting, and Artificial Neural Networks have yielded high accuracy rates across diverse data sources. Moreover, machine learning's success in discerning between differ- ent depressive disorders could revolutionize personalized treatment.

However, the quality of machine learning models is contingent on the quality of data they're trained on. Continued efforts are essential to ensure the robustness and applicability of these models across various populations and settings. The Intelligent Sensing Group at Newcastle University is conducting their own Intelligent Sensing ADHD trial (ISAT) that involves audio-visual data of controls and ADHD subjects.

The aim is for this data to be publicly available once correctly processed.

In conclusion, machine learning offers substantial potential for improving ADHD and depression diagnostics. Despite challenges related to data quality, overfitting, and algorithm interpretability, machine learnings ability to identify patterns in complex datasets makes it a valuable tool in mental health research. Future efforts should focus on creating reliable models and projects.

REFERENCES

[1] J. W. Kanter, A. M. Busch, C. E. Weeks, and S. J. Landes,"The nature of clinical depression: Symptoms, syndromes,and behavior analysis," Behav. Anal., vol. 31, no. 1, pp.1-21, Apr. 2008.

[2]Accessed: aerstanding 2023.er enlive. Available: Accessed:Jan.6, https://www.verywellmind.com/sensitivities-and-adhd-20473

[3] World Health Organisation. Depressive Disorder (Depression). Accessed: Jan. 6, 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression

[4] R. Thomas, S. Sanders, J. Doust, E. Beller, and P.Glasziou,”Preva-lenceof attention-deficit/hyperactivity disorder: A systematic review and meta-analysis," Pediatrics, vol. 135, no. 4, pp. e994- e1001, Apr. 2015.

[5] P. Song, M. Zha, Q. Yang, Y. Zhang, X. Li, and I. Rudan,"The prevalence of adult attention-deficit hyperactivity disorder: A global sys- tematic review and meta-analysis," J.Global Health, vol. 11, Feb. 2021, Art. no. 04009.

[6] U. P. Ramtekkar, A. M. Reiersen, A. A. Todorov, and R. D. Todd,"Sexandagedifferencesin attention-deficit/hyperactivity disorder symptoms and diagnoses: Implications for DSM-V and ICD-11," J. Amer: Acad. Child Adolescent Psychiatry, vol. 49, no. 3, pp. 217-283, 2010.

[7] Z. Huntley, S. Maltezos, C. Williams, A. Morinan, A.Hammon,. Ball, E. J. Marshall, F. Keaney. S. Young.P. Bolton, K. Glase . Howe-Forbes, J.Kuntsi, K. Xenitidis, D. Murphy, and P.Asherson,"Rates of undiagnosed attention deficit hyperactivity disorder in London drug and alcohol detoxification units," BMC Psychiatry, vol. 12, p. 223, Dec. 2012.

[8] BBC. ADHD Diagnosis for Adults 'Can Take Seven Years. Accessed: Jan. 3, 2022. [Online]. Available: https://www.bbc.co.uk/news/ uk-england-44956540#:~ text-About%201.5%20million%2 Oadults%20in,only%20120%2C000%20are%2 0formally%20diagnose

[9] B. T. Wymbs, W. E. Pelham, B. S. G. Molina, E. M. Gnagy,T. K. Wilson, and J. B. Greenhouse, "Rate and predictors of divorce among parents of youth with ADHD," J. Consulting Clin. Psychol., vol. 76, no. 5, pp. 735-744, Oct. 2008.

[10]K. Doheny, "Divorce more likely in ADHD families?" WebMD,Tech. Rep.,2008. [Online]. Available: https://www.cbsnews.com/ news/divorce-more-likely-in-adhd-families/

[11]J. A. Doshi, P. Hodgkins, J. Kahle, V. Sikirica, M. J. Cangelosi, J. Setyawan, M. H. Erder, and P. J. Neumann, "Economic impact of childhood and adult attention-deficit/hyperactivity disorder in the United States," J. Amer: Acad. Child Adolescent Psychiatry, vol. 51, no. 10,S.Chen,B.Mulgrew,andP.M.Grant,―Acluste ringtechniquefor digital communications channel equalization using radial basis functionnetworks,‖ *IEEETrans.on Neural Networks*, vol. 4,pp.570-578,July1993.

[12]C. Mohr-Jensen and H.-C. Steinhausen,"A meta-analysis and systematic review of the risks associated with childhood attention-deficit hyperactivity disorder onlong-termoutcome of arrests, convictions, and incarcera- tions," Clin.Psychol. Rev., vol. 48, pp. 32-42, Aug. 2016.

[13]Diagnostic and Statistical Manual of Mental Disorders(DSM-S-TR), 5th ed., Amer. Psychiatric Assoc., Arlington, VA, USA, 2013.

[14]C. K. Conners, J. Pitkanen, and S. R. Rzepa, Conners (Conners 3; Con- ners 2008), 3rd ed. New York, NY, USA:Springer, 2011, pp. 675-678.

[15]I. Macrory, Measuring National Well-Being: Life in the UK.South Wales, Wales: Office for UK National Statistics, 2016.

[16] W. H. Organisation, "Depression and other common mental disor- ders: Global health estimates,"World HealthTech.Rep.WHO/MSD/MER/2017.2, 2017. [Online]. Avail- able:https://apps.who.int/iris/handle/10665/254 610

[17]J. Alonso et al.,"Treatment gap for anxiety disorders isglobal: Results of the world mental health surveys in 21 countries," Depression Anxiety, vol. 35, no. 3, pp. 195-208, Mar. 2018.

[18]T. Beiwinkel, S. Kindermann, A. Maier, C. Kerl, J. Moock,G. Bar- bian, and W. Rössler,"Using smartphones tomonitor bipolar disorder symptoms: A pilot study," JMIR Mental Health, vol. 3, no. 1, p. e2, Jan. 2016.

[19]U.K.-Government. Labour Force Survey: Self Reported Work-Related Ill Health Workplace Injuries.Accessed:Oct.5,2021.[Online].Availabl e :https://www.hse.gov.uk/statistics/lfs/index.html

[20] D. Stevenson and P. Farmer, "Thriving at work: A review of mental health and employers," Dept. Health Social Care UK, London, U.K., Tech. Rep., Oct. 2017. [Online].Available: https://assets.publishingservice.gov.uk/governm ent/uploads/system/uploads/attachment data/file/658145/thriving-at-work-stevenson-far mer-review.pdf

[21]J. Banks and X. Xu, The Mental Health Effects of the First Two Months of Lockdown and Social Distancing During the COVID-19 Pandemic in the UK. Linköping, Sweden: IFS, Jun. 2020.

[22]S. Gallagher and M. A. Wetherell, "Risk of depression in family care- givers: Unintended consequence of COVID-19," BJPsych Open, vol. 6, no. 6, pp. 1-5, 2020, Art. no. el19.

(23]R. Stewart, A. Jewell, M. Broadbent, I. Bakolis, and J Das-Munshi, "Causes of death in mental health service users during the first wave of the COVID-19 pandemic:South London and Maudsley data from March to June 2020, compared with 2015-2019," MedRxiv, doi: 10.1101/2020.10.25.20219071.

[24]H. Yang, W. Chen, Y. Hu, Y. Chen, Y. Zeng, Y. Sun, Z. Ying,J. He, Y. Qu, D. Lu, F. Fang, U. Valdimarsdóttir, and H.Song, "Pre- pandemic psychiatric disorders and risk of COVID-19: A cohort analysis in the U.K. biobank,".[Online].Available:https://www.thelancet.com/journalslanhl/article/PIIS2666-7568(20)30013-1/fulltext

[25]J. S. Evert, C. C. Hannie, J. J. S. Kooij, M. Marieke, T. F.B. Aartjan, and J. H. D. Dorly, "The role of adverse life events on depression in older adults with ADHD," J.Affect. Disorders, vol. 174, pp. 574-579, Mar. 2015.

[26] A. J. Lundervold, S. P. Hinshaw, L. Sorensen, and M.-B.Posserud, "Co- occurring symptoms of attention deficit hyperactivity disorder (ADHD) in a population-based sample of adolescents screened for depression," BMC Psychiatry, vol. 16, no. 1, pp. 1-10, Dec. 2016.

[27]C. Nash, R. Nair, and S. M. Naqvi, "Machine learning and ADHD mental health detection—A short survey," in Proc. 25th Int. Conf. Inf. Fusion (FUSION), Jul. 2022, pp. 1-8.

[28]D. S. B. A. Hamid, S. B. Goyal, and P. Bedi,"Integration of deep learning for improved diagnosis of depression using EEG and facial features," Mater:Today, Proc., vol. 80, pp. 1965-1969, Jan. 2023

[291A. B. R. Shatte, D. M. Hutchinson, and S. J. Teague,"Machine learning in mental health: A scoping review of methods and applications," Psychol. Med., vol. 49, no. 9, pp. 1426-1448, Jul. 2019.

[30]S. Graham, C. Depp, E. E. Lee, C. Nebeker, X. Tu, H.-C.D. V. Jeste, "Artificial intelligence for mental health and mental ill- nesses: An overview," Current PsychiatryRep., vol. 21, no. 11, p. 116, Nov. 2019.

[31]H. W. Loh, C. P. Ooi, P. D. Barua, E. E. Palmer, F.Molinari, and U. R. Acharya, "Automated detection of ADHD: Current trends and future perspective," Comput.Biol. Med., vol. 146, Jul. 2022, Art. no. 105525.

[32]N. Sethu and R. Vyas, "Overview of machine learning methods in ADHD prediction," in Advances in Bioengineering. Springer, May 2020, pp. 51-71.

[33] The ICD-10 Classification of Mental and Behavioural Disorders: Diag- nostic Criteria for Research, World Health Organization, Geneva, Switzerland, 1993.

[34] Attention Deficit Hyperactivity Disorder, National Institute of Mental Health, Bethesda, MD, USA, 2022.

[35]K. Kroenke, R. L. Spitzer, and J. B. W. Williams, "The PHQ-9: Validity of a brief depression severity measure,"J. Gen. Internal Med., vol. 16, no. 9, pp. 606-613, Sep.2001.

[36]J. Ware and C. D. Sherbourne, "The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection," Med. Care, vol. 30, no.6, pp. 473-483, 1992.

[37]J. Upton, Beck Depression Inventory (BDI). New York,NY, USA: Springer, 2013, pp. 178-179.

[38]G. H. Glover, "Overview of functional magnetic resonance imaging,"Neurosurg. Clinics North Amer., vol. 22, no. 2, p. 133-139, 2011.