Healthcare Analytics 3 (2023) 100194

Contents lists available at ScienceDirect

Healthcare Analytics

journal homepage: www.elsevier.com/locate/health

A novel and efficient Wavelet Scattering Transform approach for primitive-stage dyslexia-detection using electroencephalogram signals

Shankar Parmar [a], Chirag Paunwala [b,*]

[a] Gujarat Technological University, Gujarat, India
[b] Electronics and Communication Department, Sarvajanik College of Engineering and Technology, Athwalines, Surat, 395001, Gujarat, India

ARTICLE INFO

Keywords:
Dyslexia
Electroencephalogram (EEG)
Machine learning
Feature extraction
Wavelet Scattering Transform
Support Vector Machine

ABSTRACT

Dyslexia is a neurological disorder affecting reading and writing abilities. Early intervention is important for affected individuals' social and academic development. The accuracy and objectivity limitations of traditional dyslexia detection systems based on behavioral symptoms and standard tests can pose challenges to the early detection of the condition. In response, an electroencephalogram (EEG) based detection method has been proposed to aid medical professionals in addressing these limitations. A comparison is made between the Wavelet Scattering Transform (WST) approach and three other approaches, namely Spectral Statistical Features (SSF), Connectivity Features with Autoencoders (CFA), and Hybrid Features (HF), using two datasets. These two datasets were chosen for various reasons, including the fact that they were collected during different tasks and from different countries. Another significant factor is that the age range of the participants was 7 and 12 years old, marking the beginning of their educational journey. This age range is ideal for detecting dyslexia in its primitive stages, making these datasets a perfect fit for this research. The performance evaluation of the approaches involved utilizing Support Vector Machine (SVM) classifiers with three non-linear kernels and k-fold cross-validation implementation. The findings suggest that the other three approaches could not achieve more than 80% accuracy, and their accuracy results were inconsistent with each dataset. In contrast, the WST approach achieved a high accuracy rate, with an average accuracy of 96.96% and 97.12% for dataset 1 and dataset 2, respectively, using the Radial Basis Function (RBF) kernel. The accuracy of WST features is further improved to 98.72% and 98.67% through the Majority Voting method. These findings demonstrate the effectiveness and generalization of the WST approach.

1. Introduction

According to the European Dyslexia Association [1], dyslexia is a global issue that affects individuals regardless of their culture or language, with a prevalence rate of 9-12% in the general population. While the majority of those with dyslexia may only experience mild difficulties, approximately 2-4% of the population is significantly impacted by the disorder.

Dyslexia is evaluated by multiple fields of study, including Medicine, Linguistics, Psychology, Pedagogics, and Social Sciences [2]. Each field examines different aspects of dyslexia, such as the biological and neurological underpinnings, linguistic processes, cognitive and behavioral aspects, educational implications, and social and cultural factors. These interdisciplinary approaches provide a comprehensive understanding of dyslexia and inform the development of effective interventions and support [3]. Despite decades of research on the subject, the exact mechanisms underlying dyslexia are still not fully understood. However, it is believed that individuals with dyslexia have difficulty with phonological processing, or the ability to recognize and manipulate the sounds of language [4]. This can lead to difficulty with decoding and encoding written language, which in turn impacts reading comprehension and written expression. Additionally, individuals with dyslexia may also experience challenges with working memory and attention, which can further complicate their learning experiences [5]. The underlying cause of dyslexia is neurobiological in nature, meaning that the difficulties associated with the disorder are physically located in the brain [6].

Research indicates that early identification and intervention can significantly improve reading outcomes for individuals with dyslexia. With appropriate support and instruction, individuals with dyslexia can become skilled readers and lead successful lives [7].

Dyslexia has been traditionally diagnosed by examining an individual's behavioral symptoms, which can be limiting in accurately identifying the condition. The Comprehensive Test of Phonological Processing (CTOPP) measures sound processing for reading and writing. The Woodcock-Johnson (WJ) assesses overall literacy skills. The

* Corresponding author.
E-mail address: chirag.paunwala@scet.ac.in (C. Paunwala).

https://doi.org/10.1016/j.health.2023.100194
Received 25 March 2023; Received in revised form 25 April 2023; Accepted 4 May 2023

Written and Oral Language Scales (WOLS) measures oral and written language abilities, while the Wechsler Individual Achievement Test (WIAT) evaluates academic achievement in reading, writing, and math [8]. Combining the results of these tests with family history and biographical information can help to identify individuals with dyslexia.

Identifying dyslexia through behavioral symptoms has certain limitations. One limitation is that these symptoms may not manifest until a child has already struggled with reading, potentially delaying intervention. Additionally, these symptoms may overlap with those of other learning disabilities, posing a challenge in distinguishing dyslexia from other conditions [9]. Furthermore, certain children may possess compensatory abilities that conceal their symptoms, rendering it challenging to diagnose dyslexia through behavioral observations alone. Finally, using behavioral symptoms only may result in subjective assessments and discrepancies in diagnoses among medical professionals. Therefore, a more comprehensive strategy is required to identify dyslexia.

New advances in neuroscience and brain imaging techniques have enabled researchers to gain a better understanding of the underlying neurobiological mechanisms of dyslexia. These techniques may allow for early identification of dyslexia before behavioral symptoms become apparent, thus enabling early intervention and preventing long-term reading difficulties. Hence, there is a necessity for a more comprehensive method for detecting dyslexia. Identifying speech issues can be crucial for detecting dyslexia and providing appropriate support. Speech therapy can improve language skills and positively impact reading and writing abilities [10].

A range of methodologies has been implemented to identify dyslexia, incorporating brain imaging techniques such as functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and electroencephalography (EEG) [11]. Eye tracking is also emerging as a promising technique for detecting dyslexia. EEG recording has been recognized as a practical and cost-effective approach for investigating neurological deficits [12]. The process of EEG measurement involves the quantification of local electrical currents generated by synaptic transmissions between neurons in response to stimuli [13]. This technique has proven to be effective in detecting particular patterns of brain activity associated with dyslexia, such as decreased activity in the language-processing regions of the left hemisphere [14]. However, the interpretation of EEG data presents a complex challenge that demands specialized expertise and a time-intensive analytical process.

To overcome these challenges, machine learning techniques have been employed to provide a way to automatically analyze EEG data and identify patterns that can indicate dyslexia. Machine learning has shown great promise in the field of dyslexia identification. By integrating EEG with machine learning techniques, dyslexia identification at a younger age may become feasible, offering individuals the necessary resources to achieve success. Currently, research is ongoing to improve the precision and effectiveness of dyslexia detection methods. Such improvements have the potential to considerably enhance the diagnosis and management of dyslexia.

2. Literature review

In this research [15], EEG was conducted on a total of 32 individuals, with 17 of them diagnosed with dyslexia and the remaining 15 deemed neurologically healthy. Altman's Nomogram was used to validate the sample size calculation. A Cognionics EEG headset was utilized to record EEG with a sampling rate of 300 Hz during a writing and typing exercise. The FIR band-pass filters were utilized to extract the five sub-bands of the EEG signal. Each band signal was transformed into a frequency domain signal using Fast Fourier Transform (FFT) technique, and eight statistical features were extracted from each band-separated frequency-converted signal. The participants' features were categorized according to different brain regions to determine the most responsible region for identifying dyslexic individuals. The cubic Support Vector Machine classifier was utilized for classification, with a maximum accuracy of 71.88% for writing tasks and 78.13% for typing tasks. The electrodes located in the frontal region of the brain were identified as being more responsible for producing distinctive patterns that can be used to identify individuals with dyslexia. The research concluded that typing tasks are more effective than writing tasks for identifying dyslexia. This research is constrained by the fact that the participants belong to the adult age group, while our aim is to identify dyslexia in children aged between 7 and 12 years old and to provide remedial sessions at an earlier stage. One more limitation of this research is the lack of utilization of cross-validation, and the indication of maximum accuracy, which is unsuitable for any system.

The Event-Related Potential (ERP) of 32 Hebrew-speaking students in sixth and seventh grade was measured using a 64-electrode EEG system from Biosemi, while they performed a Lexical Decision Task [16]. A normalized Power Spectral Density (PSD) was achieved by transforming an EEG signal into a frequency domain using FFT. Spectral entropy was derived from the normalized PSD as a feature. Low frequency and high frequency components were obtained from the signal using Daubechies wavelet after determining the PSD. From the low frequency component, temporal characteristics such as Latency, Maximum Peak-to-Amplitude Ratio, Total Signal Energy, and Entropy were extracted. From the high frequency component, statistical characteristics like the Zero Crossing Rate, Mean, Standard Deviation, and Skewness were obtained. To increase classification accuracy and reduce computing cost, a ReliefF feature selection algorithm with $k = 10$ nearest neighbors was implemented. For the detection of dyslexic youngsters, a Support Vector Machine (SVM) classifier was used, with an average reported accuracy of 78%. A key discovery of this research is that despite using Naive Bayes and Decision Tree classifiers, they were not as effective as the SVM classifier. An attempt was made to determine which region of the brain is most responsible for detecting dyslexic patterns. Detailed component of wavelet decomposition was explored in this research, as it contains significant features that are often overlooked by most researchers. The main limitation of this study is that feature selection was conducted on the entire dataset, which is unsuitable in the context of machine learning.

In this research [17], 48 youngsters (32 proficient readers and 16 dyslexic readers) from the LEEDUCA study were selected to evaluate learning disabilities using EEG signals recorded with Brain Vision's 32-Channel device. Children's EEG data was recorded while they were exposed to white noise modulated at a fixed rate in the delta/prosodic (2 Hz) frequency band. Spectral estimates were extracted with Welch's periodogram technique, and a connectivity model comprising covariance, precision, and correlation matrices were derived from these estimates. Principal Component Analysis (PCA) was applied to reduce dimensionality. Substantial features were found in the Covariance and Precision matrices. The research has reported a maximum accuracy of 72.8% using PCA and an SVM classifier, without the implementation of any cross-validation method. An improved methodology was introduced in [18], where covariance, precision, and correlation matrices were obtained from time domain and spectral connectivity model. A denoising autoencoder with three layers was used to recover features from temporal and spectral connectivity models, and a Support Vector Machine classifier was trained with these features. The F7 electrode was identified as a key electrode. An average per-subject classification of 76.2% was achieved using the precision matrix of the spectral connectivity model.

The Mitsar 19-channel electroencephalograph system was used to collect EEG data from 30 elementary school pupils aged between 7 and 12 years old by researchers of [19]. During the data collection, the pupils were asked to perform three distinct cognitive tasks, namely the N-Back task, the Spatial Visual N-Back task, and the Visual P300 (Oddball) task. For each channel, 43 features were extracted and sent to an SVM classifier. Rather than computing accuracy, the identification of vital electrodes for each task involved the calculation of the co-occurrence matrix. The study found that gender might be a contributing

factor, and some individuals may have issues with reading, processing sounds, or comprehending word meanings. The list of effective electrodes for each task did not contain a large number of common electrodes.

Ten children aged between seven to eleven years were observed while performing two distinct activities, namely relaxing and writing letters. EEG signals were recorded at a sampling frequency of 256 Hz using gMOBIlab hardware [20]. The data was filtered using the same hardware, and Fast Fourier Transform was employed to convert EEG signals from the time domain to the frequency domain at C3, C4, P3, and P4 electrodes. Although the authors did not attempt any classification, certain patterns were observed for dyslexic and normal children. Dyslexic children exhibited two peaks in the alpha and beta bands while writing letters. One peak occurred at the beginning of the task, while the other occurred toward the end. One of the limitations of this research is that the sample size used for the research was relatively small. This could potentially affect the ability to generalize the findings.

In the research conducted by [21], 11 dyslexic students with weak learning abilities, 11 dyslexic students with capable learning abilities who underwent remedial sessions and showed improvement, and 11 students without learning disabilities were compared. EEG signals were recorded while the students performed three distinct activities: relaxation, reading and writing challenging words, and reading and writing sentences without any challenging words. The EEG signals were analyzed using Daubechies' order 8 wavelet transform, extracting detail levels 3 and 4 which represent the Beta and Theta frequency ranges of an EEG signal. Two signals were reconstructed using the level 3 and level 4 detail coefficients, and the power of the reconstructed signals in the beta and theta bands was determined. The study concludes that the theta/beta ratio can differentiate between normal, poor, and capable dyslexic students, providing distinguishing characteristics [21].

A sample population of 33 adult Hebrew-speaking university students aged 21 to 33 years was used in a study that monitored EEG activity while completing phonological and orthographic tasks [22]. P200 and P300 latencies and amplitudes were measured based on the maximum peak for these components within specific time periods. The research aimed to examine whether brain activity differed between dyslexic and normal readers during phonological and orthographic tasks. The results revealed that dyslexic readers had lower amplitude and later delay in P200 and P300 components compared to normal readers, with the strongest group differences observed for phonological representations.

The study of [23] aimed to investigate the EEG patterns of 54 children aged 7 to 12 years while performing a writing task. The subjects were divided into three groups: 18 capable dyslexic, 18 poor dyslexic, and 18 normal children. Wavelet decomposition was employed to analyze the EEG rhythms, and the ratio of theta band power to beta band power was used as a feature. Four classifiers, namely k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), Extreme Learning Machine (ELM), and Deep Learning (DL), were implemented to classify the EEG signals. The SVM classifier with the RBF kernel achieved an overall accuracy of 88.33%, outperforming the other classifiers. These findings provide valuable insights into the use of EEG signals for the classification of dyslexic and normal children. The results of this study offer significant insights into the potential of the SVM classifier for accurately differentiating between dyslexic and normal children.

In [24], it was explored that dyslexia-related reading difficulties can be attributed to a general inability to learn new sound categories. As part of this exploration, an auditory task was conducted where participants were tasked with learning pitch categories above and below a repeated reference sound. The results showed that dyslexic participants exhibited slower category learning compared to typically developing individuals, and their brains did not adapt to repeated sounds in the same manner as their peers. These findings suggest that impaired sound learning in sensory regions may lead to insufficient representations of speech and nonspeech categories in individuals with dyslexia.

The study investigated the relationship between rhythm, reading, and sound processing in the brains of preschool children using neurocognitive tasks [25]. A total of 156 children (70 females) aged between 3 and 5 years were recruited from Chicago. The study used a drumming task to examine preliteracy skills and frequency following responses (FFRs) in over 150 preschoolers who maintained an isochronous beat. The findings revealed that good synchronizers had less degradation of certain FFR measures while listening in noise, suggesting an interconnectedness between rhythm, preliteracy, and auditory processing in early childhood.

Study of [26] investigates the behavioral and neurophysiological aspects of dyslexia in 74 Chinese children aged 7 to 11 years old. EEG recorded brain activity during verbal and visual tasks. The dyslexia group had slower and less accurate responses than typically developing controls, and showed differences in EEG band power, with higher frontal midline theta and reduced upper-alpha power in the posterior region. The study aimed to find predictors of reading difficulty, and both behavioral and neurophysiological measures were shown to be useful.

A web-based gamified test was developed to screen dyslexic individuals [27]. The test consisted of 32 linguistic exercises and was taken by 3644 native Spanish speakers, including 392 dyslexic participants aged between 7 and 17. The study utilized a Random Forests-based classification technique, resulting in a maximum accuracy of 81.6% for dyslexia screening. The dataset had a class imbalance and lacked a cross-validation strategy, which limits result replication. The research may have some limitations due to the 15-minute duration of the web-based gamified assessment, which some participants may have found to be lengthy. This duration may prove to be a challenge for younger participants or those with attention difficulties, thereby affecting the reliability of their performance.

In [28], a machine learning-based computing system is proposed for measuring various cognitive abilities such as hearing and speech, vision, reading and spelling, writing and speed, arithmetic and time management, memory and cognition, behavior, health, development, and personality metrics. The dataset includes dyslexic and non-dyslexic children aged between 7 and 18 years. The Gibson brain skill exam evaluates nine aspects of cognitive ability. Classification is performed using k-means clustering, artificial neural networks, and fuzzy clustering algorithms. Major components identified are word attack and reading strategy, auditory processing, and visual processing. In this research, it has been suggested that future work should focus on utilizing Brain–Computer Interface (BCI) technology, which uses electroencephalogram (EEG) signals to identify dyslexic youngsters.

An infrared corneal reflection system called Ober-2TM was used to track eye position over time in a study by [29]. This system sampled the horizontal and vertical position of both eyes at a rate of 100 Hz. The study included 185 participants, 97 of whom were at high risk of dyslexia and 88 were considered normal. Participants were between the ages of eight and nine. Using the SVM-Recursive Feature Elimination (RFE) feature selection technique, 48 prominent features have been identified. Linear support vector machines with sequential minimal optimization were used for classification. Unlike previous screening methods that rely on paper-and-pencil protocols, this method does not require a written or vocal response or manual evaluation.

The study in [30] recorded the ASSR using EEG in 36 children aged 7 to 12 years, half of whom had dyslexia or dyslexia and developmental language disorder (DLD). Dyslexic children showed reduced ASSRs for 2 Hz stimulation, but similar ASSRs at 5 Hz and 8 Hz compared to typically-developing children. These results align with previous research indicating that dyslexic children exhibit atypical synchronization between brain oscillations and incoming auditory stimulation at 2 Hz, which may impact speech, phonological, and syntactic processing according to Temporal Sampling theory.

In a study by [31], ERP results from 86 typically developing children (CTR) and 26 children with reading difficulties (RD) were analyzed.

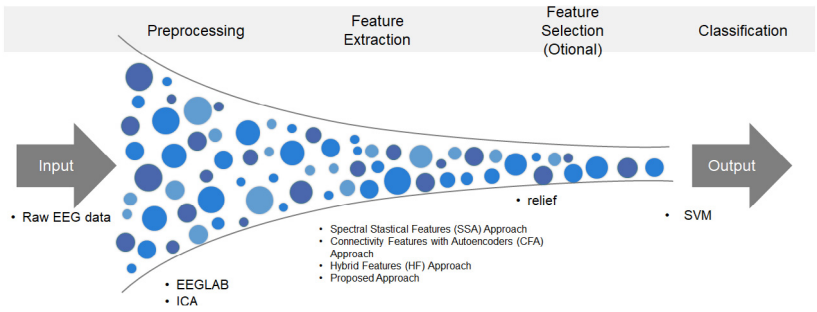


Fig. 1. Funnel diagram displaying research flow.

The participants, aged between 11–13 years, were tested using a passive auditory oddball paradigm that included both native (Finnish) and foreign (English) speech stimuli. The experiment lasted around 20 minutes and was divided into two blocks, with the Finnish stimuli presented first, followed by the English stimuli. The study focused on the change-detection-related ERP responses, specifically the mismatch response (MMR) and the late discriminative negativity (LDN). The results revealed significant differences in language processing between the CTR and RD groups, particularly in their early and late discriminatory responses. These findings suggest a connection between language processing and reading skills in both native and foreign language contexts.

The research of [32] examined the impact of treatment on dyslexic students using a sample of sixteen dyslexic students. Resting-state EEG was recorded for five minutes before and after therapy using a 19-channel Mitstar EEG machine. Various features were extracted from the EEG data, and a combination method of PCA and Sequential Floating Forward Selection (SFFS) was used to extract discriminative features. The study utilized an SVM classifier with the top 12 discriminative features, which indicated features associated with beta-band values were changed after the therapy. Notably, EEG features became more similar to those of normal individuals following therapy.

The paper [33] investigates the use of AI technology in dyslexia education, covering its development, causes and classification of learning disabilities. It analyzes how AI can improve the conditions of dyslexic and dysgraphic individuals in diagnosis, intervention, assessment, and services. The study also highlights the effectiveness of story structure teaching based on visual analysis in enhancing reading comprehension among dyslexic students. These findings can offer useful insights for future research in this field.

3. Methodology

Fig. 1 illustrates the flow of the research through the use of a funnel diagram. The research utilized a raw EEG dataset consisting of two distinct datasets. Detailed descriptions of each EEG dataset were provided, including the number of participants, selection criteria, age group, geographical information, and the EEG machine used to record the data during specific tasks. A comprehensive explanation of the preprocessing stage was presented, detailing the specific steps undertaken for each dataset.

For feature extraction, three existing approaches were implemented and tested on both datasets. Results were then compared with a proposed Wavelet Scattering Transform-based feature extraction technique.

Finally, an SVM classifier with k-fold cross-validation was utilized for classification purposes. Overall, the research employed a comprehensive approach to evaluate the efficacy of different feature extraction techniques and classifiers for EEG data classification.

Table 1
Demographic breakdown by gender and country for data.

Country	Class	Boys	Girls	Total
Germany	Dyslexic	24	22	46
	Normal	27	24	51
France	Dyslexic	34	10	44
	Normal	24	24	48
Finland	Dyslexic	36	21	57
	Normal	25	25	50
Hungary	Dyslexic	27	21	48
	Normal	24	23	47

3.1. EEG dataset

3.1.1. Dataset 1

Participants:

This research involved analyzing EEG data obtained from 391 children aged between 8 and 12 years [34]. Among them, 195 children were diagnosed with dyslexia, while the other 196 were considered normal readers. The sample group was comprised of 221 male children from Finland, Hungary, Germany, and France. A comprehensive explanation of the gender and country distribution can be found in Table 1. The children's reading level was assessed using standardized word reading tests, which were appropriate for their respective languages. The accuracy and rate of reading have been utilized to assess word reading fluency. In Finland and Germany, Electric Geodesics Inc.'s 128-channel EEG system has been used to capture EEG. A 32-channel EEG recorder has been utilized in France. In Hungary, Easycap's 32-electrode-equipped elastic cap has been utilized. In all nations, the sampling rate was 500 Hz.

Stimuli:

Using Praat Software, a stimulus consisting of three synthetic /*y*/ vowels and one /*i*/ vowel has been produced. The five lowest formants of synthetic vowels were synthesized using five tones with comparable frequency as a non-linguistic stimulus. The length of the stimulus was 150 ms, including a 10 ms onset ramp and a 15 ms offset ramp. The stimuli were delivered under two distinct passive oddball circumstances, speech and non-speech. To avoid participants from misinterpreting non-speech stimuli as speech, the non-speech condition was always presented first.

Pre-Processing:

A 50 Hz-tuned notch filter was employed to eliminate line noise. The channels containing several artifacts throughout the data were deemed defective and excluded from the average. Additionally, EEG epochs with voltage deflections exceeding ±200 μV were omitted from the averaging as artifacts. BESA 5.3.1 software is utilized in Finland, while Brainvision Analyzer Software is utilized in Hungary and Germany. Using independent component analysis, eye blink artifacts were reduced. Since the study has been done in four different countries with different EEG recorders, 21 common electrodes have been chosen.

Seven electrodes from the frontal area of the brain have been selected: F7, F3, Fz, F4, F8, FC5, and FC6. Seven electrodes from the temporal and central regions of the brain have also been selected: T7, T8, C3, Cz, C4, CP5, and CP6. Additionally, seven electrodes from the parietal and occipital regions of the brain have been selected: P7, P3, Pz, P4, P8, O1, and O2.

3.1.2. Dataset 2

Participants:

The EEG of 52 students, ranging in age from 7 to 12, has been recorded. For the purpose of identifying students who are dyslexic, Raven's IQ score has been determined [35]. In order to record the signals, a Mitstar 19-channel EEG machine with a sampling frequency of 250 Hz was utilized.

Stimulus:

Detailed explanations of the three different activities in which the 52 students participated while their EEG was recorded are provided below.

1. N-Back Cognitive Task: The objective of this exercise is to evaluate an individual's working memory. The task requires the student to hit the space bar when the preceding shape is repeated. In case the shape is not repeated, the student is required to refrain from pressing the space bar and wait for the following shape. The task was accomplished within a three-minute timeframe.
2. Spatial N-Back Task: The task involves displaying multiple shapes to the child, which can appear in different locations across the four quadrants of a centrally positioned cadre. The cadre is marked by large plus signs and is roughly divided into four parts. The child is required to press the space button only if they observe the same shape consecutively, irrespective of its location. If the shape is not repeated, the child is expected to wait. Within a time span of three minutes, the task was completed.
3. Visual P300 (Oddball) Task: During this task, EEG data was recorded for a duration of seven minutes. In this part of the activity, the shapes appeared randomly either at the top or bottom of the box. The student is expected to press the space button only if the shape appears at the top of the screen, otherwise, they should wait.

Pre-processing:

The following series of preprocessing processes was conducted in order to get rid of any noise that was present. The reference point for the EEG raw data was determined to be the average value. In order to reduce the amount of line noise, a notch filter set at 50 Hz has been implemented. Then, an implementation of a 0.5–50 Hz minimum phase causal bandpass filter was carried out so that only the most important signals could be passed through. The third step, which involved using Independent Component Analysis, consisted of removing artifacts from the signal using the method. Independent Component Analysis (ICA) is a widely used technique in neuroimaging for separating signals of interest from various sources of noise and artifacts. In EEG, ICA is particularly useful for removing artifacts caused by eye movements, muscle activity, and other sources of electrical interference. ICA works by decomposing the EEG signal into a set of independent components, each representing a different source of activity in the brain or in the environment. The components are ranked based on their statistical properties, and the components representing artifacts can be easily identified and removed from the data.

The segmentation of electroencephalography (EEG) data into 1-second intervals is a crucial aspect of EEG data analysis, representing the final step of preprocessing. This step is fundamental as it enables the reliable characterization of brain activity signals by characterizing the EEG signal as a stationary process, which allows for the application of various statistical and signal processing techniques to extract meaningful features from the data.

3.2. Feature extraction

3.2.1. Spectral Statistical Features (SSF) approach

A bandpass FIR filter was applied to the EEG signal in order to extract the delta, theta, alpha, beta and gamma bands for further analysis. This approach was adapted from the work of [15].

As a result, five distinct time domain signals, each representing a unique EEG band, were obtained. These five signals were then transformed into frequency domain signals using a Fast Fourier Transform (FFT).

Subsequently, eight statistical features, including mean, median, standard deviation, skewness, kurtosis, mode, maximum, and minimum, were calculated for each of the five signals separately. Thus, the five bands were obtained and analyzed individually, and the statistical features were calculated separately for each signal, channel, and participant.

3.2.2. Connectivity Features with Auto-encoder (CFA) approach

As described in [17], a feature extraction method based on the relationship between power spectral density (PSD) and EEG signal channels was utilized. This approach aimed to identify differences in brain function between individuals with dyslexia and those without the disorder by examining the connectivity of spectral density estimation at each electrode as a biomarker.

A modified version of Welch's approach [18] was employed to estimate the power spectral density of each subject's signals. This involved segmenting the signal into overlapping parts and generating a modified periodogram for each segment. The resulting periodograms were averaged to produce an adjusted periodogram for each subject's 5-second interval. To create this adjusted periodogram, a "Hanning" window was applied to the signal, and the resulting "average periodogram" was then utilized as a feature for computing "spectral connectivity" in subsequent stages.

The Ledoit–Wolf reduced covariance estimator [36] is used to compute the covariance connectivity, a commonly used measure in evaluating connectivity using neuroimaging or EEG. The shrunk covariance can be determined using a specific expression [37],

$$\beta_s = (1-\delta)S + \delta F \tag{1}$$

The degree of shrinkage is denoted by δ(which controls a bias–variance trade-off), and S is the sample covariance. F is a highly structured estimator whose diagonal elements f_{ii} are the same as those of S, $f_{ii} = s_{ii}$, and the remainder of the diagonal elements are defined as $f_{ij} = \bar{r}\sqrt{s_{ii}sjj}$, where $\bar{r}$ is defined as follows:

$$\bar{r} = \frac{2}{(N-1)N}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\left(\frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}}\right) \tag{2}$$

Once the covariance connectivity is computed, an automated decision is made to determine the appropriate shrinkage parameter δ. The formula shown in Eq. (1) was utilized to estimate the covariance (β_S) and its inverse, which is referred to as precision, was then computed.

The resulting features from the previous steps are then passed through a 3-layer denoising auto-encoder block. This block utilizes unsupervised learning techniques to learn and extract relevant features from the input data. The output from the last layer of the encoder is then fed into a classifier block to predict the presence or absence of dyslexia in the subject. This approach of combining auto-encoder blocks with a classifier block has been found to be effective in various machine-learning applications.

3.2.3. Hybrid Features (HF) approach

This approach involves a comprehensive feature extraction process, where a total of 70 features are extracted from the data. These features are categorized into different types, including statistical, autoregressive, HjOrth, Signal energy and Complexity measures, fractal

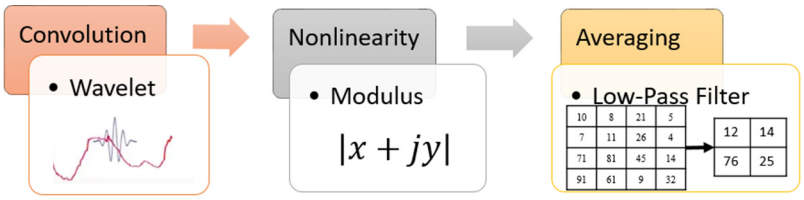


Fig. 2. Generation of coefficients for a signal with the use of the Wavelet Scattering Transform.

dimension, modified center symmetric Local Binary Pattern (LBP), Lempel–Ziv (LZ) complexity, and slow wave index features.

The statistical features comprise four distinct measures, namely arithmetic mean, kurtosis, skewness, and median. On the other hand, the auto-regressive features, which include orders 5, 4, and 3, comprise twelve different features. The Hjorth features consist of three distinct features, namely Activity, Mobility, and Complexity. Meanwhile, the Signal Energy and Complexity Measures category comprises five different features, which are Normalized First Order Difference, Mean Curve Length, Mean Energy, Shannon Entropy, and Hurst Exponent.

In addition to these features, this approach also extracts two fractal dimension features, namely Katz Fractal Dimension and Higuchi Fractal Dimension. Additionally, it includes sixteen Modified Centre Symmetric LBP features, twenty-five LZ Complexity features, and three Slow Wave Index features.

The feature extraction process is described in [38], and involves several steps to ensure that the most relevant information is captured from the data. Once the features have been extracted, a PCA block is used to reduce the size of the feature matrix.

In essence, this approach is intended to be thorough and efficient in capturing the crucial elements of the data. By amalgamating various features, it is capable of furnishing valuable insights into the fundamental patterns and attributes of the data.

3.2.4. Proposed approach

Previous approaches have not been successful in providing generalized results that can be applied to various types of datasets, and they suffer from the limitation of being less accurate. Therefore, a new approach has been proposed that overcomes these shortcomings by using Wavelet Scattering Transform (WST) to extract features in a precise and generalized manner. The pre-processed EEG signals were then transformed using the WST, which is a multiresolution analysis technique. The WST was applied to the EEG signals to extract relevant features for dyslexia detection. The Wavelet Scatter Transform (WST) is a time–frequency analysis technique that provides a robust representation of nonlinear and nonstationary signals. It combines the wavelet transform with a scatter plot to provide a more comprehensive representation of the signal's characteristics.

Due to its structural similarities with a Convolutional Neural Network (CNN), the Wavelet Scattering Transform (WST) can be interpreted as a CNN. Similar to CNN, WST operates in layers, which consist of Convolutions, Nonlinearity, and Pooling. In contrast to CNN, WST does not require the training of networks, and one of the primary benefits of WST is that it performs exceptionally well even with relatively limited datasets. A time domain EEG signal has been processed by a series of layers, and each layer has been broken down into its component parts, which are convolution, modulus as non-linearity, and averaging as pooling respectively. These three steps are described in Fig. 2. WST does not require any learning parameters as input, and as a result, it is suitable for small datasets.

Wavelet Scattering Transform

A time domain signal z can be transformed into frequency domain using Fourier transform,

$$Z(\omega) = \int z(v) e^{(-j\omega v)} dv \tag{3}$$

The Fourier transform of a time-shifted signal $z_\tau(t) = z(t-\tau)$, where a time delay of τ is introduced in the original signal, is given by $Z_\tau(\omega) = e^{(-j\omega\tau)} Z(\omega)$. In this instance, if the modulus of the Fourier transform is determined in both cases, it will be translation invariant, i.e. $|Z_\tau(\omega)| = |Z(\omega)|$. It is possible to localize this shifting by computing a Spectrogram with a window m of duration D in such a way that the entire area under $m(v)$ is equal to unity. Spectrogram can be computed using the provided formula.

$$|Z(t,\omega)| = \left| \int z(v) m(v-t) e^{(-j\omega v)} dv \right| \tag{4}$$

The average amount of energy in a spectrogram Mz is determined using a mel frequency spectrogram by the application of mel scale filters Ψ_λ, with λ serving as the center frequency of each $\Psi_\lambda(\omega)$:

$$Mz(t,\lambda) = \frac{1}{2\pi} \int |Z(t,\omega)|^2 \left|\Psi_\lambda(\omega)\right|^2 d\omega \tag{5}$$

Band pass filters Ψ_λ at high frequencies will have a bandwidth of Q Hz that will not change over the course of the filter's operation. The bandwidth will have an order of Q as a result of the fact that the central frequency is λ. At lower frequencies, rather than having constant Q bandwidth, the bandwidth will be equal to $2\pi/D$. The Mel scale averaging procedure, as explained above, clearly gives temporal warping stability, but at the expense of information loss. To compensate for this loss, frequency averaging can be accomplished by time averaging the output of a filter bank. By applying Plancherel's formula [39],

$$\begin{aligned} Mz(t,\lambda) &= \frac{1}{2\pi} \int |Z(t,\omega)|^2 \left|\Psi_\lambda(\omega)\right|^2 d\omega \\ &= \int |z_t * \psi_\lambda(u)|^2 du \\ &= \int \left| \int z(v) m(v-t) \psi_\lambda(u-v) dv \right|^2 du \end{aligned} \tag{6}$$

If $\lambda \gg Q/D$, then $m(t)$ is nearly constant and $m(v-t)\psi_\lambda(u-v) \approx m(u-t)\psi_\lambda(u-v)$.

Hence,

$$\begin{aligned} Mz(t,\lambda) &\approx \int \left| \int z(v) \psi_\lambda(u-v) dv \right|^2 |m(u-t)|^2 du \\ &= \left| z * \psi_\lambda \right|^2 * |m(t)|^2 \end{aligned} \tag{7}$$

Large wavelet coefficients are amplified using the square operator in the preceding equation. To eliminate outliers, $|z * \psi_\lambda| * m(t)$ has been calculated. Using this equation, high frequency components that were eliminated by the low pass filter m are recovered using a new set of wavelet modulus coefficients. This technique can be cascaded to obtain WST.

$W_0 z(t) = z * m(t)$ can be used to create a translation invariant feature of $z(t)$, which removes all high frequencies. These high frequencies can be recovered by calculating approximate mel frequency spectral coefficients by averaging the wavelet modulus coefficient with m:

$$W_1 z\left(t, \lambda_1\right) = \left| z * \psi_{\lambda_1} \right| * m(t) \tag{8}$$

It is calculated using wavelets ψ_{λ_1} with an octave frequency resolution Q_1. These particular coefficients are referred to as first-order scattering coefficients. Similar to first order coefficients, second order coefficients can be determined by using

$$W_2 z\left(t, \lambda_1, \lambda_2\right) = \left| \left| z * \psi_{\lambda_1} \right| * \psi_{\lambda_2} \right| * m(t) \tag{9}$$

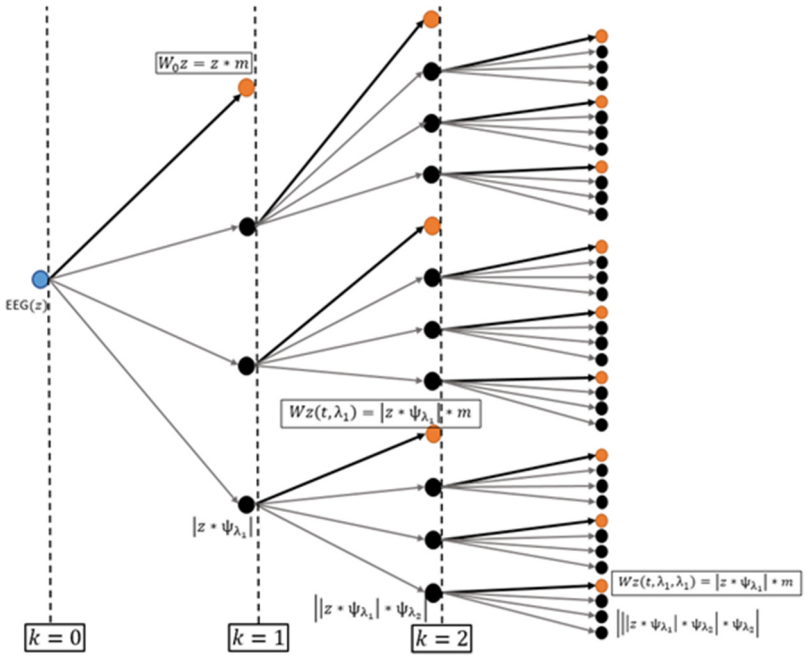


Fig. 3. Generation of coefficients for a signal with the use of the Wavelet Scattering Transform.

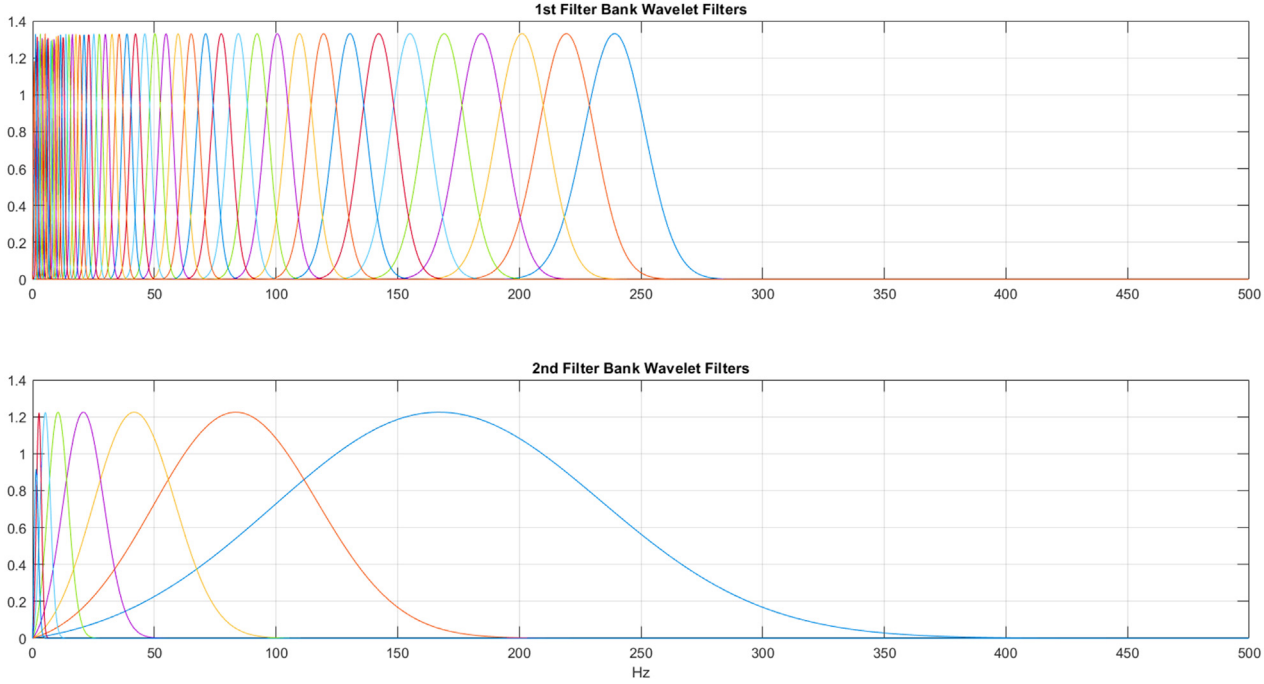


Fig. 4. Wavelet filters.

The octave resolution of wavelets ψ_{λ_2} is Q_2, which may differ from Q_1. Iterating this technique yields scattering coefficients of any order k. Its generic form can be expressed as,

$$W_k z\left(t, \lambda_1, \cdots \lambda_k\right) = \left|\left|\left|z * \psi_{\lambda_1}\right| * \psi_{\lambda_2}\right| \cdots * \psi_{\lambda_k}\right| * m(t) \tag{10}$$

This scattering can also be interpreted as a convolution neural network due to its cascading nature, which consists of convolutions and non-linearities.

The Gabor wavelet, sometimes known as the Morlet wavelet, is the fundamental component of the Wavelet Scattering Transform. This wavelet was selected because of its near similarities to the mammal auditory and visual systems, as well as its deep connections to the visual cortex [40]. The energy of scattering coefficients decreases significantly with increasing layer level, and research has shown that the first two layers capture 99% of the system's total energy [41]. As energy dissipates with each iteration, two layers are sufficient for most applications. Consequently, this study employs a two-layer decomposition with a decomposition order of $k = 2$. In order to obtain the wavelet coefficients for the subsequent layer, the wavelet octave resolutions are tuned at each layer k. It has been demonstrated that selecting $Q_1 = 8$ wavelets per octave for an EEG signal $z[n]$ can generate sparse representations of a combination of physiological signals [39]. When it comes to the second order, selecting $Q_2 = 1$ will create wavelets with narrower time support. These wavelets are more suited to characterize

Table 2
Performance assessment of SSF approach on Dataset1 (Accuracy in %).

	RBF kernel	Polynomial kernel	MLP kernel
Dataset 1	56.02 ± 1.29	54.97 ± 1.49	50.9 ± 1.81

Table 3
Performance assessment of SSF approach on Dataset2 (Accuracy in %).

Dataset 2	RBF kernel	Polynomial kernel	MLP kernel
N-Back task	79.46 ± 3.22	66 ± 5.04	56.73 ± 6.04
Spatial N-Back task	66.37 ± 6.68	57.82 ± 9.57	52 ± 4.97
P300 (Oddball) task	64.73 ± 5.42	61.28 ± 8.66	61.1 ± 3.38

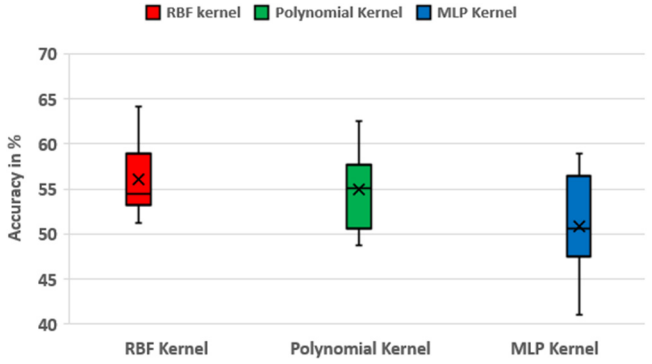


Fig. 5. Box plot illustrating accuracy variation in stratified k-fold cross validation for SSF approach on Dataset 1, with mean accuracy highlighted by cross sign.

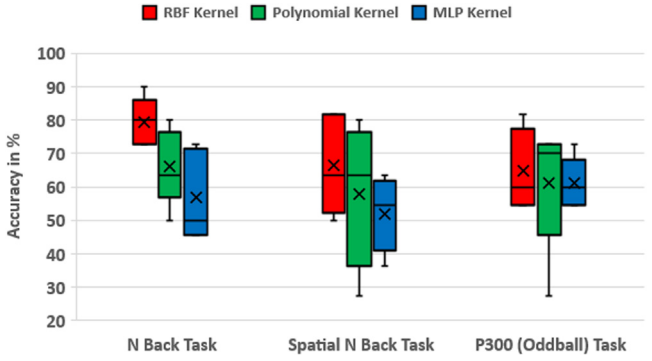


Fig. 6. Box plot showing accuracy variation in stratified k-f old cross validation for SSF approach on Dataset 2, with mean accuracy highlighted by cross sign.

transients and attacks. The time scale of the scaling (lowpass) filter in a wavelet time scattering network can be customized through the use of the Invariance Scale attribute. According to [42], Invariance Scale is equal to half of signal length in samples. The bandwidth of low-pass filters tends to increase as the invariance scale decreases. The graphic in Fig. 3 presents an illustration of the hierarchical process of wavelet scattering transform. Wavelets depicted in the illustration of Fig. 4 have Q_1 equal to 8, and Q_2 equal to 1.

3.3. Support Vector Machine

Support Vector Machine (SVM) has been widely used as an effective classifier for EEG signal classification due to its ability to handle high-dimensional data and nonlinearity of the data [43]. In particular, SVM with non-linear kernels has been shown to outperform linear SVM in EEG signal classification tasks [44].

The mathematical formulation of SVM with non-linear kernel involves mapping the input data into a higher dimensional feature space, where the data can be separated by a hyperplane. The decision function can be expressed as:

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i \kappa(x_i, x) + b \tag{11}$$

where x is a new input data point, n is the number of support vectors, α_i is the Lagrange multiplier, y_i is the corresponding class label of the ith support vector, κ is the kernel function, and b is the bias term.

Three distinct kernels, namely the Radial Basis Function (RBF) kernel, polynomial kernel, and sigmoid or Multilayer perceptron (MLP) kernel, have been employed to accomplish the classification task. These kernels are widely employed in different machine learning algorithms and are renowned for their efficacy in enhancing the accuracy of classification tasks. The mathematical equations for each of these kernels are presented below:

- The RBF kernel function is widely used for EEG signal classification with SVM. It is defined as:

$$\kappa(x, y) = exp(-\gamma \|x - y\|^2) \tag{12}$$

  where γ is the kernel parameter that controls the smoothness of the decision boundary. In this research, the hyperparameter γ was varied with values 0.1, 1, 10, and 100 for the RBF kernel. The analysis of results showed that the highest classification accuracy was obtained when $\gamma = 1$ was used, which is consistent with the previous research studies exploring EEG signal classification with SVM and the RBF kernel [45]. Hence, $\gamma = 1$ has been chosen as the optimal value for the RBF kernel hyperparameter to improve classification performance in this paper.
- The polynomial kernel is defined as:

$$\kappa(x, y) = (1 + x^T y)^d \tag{13}$$

  where, d is the degree of the polynomial. The effect of the polynomial order was investigated by testing the values of $d = 2$ and 3. The results demonstrated that the classification accuracy was significantly better with $d = 2$ than with $d = 3$. Therefore, the polynomial order of $d = 2$ was selected as the optimal value to improve classification performance.
- The sigmoid or Multilayer perceptron (MLP) kernel is defined as:

$$\kappa(x, y) = tanh(\alpha x^T y + \beta) \tag{14}$$

  where, $\alpha = 0.1$ and $\beta = -0.1$ has been chosen for optimization [46].

4. Implementation results and discussion

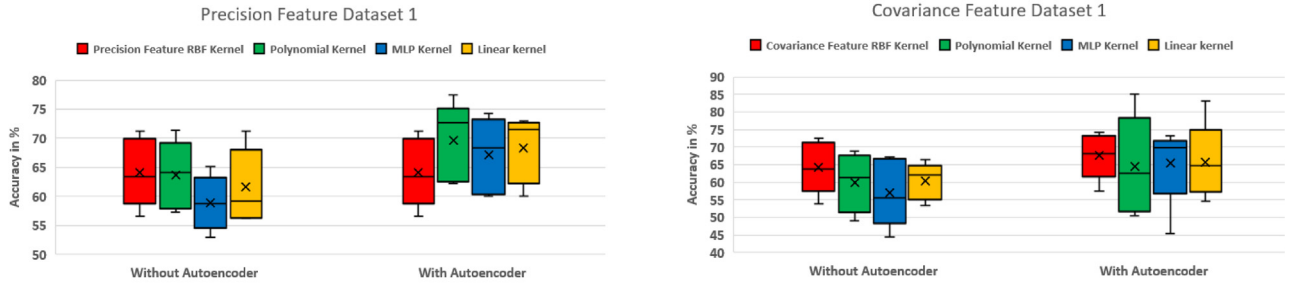
This paper compares the performance of three existing approaches and one proposed approach for a classification task on two datasets — one from children performing an auditory oddball task, and the other from children performing three different tasks. The following section outlines the implementation details of each approach.

4.1. Spectral Statistical Features (SSF) approach

The EEG signals underwent filtering to extract delta, theta, alpha, beta, and gamma bands, which were then converted to frequency domain signals using Fast Fourier Transform. Statistical features were calculated separately for each signal, as explained in Section 3.2.1, and the process was repeated for each individual and channel. This repeated process resulted in a feature matrix size of $N \times ch \times 5 \times 8$, where N is the number of participants and ch is the number of channels used for EEG captures. In other words, for each participant, there were ch number of

Table 4
Performance assessment of CFA approach on Dataset1 and Dataset 2 (Accuracy in %).

		Precision feature			
		RBF kernel	Polynomial kernel	MLP kernel	Linear kernel
Without autoencoder	Dataset 1	65.23 ± 2.26	62.87 ± 3.37	60.23 ± 2.43	61.38 ± 5.43
With autoencoder	Dataset 1	72.54 ± 3.45	68.35 ± 4.45	67.43 ± 3.85	67.78 ± 2.93
		Covariance feature			
		RBF kernel	Polynomial kernel	MLP kernel	Linear kernel
Without autoencoder	Dataset 1	62.58 ± 3.78	59.48 ± 4.63	58.38 ± 5.38	59.87 ± 3.85
With autoencoder	Dataset 1	67.55 ± 3.87	64.57 ± 4.88	65.38 ± 4.53	65.78 ± 3.73
	Dataset 2	Precision feature			
		RBF kernel	Polynomial kernel	MLP kernel	Linear kernel
Without autoencoder	N-Back task	66.43 ± 4.02	61.59 ± 7	58.95 ± 4.35	59.68 ± 3.69
	Spatial N-Back task	56.5 ± 5.96	56.37 ± 3.45	53.02 ± 7.15	54.72 ± 0.37
	P300 (Oddball) task	64.15 ± 4.87	62.44 ± 2.38	58.9 ± 2.6	58.22 ± 5.58
With autoencoder	N-Back task	77.28 ± 4.8	69.75 ± 6.78	66.97 ± 2.04	69.71 ± 3.61
	Spatial N-Back task	66.67 ± 8.34	64.77 ± 4.93	61.1 ± 5.81	61.5 ± 4.26
	P300 (Oddball) task	68.73 ± 4.98	68.61 ± 6.61	59.36 ± 6.47	62.94 ± 5.3
	Dataset 2	Covariance feature			
		RBF kernel	Polynomial kernel	MLP kernel	Linear kernel
Without autoencoder	N-Back task	57.07 ± 1.2	55.22 ± 5.28	52.77 ± 4.45	54.72 ± 5.28
	Spatial N-Back task	56.53 ± 1.85	56.47 ± 6.02	52.64 ± 4.58	52.7 ± 4.5
	P300 (Oddball) task	58.33 ± 2.57	54.92 ± 2.03	52.96 ± 2.71	54.72 ± 0.98
With autoencoder	N-Back task	68.4 ± 4.5	67.11 ± 5.75	61 ± 3.51	67.08 ± 3.43
	Spatial N-Back task	67 ± 7.19	64.03 ± 3.97	60.23 ± 6.45	61.59 ± 9.3
	P300 (Oddball) task	68.44 ± 3.49	67.89 ± 1.56	60.02 ± 4.52	60.62 ± 3.59



(a) Precision feature for Dataset 1

(b) Covariance feature for Dataset 1

Fig. 7. Box plot demonstrating accuracy variation in CFA approach on Dataset 1 using stratified k-fold cross validation.

EEG channels, each with 5 bands, and for each band, eight statistical features were calculated.

To feed this information into a machine learning algorithm, the matrix size was reduced to $N \times m$, where m was produced by concatenating $ch \times 5 \times 8$. This approach ensured that each participant was represented by a unique set of features that captured the EEG activity across different bands and channels. The resulting matrix was used for subsequent analysis and model development.

In order to assess the efficacy of SSF approach, a 5-fold stratified cross-validation method was employed. During this evaluation, the EEG signals were classified into their respective classes by utilizing the SVM classifier, which had been trained on the extracted features. To accomplish this classification task, a range of kernels, including RBF, polynomial, and MLP, were implemented in the SVM classifier.

Table 2 shows the accuracy results for Dataset 1 obtained using SSF approach. The maximum mean accuracy achieved was 56.02% with the RBF kernel. Although the accuracy is not very high, it is still significant given the complexity of the task and the limited number of features used.

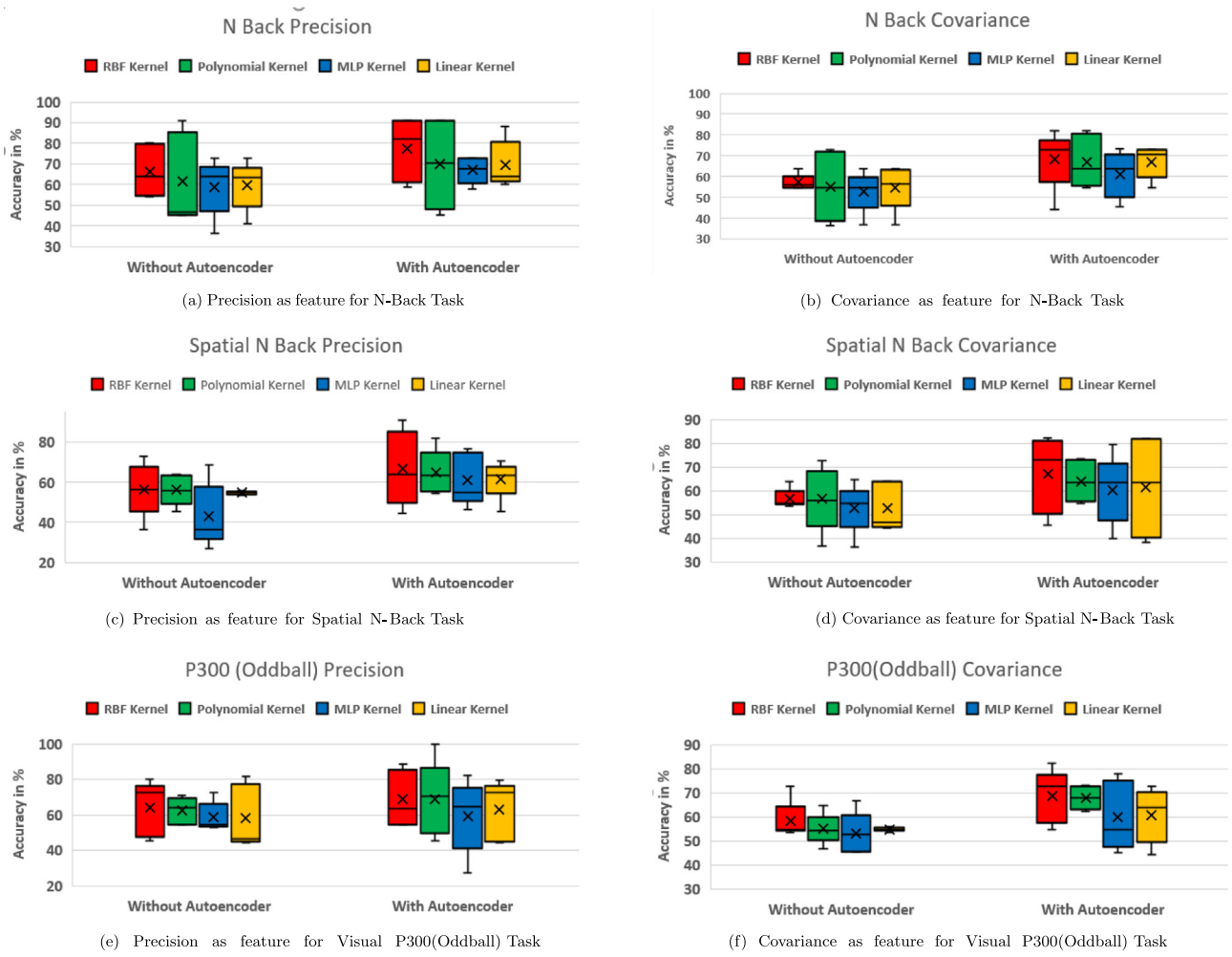
Table 3 shows the accuracy results for Dataset 2 for different tasks. The maximum mean accuracy achieved was 79.46% for the N-Back task with the RBF kernel. These results demonstrate the effectiveness of SSF approach in classifying EEG signals for different tasks and highlight the potential of using statistical features extracted from different frequency bands for EEG signal classification.

The box plot helps to identify the distribution of accuracy values and highlights any outliers or anomalies in the data. The stratified k-fold cross-validation technique is used to evaluate the performance of the classifier, and the resulting accuracy values are displayed in the box plot of Fig. 5 for dataset 1 and in Fig. 6 for dataset 2. The cross sign within the box plot indicates the mean accuracy value.

The high variation in box plot in accuracy in EEG classification can be attributed to the complex and dynamic nature of EEG signals, which can vary significantly depending on the individual, the type of task being performed, and other factors such as the electrode placement and environmental noise.

4.2. Connectivity Features with Auto-encoder (CFA) approach

In Section 3.2.2, the Power Spectral Density (PSD) and connectivity measures of Covariance and Precision were obtained from EEG signals by utilizing the Welch's technique. The obtained features were then provided as input to a three-layer autoencoder block, which consisted of an encoder and a corresponding decoder. The encoder was implemented as a three-layer perceptron with layer sizes of 512, $nhid$, and 3, while the decoder had layer sizes of 3, $nhid$, and 512. The value of $nhid$ was set to 64. The Z-layer of the autoencoder had 3 neurons.



(a) Precision as feature for N-Back Task

(b) Covariance as feature for N-Back Task

(c) Precision as feature for Spatial N-Back Task

(d) Covariance as feature for Spatial N-Back Task

(e) Precision as feature for Visual P300(Oddball) Task

(f) Covariance as feature for Visual P300(Oddball) Task

Fig. 8. Box plot demonstrating accuracy variation in CFA approach on Dataset 2 using stratified k-fold cross validation.

To implement the encoder, the Exponential Linear Unit (ELU) function, which is a widely used activation function, was utilized.

In order to assess the effectiveness of the approach, the features obtained from the PSD and connectivity measures of Covariance and Precision were utilized to measure accuracy. The obtained results were presented in Table 4 and were evaluated using three different non-linear kernels. The performance of the approach was evaluated by measuring the accuracy of classification using SVM classifier. In addition to the aforementioned non-linear kernels, the accuracy of the approach was also evaluated using a linear kernel, and the corresponding results were included in the findings.

The authors of this approach opted for a 5-fold cross-validation approach to evaluate the accuracy of their method. Consequently, a 5-fold cross-validation technique was implemented to assess the performance of the approach.

The obtained results provide valuable insights into the effectiveness of the proposed approach for both datasets. Specifically, for dataset 1, it was observed that the use of the precision feature, in conjunction with an RBF kernel and features extracted from an autoencoder, resulted in a maximum mean accuracy of 72.58%. This represents a notable improvement over the results obtained without the use of an autoencoder, which yielded a mean accuracy of 65.23%.

Similarly, for dataset 2, the utilization of the Precision feature and an autoencoder for the N-Back task resulted in a maximum mean accuracy of 77.28%, demonstrating the high effectiveness of the approach. On the other hand, when the encoder block was not utilized for dataset 2, the maximum mean accuracy achieved was 66.43% using the Precision feature for the N-Back task with an RBF kernel. These findings suggest that the incorporation of an autoencoder in the approach contributes significantly to enhancing its performance.

Fig. 7 illustrates the accuracy variation for dataset 1, where Fig. 7(a) represents the variation with precision as a feature, and Fig. 7(b) displays the variation in accuracy with covariance as a feature. On the other hand, Fig. 8 illustrates the accuracy variation for dataset 2. Specifically, Figs. 8(a) and 8(b) depict the variation in accuracy using box plots for the N-Back task, where Fig. 8(c) represents the variation with precision as a feature, and Fig. 8(d) displays the variation with covariance as a feature. Additionally, Figs. 8(e) and 8(f) also demonstrate the variation in accuracy using box plots for the Spatial N-Back task and Visual P300 (Oddball) task.

4.3. Hybrid Features(HF) approach

As outlined in Section 3.2.3, the third method employed in the research was the derivation of a series of hybrid characteristics from the EEG signals. These features were then used to build a classification model to predict the cognitive tasks performed by the subjects. To reduce the dimensionality of the feature space, a PCA block was employed, which retained 95% of the energy with selected eigenvectors [47].

To evaluate the performance of the approach, five-fold stratified cross-validation was performed using SVM with non-linear kernels such

Table 5
Performance assessment of HF approach on Dataset1 (Accuracy in %).

Dataset 1	RBF kernel	Polynomial kernel	MLP kernel
Without PCA	74.95 ± 2.23	73.13 ± 1.47	52.6 ± 3.24
With PCA	75.6 ± 1.92	74.97 ± 2.37	66.42 ± 2.76

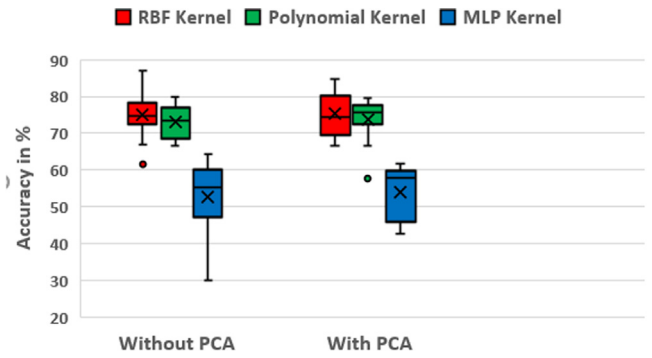


Fig. 9. Box plot showing accuracy variation in stratified k-fold cross validation for HF approach on Dataset 1, with mean accuracy highlighted by cross sign.

as RBF, MLP, and Polynomial. Tables 5 and 6 present the accuracy measurements for each dataset, both with and without the use of a PCA block.

In the first dataset, the maximum mean accuracy of 75.6% was attained through the use of principal component analysis (PCA), representing a noticeable enhancement from the 74.96% accuracy achieved without PCA. In contrast, the second dataset yielded a maximum mean accuracy of 71.27% for the N-Back task when employing the PCA block.

This approach leverages a more comprehensive and diverse set of hybrid features. This extensive incorporation of features in the classification of cognitive tasks based on EEG signals highlights the vast potential of this approach.

The accuracy variation of Dataset 1 and Dataset 2 are presented in Figs. 9 and 10, respectively. The figures illustrate the impact of Principle Component Analysis (PCA) on the accuracy of the datasets under two different scenarios. The accuracy variation of each cognitive task examined in Dataset 2, namely the N-back task (Fig. 10(a)), the Spatial N-Back task (Fig. 10(b)), and the P300 (Oddball) task (Fig. 10(c)), is represented using the box plot technique.

4.4. Proposed approach

In the proposed approach, Wavelet Scattering Transform (WST) based features are extracted from the EEG signals. These features are then supplied to SVM with non-linear kernels.

The output of WST is a tensor which can be represented as a three-dimensional matrix of size, coefficients × scattering window × number of channels. The concatenation technique was utilized so that this three-dimensional tensor could be transformed into a two-dimensional matrix. The number of coefficients and the number of channel dimensions were concatenated to form N. The final matrix size will be window $\times N$, which is an appropriate input for the machine learning method. Each scattering window represents a path of WST, and features are obtained for each path. Consequently, the number of rows that will be sent into the machine-learning algorithm will increase.

Accuracy has been reported using cross-validation method and majority voting method. When using majority voting method, each participant is associated with multiple windows, where each window represents a set of features. If there are W windows, each with F features, and N participants, a label vector of length (WN) is created by repeating each participant's label W times. The input to an SVM with 10-fold cross-validation is a feature matrix of size $(WN) \times F$ and a label vector of size (WN), and the output is a predicted label vector of size (WN). To obtain the final predicted labels, the predicted label vector is grouped into W elements per group, and the majority class in each group is determined using majority voting. The resulting predicted vector of size N is used to calculate the accuracy of the method. Due to majority voting, this method typically achieves higher accuracy.

For dataset 1, the proposed approach achieved a mean accuracy of 96.96% using the cross-validation method, and 98.72% accuracy using the majority voting method. For dataset 2, the proposed approach achieved a mean accuracy of 97.12% using the cross-validation method, and 98.85% accuracy using the majority voting method for N-Back task (see Table 7).

The accuracy variation in the Proposed Approach using Dataset 1 and Dataset 2 is demonstrated through a box plot representation of stratified k-fold cross-validation in Figs. 11 and 12. The box plot includes descriptive statistics of accuracy values for each fold and highlights the mean accuracy through a cross sign.

4.5. Discussion

To achieve optimal performance in analyzing EEG signals, it is essential to consider both spectral and temporal features, as the exclusive focus on spectral features in the SSF approach may lead to lower accuracy. Temporal features contain important information that should not be neglected.

Achieving optimal performance in the CMA approach relies heavily on selecting the appropriate hyperparameters, such as learning rate, batch size, and number of layers. However, identifying the optimal hyperparameters can be computationally intensive and time-consuming. Poor selection of these parameters can significantly reduce the accuracy of the model's output.

The integration of various feature sets and PCA in the Hybrid Features approach can lead to high dimensionality reduction, resulting in the loss of critical information. This can pose a challenge in modeling the underlying feature relationships, which can ultimately lead to lower accuracy in the model's output.

In contrast to the preceding three methods, the WST approach boasts an advantageous characteristic of invariance to signal translations and dilations. This particular property plays a critical role in precisely classifying signals that exhibit variations in both the temporal and frequency domains. Consequently, the WST approach proves to be effective in managing such signals and delivering high classification accuracy.

Individuals with dyslexia are likely to exhibit inadequate performance on tasks that rely on working memory processes, owing to the association between dyslexia and working memory deficits. The N-Back task and the Spatial N-Back task both involve working memory processes, but the Spatial N-Back task requires an additional spatial processing component. This added spatial processing demand may cause additional cognitive load for individuals with dyslexia, who are already struggling with working memory deficits. As a result, individuals with dyslexia may perform poorer on the Spatial N-Back task than on the N-Back task, leading to less accuracy in dyslexia detection when using the Spatial N-Back task as a diagnostic tool. On the other hand, the Visual P300 (Oddball) task necessitates attention and decision-making abilities, which may not be as impaired in individuals with dyslexia.

The RBF kernel outperforms polynomial and MLP kernels in separating complex data distributions by capturing nonlinear feature interactions in higher dimensional space, yielding superior accuracy. With fewer hyperparameters, it is less prone to overfitting and produces smoother decision boundaries that minimize the risk of misclassification.

Table 6
Performance assessment of HF approach on Dataset 2 (Accuracy in %).

	Dataset 2	RBF kernel	Polynomial kernel	MLP kernel
N-Back task	Without PCA	68.18 ± 3.37	65.77 ± 2.55	62.59 ± 3.24
	With PCA	71.27 ± 2.56	68.45 ± 3.45	65.42 ± 3.12
Spatial N-Back task	Without PCA	65.18 ± 1.95	61.67 ± 4.23	58.56 ± 2.84
	With PCA	68.6 ± 3.24	65.45 ± 3.23	62.34 ± 2.78
P300 (Oddball) task	Without PCA	66.27 ± 3.28	63.73 ± 4.23	59.37 ± 3.87
	With PCA	69.85 ± 3.56	67.12 ± 4.27	64.54 ± 3.56

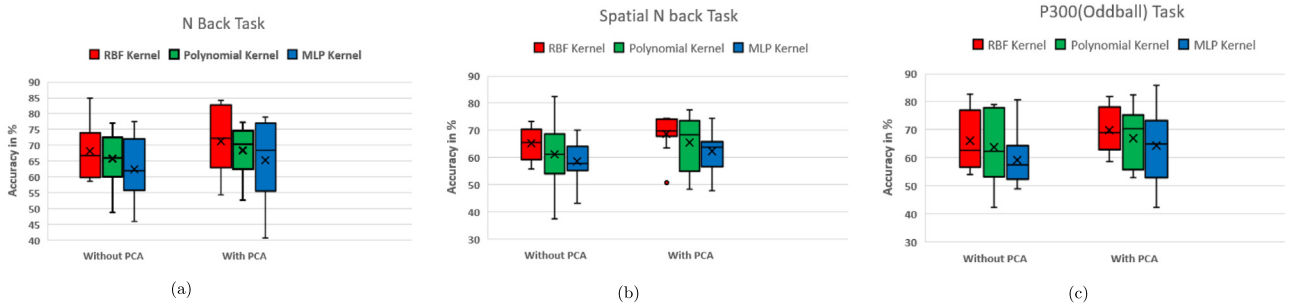


Fig. 10. Box plot showing accuracy variation in stratified k-fold cross validation for HF approach on Dataset 2, with mean accuracy highlighted by cross sign.

Table 7
Performance assessment of proposed approach on Dataset 1 and Dataset 2 (Accuracy in %).

	Cross-validation approach			Majority voting approach		
Dataset 1	RBF kernel	Polynomial kernel	MLP kernel	RBF kernel	Polynomial kernel	MLP kernel
	96.96 ± 1.45	94.52 ± 1.78	90.15 ± 1.56	98.72	96.45	92.34
Dataset 2	RBF kernel	Polynomial kernel	MLP kernel	RBF kernel	Polynomial kernel	MLP kernel
N Back task	97.12 ± 1.85	94.32 ± 2.21	92.55 ± 1.56	98.67	95.67	94.25
Spatial N Back task	93.78 ± 1.97	91.65 ± 2.45	90.87 ± 2.11	94.59	93.48	92.72
P300 (Oddball) task	94.19 ± 1.28	92.76 ± 2.65	91.76 ± 1.89	95.48	94.56	93.63

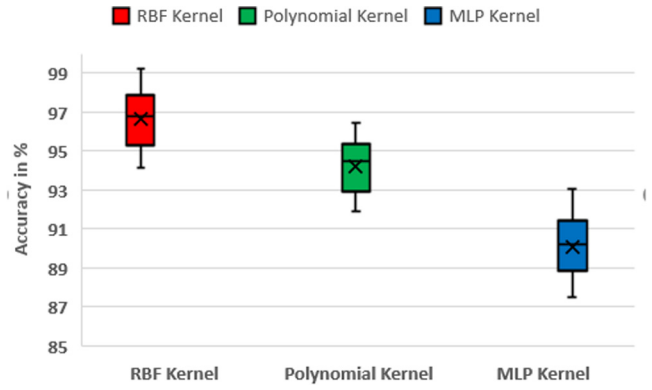


Fig. 11. Box plot showing accuracy variation in stratified k-fold cross validation for proposed approach on Dataset 1, with mean accuracy highlighted by cross sign.

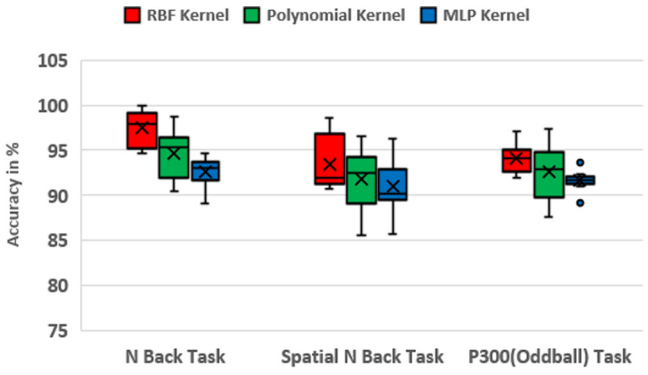


Fig. 12. Box plot showing accuracy variation in stratified k-fold cross validation for proposed approach on Dataset 2, with mean accuracy highlighted by cross sign.

5. Conclusion and future scope

To summarize, the research paper has undertaken a thorough investigation of several kernel functions and feature extraction techniques for identifying dyslexia using EEG data. The evaluation of the results indicates that the combination of WST features and RBF kernel is the most effective approach in distinguishing individuals with dyslexia from those without dyslexia. The findings suggest that this approach outperforms the other three techniques assessed in terms of accuracy.

Additionally, the research has uncovered that the use of the N-Back task in dataset 2 has significantly enhanced the performance of the WST features and RBF kernel approach. This finding provides valuable insights into the potential of this cognitive task to facilitate the identification of dyslexia in EEG data.

The potential of brain area localization in reducing the number of electrodes used in EEG data collection without sacrificing accuracy presents a promising avenue for future research, leading to significant cost and time savings. Additionally, larger datasets incorporating a variety of tasks can further validate the effectiveness of the WST approach and identify the root cause of dyslexia. In future studies, it would be beneficial to compare the efficacy of the WST approach with other machine learning techniques, such as CNN, using larger datasets.

In conclusion, this research provides valuable insights into the effective use of EEG data for dyslexia identification and highlights promising areas for future research. The potential of brain area localization, larger datasets, and comparison with other machine learning techniques can significantly advance understanding of dyslexia detection and contribute to the development of more effective diagnostic and intervention strategies.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

References

[1] What is dyslexia. https://eda-info.eu/what-is-dyslexia/.

[2] T. Gull, S. Khan, A systematic review of research dimensions towards dyslexia screening using machine learning, J. Inst. Eng. (India): Ser. B 104 (2023) http://dx.doi.org/10.1007/s40031-023-00853-8.

[3] T. Helland, Trends in dyslexia research during the period 1950 to 2020-theories, definitions, and publications, Brain Sci. 12 (10) (2022) 1323.

[4] P. Paz-Alonso, M. Oliver, G. Lerma-Usabiaga, C. Caballero, I. González, P. Suárez-Coalla, J.A. Duñabeitia, F. Cuetos, M. Carreiras, Neural correlates of phonological, orthographic and semantic reading processing in dyslexia, NeuroImage: Clin. 20 (2018) http://dx.doi.org/10.1016/j.nicl.2018.08.018.

[5] M.S. Farooq, A. Zulfiqar, S. Riaz, Epileptic seizure detection using machine learning: Taxonomy, opportunities, and challenges, Diagnostics 13 (6) (2023) http://dx.doi.org/10.3390/diagnostics13061058.

[6] M.J. Snowling, C. Hulme, K. Nation, Defining and understanding dyslexia: past, present and future, Oxf. Rev. Educ. 46 (4) (2020) 501–513, http://dx.doi.org/10.1080/03054985.2020.1765756, PMID: 32939103.

[7] S.E. Shaywitz, B.A. Shaywitz, Dyslexia (specific reading disability), Biol. Psychiatry 57 (11) (2005) 1301–1309.

[8] S.K. Parmar, O.A. Ramwala, C.N. Paunwala, Performance evaluation of SVM with non-linear kernels for EEG-based dyslexia detection, in: 2021 IEEE 9th Region 10 Humanitarian Technology Conference (R10-HTC), 2021, pp. 1–6, http://dx.doi.org/10.1109/R10-HTC53172.2021.9641696.

[9] S.E. Shaywitz, R. Morris, B.A. Shaywitz, The education of dyslexic children from childhood to young adulthood, Annu. Rev. Psychol. 59 (1) (2008) 451–475.

[10] U. Shah, M. Alzubaidi, F. Mohsen, A. Abd-Alrazaq, T. Alam, M. Househ, The role of artificial intelligence in decoding speech from EEG signals: A scoping review, Sensors (Basel) 22 (18) (2022).

[11] M. Vandermosten, F. Hoeft, E.S. Norton, Integrating MRI brain imaging studies of pre-reading children with current theories of developmental dyslexia: A review and quantitative meta-analysis, Curr. Opin. Behav. Sci. 10 (2016) 155–161.

[12] A. Gupta, J.S. Kirar, A novel approach for extracting feature from EEG signal for mental task classification, in: 2015 International Conference on Computing and Network Communications (CoCoNet), 2015, pp. 829–832, http://dx.doi.org/10.1109/CoCoNet.2015.7411284.

[13] N. Ahire, R.N. Awale, S. Patnaik, A. Wagh, A comprehensive review of machine learning approaches for dyslexia diagnosis, Multimedia Tools Appl. (2022) http://dx.doi.org/10.1007/s11042-022-13939-0.

[14] J. Dushanova, Y. Lalova, A. Kalonkina, S. Tsokov, Speech-brain frequency entrainment of dyslexia with and without phonological deficits, Brain Sci. 10 (12) (2020) 920.

[15] H. Perera, M.F. Shiratuddin, K. Wong, K. Fullarton, EEG signal analysis of writing and typing between adults with dyslexia and normal controls, Int. J. Interact. Multimed. Artif. Intell. 5 (2018) 62, http://dx.doi.org/10.9781/ijimai.2018.04.005.

[16] A. Frid, L.M. Manevitz, Features and machine learning for correlating and classifying between brain areas and dyslexia, 2018, http://dx.doi.org/10.48550/ARXIV.1812.10622, arXiv https://arxiv.org/abs/1812.10622.

[17] F.J. Martinez-Murcia, A. Ortiz, R. Morales-Ortega, P.J. López, J.L. Luque, D. Castillo-Barnes, F. Segovia, I.A. Illan, J. Ortega, J. Ramirez, J.M. Gorriz, Periodogram connectivity of EEG signals for the detection of dyslexia, in: J.M. Ferrández Vicente, J.R. Álvarez-Sánchez, F. de la Paz López, J. Toledo Moreo, H. Adeli (Eds.), Understanding the Brain Function and Emotions, Springer International Publishing, Cham, ISBN: 978-3-030-19591-5, 2019, pp. 350–359.

[18] F.J. Martinez-Murcia, A. Ortiz, J.M. Gorriz, J. Ramirez, P.J. Lopez-Abarejo, M. Lopez-Zamora, J.L. Luque, EEG connectivity analysis using denoising autoencoders for the detection of dyslexia, Int. J. Neural Syst. 30 (07) (2020) 2050037, http://dx.doi.org/10.1142/S0129065720500379, PMID: 32466692.

[19] R. Kheyrkhah, S. Setarehdan, The Effective Brain Areas in recognition of dyslexia, Int. Clin. Neurosci. J. 7 (2020) 147–152, http://dx.doi.org/10.34172/icnj.2020.16.

[20] N.B. Mohamad, K.Y. Lee, W. Mansor, Z. Mahmoodin, C.W.N.F. Che Wan Fadzal, S. Amirin, Dyslexic frequency signatures in relaxation and letter writing, in: N.T. Nguyen, K. Jearanaitanakij, A. Selamat, B. Trawiński, S. Chittayasothorn (Eds.), Intelligent Information and Database Systems, Springer International Publishing, Cham, 2020, pp. 109–119.

[21] Z. Mahmoodin, W. Mansor, K.Y. Lee, A.Z.A. Zainuddin, Electroencephalogram theta-beta band power features generated from writing for the classification of dyslexic chidren, in: 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences, IECBES, 2018, pp. 288–292, http://dx.doi.org/10.1109/IECBES.2018.8626608.

[22] A. Meyler, Z. Breznitz, Impaired phonological and orthographic word representations among adult dyslexic readers: evidence from event-related potentials, J. Genet. Psychol. 166 (2) (2005) 215–238.

[23] A.Z. Ahmad Zainuddin, W. Mansor, Y.K. Lee, Z. Mahmoodin, Machine learning and deep learning performance in classifying dyslexic children's electroencephalogram during writing, Int. J. Electr. Comput. Eng. (IJECE) 12 (2022) 6614, http://dx.doi.org/10.11591/ijece.v12i6.pp6614-6624.

[24] A. Gertsovski, M. Ahissar, Reduced learning of sound categories in dyslexia is associated with reduced regularity-induced auditory cortex adaptation, J. Neurosci. 42 (7) (2022) 1328–1342.

[25] S. Bonacina, S. Huang, T. White-Schwoch, J. Krizman, T. Nicol, N. Kraus, Rhythm, reading, and sound processing in the brain in preschool children, NPJ Sci. Learn. 6 (1) (2021) 20.

[26] J. Wang, S. Huo, K.C. Wu, J. Mo, W.L. Wong, U. Maurer, Behavioral and neurophysiological aspects of working memory impairment in children with dyslexia, Sci. Rep. 12 (1) (2022) 12571.

[27] L. Rello, R. Baeza-Yates, A. Ali, J.P. Bigham, M. Serra, Predicting risk of dyslexia with an online gamified test, PLOS ONE 15 (12) (2020) 1–15, http://dx.doi.org/10.1371/journal.pone.0241687.

[28] H.M. Al-Barhamtoshy, D.M. Motaweh, Diagnosis of dyslexia using computation analysis, in: 2017 International Conference on Informatics, Health & Technology, ICIHT, 2017, pp. 1–7, http://dx.doi.org/10.1109/ICIHT.2017.7899141.

[29] M. Nilsson Benfatto, G. Öqvist Seimyr, J. Ygge, T. Pansell, A. Rydberg, C. Jacobson, Screening for dyslexia using eye tracking during reading, PLoS One 11 (12) (2016) e0165508.

[30] V. Peter, U. Goswami, D. Burnham, M. Kalashnikova, Impaired neural entrainment to low frequency amplitude modulations in english-speaking children with dyslexia or dyslexia and DLD, Brain Lang. 236 (2023) 105217, http://dx.doi.org/10.1016/j.bandl.2022.105217.

[31] V. Peter, U. Goswami, D. Burnham, M. Kalashnikova, Impaired neural entrainment to low frequency amplitude modulations in english-speaking children with dyslexia or dyslexia and DLD, Brain Lang. 236 (2023) 105217, http://dx.doi.org/10.1016/j.bandl.2022.105217.

[32] A. Oliaee, M. Mohebbi, S. Shirani, R. Rostami, Extraction of discriminative features from EEG signals of dyslexic children; before and after the treatment, Cogn. Neurodyn. 16 (6) (2022) 1249–1259, http://dx.doi.org/10.1007/s11571-022-09794-2.

[33] H. Yang, Effect of story structure instruction based on visual analysis on reading comprehension intervention for dyslexic students, Comput. Intell. Neurosci. 2022 (2022) 9479709.

[34] P.H.T. Leppänen, D. Tóth, F. Honbolygó, K. Lohvansuu, J.A. Hämäläinen, J. Bartling, J. Bruder, Y. Chaix, S. Iannuzzi, R. Nenert, N. Neuhoff, S. Streiftau, A. Tanskanen, J. Tuomainen, J.-F. Demonet, G. Schulte-Körne, V. Csépe, N.W. group, Reproducibility of brain responses: High for speech perception, low for reading difficulties, Sci. Rep. 9 (1) (2019) 8487, http://dx.doi.org/10.1038/s41598-019-41992-7.

[35] R. Kheyrkhah Shali, S.K. Setarehdan, The impact of electrode reduction in the diagnosis of dyslexia, in: 2020 27th National and 5th International Iranian Conference on Biomedical Engineering, ICBME, 2020, pp. 118–125, http://dx.doi.org/10.1109/ICBME51989.2020.9319431.

[36] O. Ledoit, M. Wolf, A well-conditioned estimator for large-dimensional covariance matrices, J. Multivariate Anal. 88 (2) (2004) 365–411, http://dx.doi.org/10.1016/S0047-259X(03)00096-4.

[37] V. Sakkalis, Review of advanced techniques for the estimation of brain connectivity measured with EEG/MEG, Comput. Biol. Med. 41 (12) (2011) 1110–1117.

[38] S.K. Parmar, C.N. Paunwala, Significance of pre-processing phase and dimensionality reduction in EEG-based dyslexia diagnosis with novel features, in: 2022 IEEE 19th India Council International Conference, INDICON, 2022, pp. 1–6, http://dx.doi.org/10.1109/INDICON56171.2022.10039942.

[39] J. Andén, S. Mallat, Deep scattering spectrum, 2013, CoRR arXiv:1304.6763, 1304.6763.

[40] T.S. Lee, Image representation using 2D gabor wavelets, IEEE Trans. Pattern Anal. Mach. Intell. 18 (10) (1996) 959–971, http://dx.doi.org/10.1109/34.541406.

[41] J. Andén, S. Mallat, Multiscale scattering for audio classification, 2011, pp. 657–662.

[42] S. Mallat, Group invariant scattering, 2011, http://dx.doi.org/10.48550/ARXIV.1101.2286, arXiv https://arxiv.org/abs/1101.2286.

[43] M.C. Guerrero, J.S. Parada, H.E. Espitia, EEG signal analysis using classification techniques: Logistic regression, artificial neural networks, support vector machines, and convolutional neural networks, Heliyon 7 (6) (2021) e07258, http://dx.doi.org/10.1016/j.heliyon.2021.e07258.

[44] A. Al-Nafjan, M. Hosny, Y. Al-Ohali, A. Al-Wabil, Review and classification of emotion recognition based on eeg brain-computer interface system research: A systematic review, Appl. Sci. 7 (12) (2017) http://dx.doi.org/10.3390/app7121239.

[45] A. Subasi, M. Ismail Gursoy, EEG signal classification using PCA, ICA, LDA and support vector machines, Expert Syst. Appl. 37 (12) (2010) 8659–8666, http://dx.doi.org/10.1016/j.eswa.2010.06.065.

[46] A. Vora, C.N. Paunwala, M. Paunwala, Statistical analysis of various kernel parameters on SVM based multimodal fusion, in: 2014 Annual IEEE India Conference, INDICON, 2014, pp. 1–5, http://dx.doi.org/10.1109/INDICON.2014.7030414.

[47] Choosing a subset of principal components or variables, in: Principal Component Analysis, Springer, New York, NY, 2002, pp. 111–149, http://dx.doi.org/10.1007/0-387-22440-8_6.