Beetle Swarm-based Multi Verse Optimization for the Autism Spectrum Disorder Detection via EEG Signal Analysis: A Novel Hybrid Deep Learning Concept

Nagendran Krishnaraj * **Mohanraj Elangovan** **

* *Assistant Professor, Department of Artificial Intelligence and Data Science, Saveetha Engineering College, Thandalam, Chennai-602105, Tamil Nadu,India.(e-mail: nagendrankcse@gmail.com).*
** *Professor and Head, Department of Artificial Intelligence and Data Science, K.S.Rangasamy College of Technology, Tiruchengode,Tamil Nadu, India. (e-mail: csemohanraj@gmail.com)*

Abstract: Traditional automated EEG-based diagnosis of ASD typically focuses on distinguishing children with autism spectrum disorder (ASD) from their typically developing peers without addressing the severity of their autistic traits. Using EEG research to identify possible variations among children having mild and severe ASD is a serious challenge. Therefore, this paper plans to diagnose ASD through EEG signal processing and analysis. This dataset was initially collected from the ASD dataset. Median filtering and normalization complete the preprocessing. The signal is converted into an image using the STFT method. GLCM is used to extract features from the transformed image. Since the extracted features are very long, PCA is used to select the essential features. The classification uses a hybrid deep learning model based on these ideally picked features. In this case, SVM replaces CNN's fully connected layer. By integrating BSO and MVO, the novel Optimized BS-MVO improves CNN's learning rate and the number of iterations of SVM, achieving an improved CNN that increases accuracy as an exercise or objective function. The final output is classified (ASD or not) using this enhanced CNN. The accuracy of the proposed BS-MVO at 90% learning percentage is higher at 94.49% compared to GA with 92.90%, WOA at 89.99%, MVO at 93.91%, and BSO at 91.60%, respectively. The simulation results show that the proposed method outperforms the traditional methods in all aspects.

Keywords: autism spectrum disorder; EEG Signal Analysis; Novel Hybrid Deep Learning Model; Beetle Swarm-based Multi Verse Optimization

1. INTRODUCTION

The ability to describe and diagnose a variety of neurological illnesses has been demonstrated by representing the brain's information processing systems as a huge network Ibrahim et al. (2017). ASDs are one instance of abnormalities in this network's normal operation Sheikhani et al. (2010). The existence of compulsive movements of narrow impairment and interests in socio-communicative capabilities are the basic signs of an ASD Sheikhani et al. (2008). Because ASDs represent a wider range of illnesses rather than a singular diagnosis, treatment must be tailored to every kid individually. For determining the success of a therapy, one should preferably depend on objective metrics that quantitatively represent physiological variations in the brain Ahmadlou et al. (2012a). Presently, the success of a therapy is determined by behavioral method (i.e., seeing if any change happens as a result of the treatment), which is susceptible to subjective assessment and hence may add systematic bias Alhaddad et al. (2012). DTI, resting-state fMRI, fMRI, or EEG have all been used to measure the impact of intensive behavioral therapy on the function and brain structure of ASD patients Vissers et al. (2011). This explains how brain connection analysis might be employed to provide a group of objective measurements for defining an autistic child's true patho physiological condition.

When choosing an imaging approach, it is critical for knowing the basic processes of illnesses like ASD Matthew et al. (2015). As a result, the knowledge on brain networks offered by fMRI is inadequate for a complete and appropriate characterization, necessitating the use of EEG data Billeci et al. (2016). True ubiquitous EEG systems, like the ENOBIO system, have recently appeared, including wireless data transfer and dry electrodes. This enables EEG recorded as much as needed at home Ahmadlou et al. (2012b). Nevertheless, there are no convergent results, necessitating a multimodal investigation to establish a technique that may disclose network alterations comparable to those discovered by DTI or fMRI Bosl et al. (2011). Prior research, on the other hand, was confined

to just distinguishing children having autism from those who were developing normally on the basis of the ASD - (DSM) and without considering the symptomsAlsaggaf and Kamel (2014). Even though the DSM defines the gold standard for distinguishing children having autism from those who are normally developing, earlier studies have suggested that much socio-behavioral scales, like the CARS or the ATEC, be used instead Jamal et al. (2014). Furthermore Kang et al. (2020), it is critical to investigate a variety of analytic approaches. Since EEG signals contain both non stationary and nonlinear characteristics, a novel approach on the basis of EMD, SODP, and ellipse area was developed to evaluate non stationary and nonlinear EEG signalsKana et al. (2011). The novel SODP and EMD approach has been shown to be quite helpful in distinguishing among several non-neurological and neurological illnesses, although it has yet to be used to ASD Dawson et al. (2012).

This novel is summarized as follows: Section II describes the literature survey; Section III defines the proposed model and pre-processing for ASD detection; Section IV discusses signal-to-image conversion and feature extraction for ASD detection; Section V elucidates hybrid deep learning model-based EEG signal processing for ASD detection with results and discussion in Section VI; and Section VII concludes with future work.

2. LITERATURE SURVEY

Cociu et al. (2017) have developed a unique method for characterizing an autistic brain by combining several neuro imaging methods. Various methodologies have historically been utilized to know the neurobiological basis of ASD using a multimodal or combined technique of fMRI, EEG, and DTI. Here, we look at how only those can reveal the information shared by the two. Initially, source localization from EEG was performed utilizing ICA and dipole fitting. This enables time changing to be estimated using EEG's high temporal resolution. The predicted functional connection among two active areas can be associated according to the findings. This defines an initial step toward employing ubiquitous EEG without having to do costly DTI or fMRI studies on a regular basis.

Alturki et al. (2021) the examination and analysis of EEGs has been employed to diagnose epilepsy and ASDs. Artifacts were eliminated from EEG datasets employing ICA, while noise and interference were eliminated utilizing a fifth-order band-pass Butterworth filter. Then, employing a CSP, novel approaches were employed to extract the characteristics of EEGs. The usage of variance in traditional CSP is well recognized. Nevertheless, to extract features from EEGs, the usage of energy, entropy, and band power having CSP was recommended. Furthermore, in the study, four classification approaches were used: LDA, SVM, KNN, and ANN, with the goal of evaluating the techniques and selecting the best combination for ASD and epilepsy diagnosis. Furthermore, we looked at the impacts of frequency band, segment length, and reduction count on the findings. The suggested approaches were tested using two EEG datasets.

? youngsters having autism and healthy children had their resting-state EEG data compared. The functional connectivity was calculated using the dDTF approach. The directed temporal network metrics were established and employed to measure the functional brain connections in the frequency bands. The findings revealed that every one among the global measurements was successful in establishing a substantial difference. In the entire groups, the burstiness differed across complete frequency ranges. In every group, the important nodes and edges were also identified. In the Alpha, Beta1, Delta, and Theta bands, the count of important bursty was lower than in the healthy group.

Shillington and Capal (2020) in children having abnormal EEGs and ASD, the genetic test technique was investigated. Data was gathered from the medical records of an earlier reported cohort of kids with very well ASD who were evaluated by Developmental Pediatrics at CCHMC. ASD without epilepsy but having aberrant EEG data was split into two groups: ASD with epilepsy and ASD without epilepsy. Data on EEG was extracted from reports. We investigated genetic testing data in this follow-up research, specifically the percentage of this cohort.

Lavanga et al. (2021) have looked at the ability of early-life EEG backdrop anomalies or "dysmature" qualities to forecast ASD or neuro developmental success in newborns with TSC. During the EPISTOP experiment, EEG data from TSC patients were obtained continuously (NCT02098759). Neuro developmental performance and ASD risk were measured at the age of two years in subjects who were less than four months. MSE, MFA were used to assess the EEG at the initial visit. Applying LDA, these characteristics were linked to both abnormal Bayley and ASD outcomes in the babies.

Vettori et al. (2019) to resolve this concern, researcher observed eye tracking and EEG. While both groups had stronger frequency-tagged neural activity to features than to concurrently displayed homes, this impact was substantially higher. After 5 seconds of sensory stimulation, there exists a substantial variation in group sensitivity of social vs non-social processing, which persists all via the experiment. Eye tracking as well as EEG values were substantially associated. We suggest that frequency-tagging EEG offered a quick, objective, and computed the real of reduced desire for social information in ASD by permitting tracking of both hidden and overt processes.

Hadoush et al. (2019) have utilized EEG analysis and EMD and SODP methods. The IMFs features, elliptical area, SODP patterns, and CTM values were identified using EMD research on their EEG data. The accuracy of this method's primary finding in discriminating among the two ASD groups was then determined using an ANN.

Capal et al. (2018) have discovered a link among seizures and/or abnormalities in EEG data, as well as disease-related deficits in children having ASD. Data was gathered from the medical records of kids having well-defined ASD at CCHMC. The assessments included formative (MSEL), as well as medications, developmental and birth histories, and medical comorbidities. The existence, type, and location of anomalies in EEG data were extracted from reports.

Abdolzadegan et al. (2020) the goal of this study was to offer a reliable approach for detecting ASD early in the

EEG data. Some of the nonlinear and linear properties were used to describe the EEG signal in the suggested method. Furthermore, for artefact removal and resilience, Density-oriented Clustering was used. Furthermore, other criteria like IG, MI, mRMR, and GA are used to pick features. For the final choice, the KNN and SVM classifiers were utilized. As a consequence, the inquiry found that the strategy employing SVM has a classification accuracy of 90.57 percent, whereas KNN has a classification accuracy of 72.77 percent. Furthermore, the suggested technique has a sensitivity of 99.91 percent for SVM and 91.96 percent for KNN. Experiments also reveal that DFA, Entropy, LE, and SL characteristics contained a significant impact on classification accuracy.

Demuru and Fraschini (2020) employed characteristics collected from two modalities (eye-tracking and EEG) as input to SVM to identify autistic toddlers. The current study included 97 children ranging in age from 3 to 6. For the classification of autistic vs generally functioning youngsters, the MRMR feature selection approach was paired with SVM classifiers. Katukuri Arun Kumar et al. **?** used a Median filter (MF), Otsu thresholding, and RBS-MVO techniques for pre-processing, feature selection, and classification process, respectively.

The findings allow for associating the anticipated functional relationship between two active zones. This outlines a first step towards using ubiquitous EEG without needing to do expensive DTI or fMRI experiments regularly. The results showed that every single worldwide measurement had been able to demonstrate a sizable difference. The burstiness varied across all frequency ranges in the full groupings. The number of significant bursts was lower than in the group of healthy people. We argue that by enabling the surveillance of both covert and overt activities, frequency-tagging EEG provided a rapid, accurate, and computed the actual diminished need for social information in ASD. These impacts lead to developing a novel hybrid methodology and intelligent deep learning models for ASD detection via EEG signal. The significant contributions of the work are,

- A novel presents a hybrid deep learning approach to diagnose ASD through EEG signal processing and analysis.
- To use the improved CNN for classification, we replace the fully connected layers of the CNN with the SVM and improve the CNN's learning rate and the number of iterations of the SVM to maximize accuracy.
- To propose a novel optimization approach called BS-MVO to improve the classification status of the developed EEG model and compare the proposed model with existing methods to demonstrate its superiority.

3. PROPOSED MODEL AND PRE-PROCESSING FOR AUTISM SPECTRUM DISORDER DETECTION

A. Proposed Model

The proposed model mainly consists of six phases such as data collection, pre processing, signal to image conversion, feature extraction, optimal feature selection, and classification. First, the dataset is collected from ASD dataset. Next, the pre processing is accomplished using median filtering and normalization. Then, the signals are converted to image using STFT. From the converted images, the features are extracted employing GLCM. As the extracted features are lengthy, significant features are selected from it using PCA. From these optimally selected features, the classification is done using hybrid deep learning model. Here, the fully connected layer of CNN is replaced by SVM. The learning rate in CNN as well as iteration in SVM is optimized by BS-MVO with the intention of accuracy maximization as the fitness or objective function thus referred as enhanced CNN. This enhanced CNN classifies the final output. The developed model for the EEG signal-based autism spectrum disorder is shown in Fig. 1.

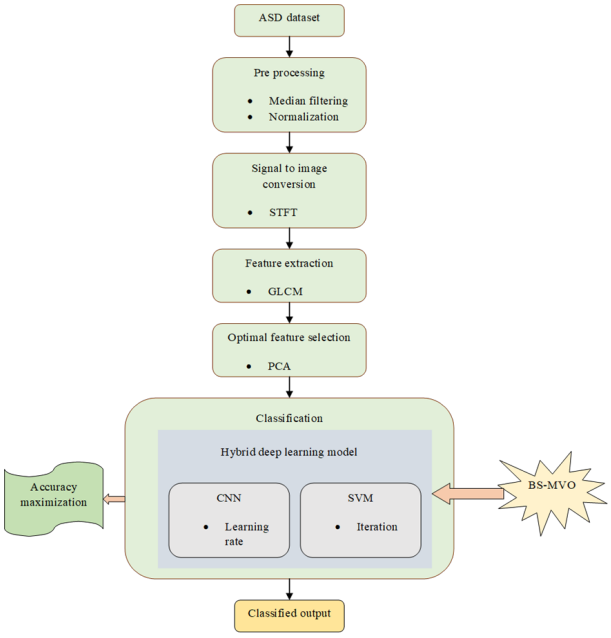

Fig. 1. Proposed architectural model for EEG signal-based ASD

B. Pre processing

The pre processing of the introduced ASD model is done by the median filtering and normalization techniques.

Median filtering: The median filterZhu and Huang (2012)describes a statistically oriented nonlinear signal processing method. The median filtering result is $h(y,z) = \text{med}\{g(y-j\cdot z-k) \mid j,k \in X\}$ where $(y,z), h(y,z)$ respectively, describes the original as well as output images. The 2D mask is$X$: The mask size is 00in which O is usually odd); the mask form is square, linear, cross, circular, and so on.

The statistical modelling of the median filter for the image having random noise is quite hard since it defines a nonlinear filter filter Chen and Xia (1994).

$$\sigma_{\text{med}}^2 = \frac{1}{4\,\text{cg}^2\,\overline{(o)}} \approx \frac{\sigma_j^2}{o+\frac{\pi}{2}-1}\cdot\frac{\pi}{\overline{2}} \tag{1}$$

Here, σ_j^2 shows the variance linked with the input noise power, oshows the size linked with the median filtering

mask, and $g(\bar{o})$ shows the noise density function. And the average filtering's noise variance is

$$\sigma_0^2 = \frac{1}{o}\sigma_j^2 \tag{2}$$

When comparing (1) and (2), the median filtering impacts are determined by two factors: the noise distribution and the mask size.

Normalization Patro and Sahu (2015) Normalization can take several forms, and minimal mixture regularization **?**Şen et al. (2014)Eslami et al. (2019)Craddock et al. (2011)describes a technique for normalizing data that applies a linear transformation to an initial range of data.

$$B' = \left(\frac{B - \text{min value of } B}{\text{max value of } B - \text{min value of } B}\right) * (E - D) + D \tag{3}$$

Where A' is the min-max normalized dataset. For the given ranges [C, D], A represents the original data range, and B represents the map data range. Z-score standardization **?**Şen et al. (2014)Eslami et al. (2019)Craddock et al. (2011)defines a technique to derive standardized values or data ranges from raw unstructured data using concepts such as standard deviation and mean. So, the unstructured data can be normalized using the Z-score parameter, as shown below.

$$w_j' = \frac{w_j - \bar{F}}{\text{std}(F)} \tag{4}$$

Here, w_j shows the normalised Z-score one value. $w_j$defines the value of the j^{th} column's row F.

$$\text{std}(F) = \sqrt{\frac{1}{(o-1)}\sum_{j=1}^{o}\left(w_j - \bar{F}\right)^2} \tag{5}$$

The mean value is shown as follows.

$$\bar{F} = \frac{1}{o}\sum_{j=1}^{o} w_j \tag{6}$$

Decimal Scaling **?**Şen et al. (2014)Eslami et al. (2019)Craddock et al. (2011) represents the approach that offers the range among -1 and 1. As a result of the decimal scaling method,

$$w^j = \frac{w}{10^k} \tag{7}$$

The scaled values are represented by w^j. The range of values is denoted by w. The smallest integer is shown by $k \cdot \text{Max}\left(\left|w^j\right|\right) < 1$.

4. SIGNAL TO IMAGE ONVERSION AND FEATURE EXTRACTION FOR AUTISM SPECTRUM DISORDER DETECTION

A. Signal to image conversion

The signal is converted to image in this step by STFT. The Fourier transform linked with overlapping windowed chunks of a signal is called the STFT Selesnick (2009). In this section, we calculate the perfect reconstruction condition for the window function, designated$\omega(0)$. For $o = 0, \cdots, O-1$ is considered to be supported (non-zero). The window in this case is a symmetric half-cycle sine window. The half-cycle sine window for Opoints is defined by

$$\omega(o) = \sin\left(\frac{\pi}{o}(o + 0.5)\right), \quad o = 0, \cdots, 0-1 \tag{8}$$

The relocated windows have been moved by half their length.$y(o) \cdot \omega(o - n \cdot 0/2)$ represents the n^{th} thwindowed block of the signal $y(o)$.$t(n, o)$ denotes the n^{th} windowed block:

$$t(n, o) := y(o) \cdot \omega(o - n \cdot 0/2) \tag{9}$$

Taking the DTFT of every windowed block yields the STFT:

$$T(n, \omega) := \text{DTFT}\{y(o) \cdot \omega(o - n \cdot o/2)\} \tag{10}$$

The STFT related with a discrete-time signal y(o)is specified by

$$T(n, \omega) = \text{STFT}\{y(o)\} \tag{11}$$

B. Feature Extraction

The features are extracted in the developed ASD model utilizing GLCM. The GLCM approach Albregtsen (2008)describes a technique for obtaining statistical texture features linked with second order. The method has been applied in a variety of situations. The count of rows and columns in a GLCM matrix equals the count of grey levels H in the image. The matrix member $Q(j, k \mid \Delta y, \Delta z)$represents the relative frequency, where two pixels spaced by a pixel distance$(\Delta y, \Delta z)$, one having intensity j and the other having intensity k, appear inside a specified neighbourhood. The second order statistical probability values for variations among grey levels $j$and $k$at a certain displacement distance e and at a specific angle (0) are contained in the matrix element onsider $g(n; o)$

be the intensity at sample$n$, line $o$of a N x o Oneighbourhood of an input image having H grey levels ranging from 0 to H-1. Hence,

$$Q(j, k \mid \Delta y, \Delta z) = XR(j, k \mid \Delta y, \Delta z) \tag{12}$$

$$X = \frac{1}{(N - \Delta y)(O - \Delta z)} \tag{13}$$

$$R(j, k \mid \Delta y, \Delta z) = \sum_{o=1}^{O-\Delta z}\sum_{n=1}^{N-\Delta y} B \tag{14}$$

$$B = \begin{cases} 1 \text{ if } g(n, o) = j \text{ and } g(n + \Delta y, o + \Delta z) = k \\ 0 \text{ elsewhere} \end{cases} \tag{15}$$

Various features that are utilized are as below.

Homogeneity, ASM: It defines a metric for image uniformity. A homogenous image will have some grey levels, resulting in a GLCM having few but greater values $Q(j, k)$. As a result, the sum of squares would be large.

$$ASM = \sum_{j=0}^{H-1}\sum_{k=0}^{H-1}\{Q(j, k)\}^2 \tag{16}$$

Contrast:$Q(j, k)$ contributions distant from the diagonal, $j \neq k$, will be favoured by this metric of contrast

or local intensity fluctuation.

Local Homogeneity, IDM: The uniformity of the image has an impact on IDM. IDM will receive minimal benefits from inhomogeneous locations $j \neq k$ due to the weighting factor $\left(1+(j-k)^2\right)^{-1}$. As a consequence, inhomogeneous images have a lower IDM value, whereas homogeneous images have a significantly greater value.

$$IDM = \sum_{j=0}^{H-1}\sum_{k=0}^{H-1}\frac{1}{1+(j-k)^2}Q(j,k) \tag{17}$$

Entropy: The entropy of inhomogeneous situations is small, whereas the entropy of homogeneous scenes is large.

$$Entropy = -\sum_{j=0}^{H-1}\sum_{k=0}^{H-1}Q(j,k)\times\log(Q(j,k)) \tag{18}$$

Correlation: The grey level linear dependency among the pixels at the given places relative to one another is measured via correlation.

$$Correlation = \sum_{j=0}^{H-1}\sum_{k=0}^{H-1}\frac{\{j\times k\}\times Q(j,k)-\{\mu_y\times\mu_z\}}{\sigma_y\times\sigma_z} \tag{19}$$

Sum of Squares, Variance: This feature gives the items that deviate from the average value of $Q(j,k)a$ lot of weight.

$$Variance = \sum_{j=0}^{H-1}\sum_{k=0}^{H-1}(j-\mu)^2Q(j,k) \tag{20}$$

Sum Average:

$$Aver = \sum_{j=0}^{2H-2}jQ_{y+z}(j) \tag{21}$$

Sum Entropy:

$$Sent = -\sum_{j=0}^{2H-2}Q_{y+z}(j)\log\left(Q_{y+z}(j)\right) \tag{22}$$

Difference Entropy:

$$Dent = -\sum_{j=0}^{H-1}Q_{y+z}(j)\log\left(Q_{y+z}(j)\right) \tag{23}$$

Inertia:

$$Inertia = \sum_{j=0}^{H-1}\sum_{k=0}^{H-1}\{j-k\}^2\times Q(j,k) \tag{24}$$

Cluster Shade:

$$Shade = \sum_{j=0}^{H-1}\sum_{k=0}^{H-1}\{j+k-\mu_y-\mu_z\}^3\times Q(j,k) \tag{25}$$

Cluster Prominence:

$$Prom = \sum_{j=0}^{H-1}\sum_{k=0}^{H-1}\{j+k-\mu_y-\mu_z\}^4\times Q(j,k) \tag{26}$$

C. Optimal feature selection

The optimal features are selected from the extracted features using PCA. PCA Puyati and Walairacht (2008) employs n×odimensional vectors to describe an image of size n×opixels and utilizes those vectors to compute the mean of vectors, which shows the average throughout all dimensions. The vector mean is removed from the entire vectors. The covariance matrix Ddefines the product matrix of the subtracted mean vector together with its transpose. It is written as follows:

$$D = Y'Y'^T \tag{27}$$

Here, Y' defines the subtracted mean data matrix, and the trained vectors are the column. The covariance matrix is used to calculate the eigenvectors as well as eigen values. For Lvector, the eigenvectors having the greatest eigenvalue are chosen. Eigenfaces describes the names given to these eigenvectors. These eigenfaces are used to create the transformation matrix in a row fashion. The below equation may be used to determine the extracted feature, z_j of the input data, y_j.

$$z_j = X\left(y_j-\bar{y}\right) \tag{28}$$

Here, X denotes the transformation matrix and y denotes the data as well as its mean value.

5. HYBRID DEEP LEARNING MODEL-BASED EEG SIGNAL PROCESSING FOR AUTISM SPECTRUM DISORDER DETECTION

A CNN

A typical CNN is made up of one or multiple convolutions as well as subsampling layers, followed by one or multiple fully connected layers, and finally an output layer.

Convolutional Layer:

The essential portion of a CNN Sultana et al. (2018) defines the convolutional layer (conv layer). In reality, images are usually static. That would be to say, the formation of one portion of the image is like the formation of any remaining part. As a result, a feature learned in one location may be used to match a pattern in another. We consider a tiny segment of a huge image and pass it via entire of the points in the large image (Input). We condense them into a single position when traveling through any location (Output). Filters are tiny sections of an image that pass across a larger image (Kernel). The back propagation approach is then used to construct the filters.

Pooling or sub-sampling layer:

Down sampling associated with an image is referred to as pooling. It sub-samples a tiny section of the convolutional output as input to create a single output. There exist several pooling strategies available, like mean pooling, maximum pooling, and average pooling. The biggest of a region's pixel values is used in max pooling. Pooling decreases the count of parameters that must be calculated while also making the network insensitive to changes in size, form, and scale. Here, the fully connected layer is replaced by SVM.

B SVM

SVMsSVMsSrivastava and Bhambhu (2005) define a collection of supervised learning techniques for classification as well as regression **?**. Data points of the type $\{(y_1,z_1),(y_2,z_2),(y_3,z_3),(y_4,z_4)\cdots\cdots\cdots,(y_0,z_o)\}$ are considered. Here,$z_o = 1/-1$ is a constant indicating the class that $y_o$corresponds to. ois the count of samples

taken. Every y_o represents a q-dimensional real vector. The scaling is necessary to avoid variables (attributes) having a lot of variation. We can see this Training data by using the dividing (or separating) hyperplane, which divides the data into two halves.

$$x \cdot y + c = 0 \tag{29}$$

Here, x shows a q-dimensional Vector and cdefines a scalar. The separating hyperplane is perpendicular to the vectorx. We may extend the margin by using the offset parameter c. The hyperplane is forced to pass via the origin in the absence of c, limiting the solution.

$$x \cdot y + c = 1 \tag{30}$$

$$x \cdot y + c = -1 \tag{31}$$

The distance among the hyperplane is$2/|x|$, as determined by geometry. As a result, we aim to keep $|x|$to a minimum. To pique the interest of data points, we must guarantee that for every jeither

$$x \cdot y_j - c \geq 1 \quad or \quad x \cdot y_j - c \leq -1 \tag{32}$$

This can be expressed as:

$$z_j \left(x \cdot y_j - c\right) \geq 1, \quad 1 \leq j \leq o \tag{33}$$

C Hybrid Deep Learning Model

The hybrid deep learning model for the introduced ASD is performed by enhanced CNN. Here, the fully connected layer of CNN is replaced by SVM thus called as enhanced CNN. The major objective of hybrid deep learning model is to optimize the learning rate of CNN as well as iteration of SVM with the intention of accuracy maximization. This is shown as below.

$$fit = \underbrace{\text{argmax}}_{\{LR_{CNN}, IT_{SVM}\}} (Accuracy) \tag{34}$$

Here, the fitness or the objective function is revealed byfit, learning rate of CNN is given by LR_{CNN}, iteration of SVM is given byIT_{SVM}, and accuracy is shown byAccuracy respectively. Accuracy is formulated as below.

$$Accuracy = \frac{T^{PE} + T^{NE}}{T^{PE} + T^{NE} + F^{PE} + F^{NE}} \tag{35}$$

Here, the true positive, true negative, false positive, and false negative is shown by T^{PE}, T^{NE}, F^{PE},and F^{NE} respectively.

D Proposed Optimized BS-MVO

The proposed BS-MVO uses a hybrid deep learning model to improve the classification status of the introduced ASD model. It optimizes the learning rate of ANN and iteration of SVM with the intention of accuracy maximization. MVO Mirjalili et al. (2015) algorithm's major influences are three cosmological concepts: white hole, black hole, and wormhole. These three notions have statistical formulas that are used to execute exploitation, exploration, and local search, accordingly. The MVO has various advantages such as better exploitation, exploration, and convergence. But, it limits from binary as well as multi objective versions. Hence, to overcome its shortcomings, BSO is integrated into it and the resulting algorithm is called as BS-MVO.

BSOChen et al. (2018), which is dependent on the BAS Jiang and Li (2017) and the conventional PSO Eberhart and Kennedy (1995), is drawn from PSO, but the update criteria for every particle come from BAS.

Hence, BS-MVO is designed on the basis of random concept. If$rand \geq 0.5$,then the update takes place by BSO as below.

$$wc_j = -\delta^u \cdot \boldsymbol{c} \cdot \text{sign}\left(g\left(y_{\text{right}}\right) - g\left(y_{\text{left}}\right)\right) \tag{36}$$

Here, the update rate is shown by wc_j, particle is shown by j, fitness function value of left and right side is shown by $g\left(y_{left}\right)$and $g\left(y_{right}\right)$, iteration is shown by u, step size is shown by$delta$, and the symbolic function is shown by sign respectively.

Otherwise, if $rand < 0.5$, then the update takes place by MVO as in Eq. (37).

$$wc_u^j = \begin{cases} WC_j + TDR \times \left((ud_j - ld_j) \times q4 + ld_j\right) & q3 < 0.5 \\ & q2 < WEP \\ wc_u^j WC_j - TDR \times \left((ud_j - ld_j) \times q4 + ld_j\right) & q3 \geq 0.5 \\ & q2 \geq WEP \end{cases} \tag{37}$$

Here, WC_j defines the j^{th} parameter of best universe, W E P and T D R are coefficients, upper and lower bound is revealed by ud_j, ld_j, random numbers are $q2, q3$, and $q4, j^{\text{th}}$ parameter of u^{th} universe is depicted by wc_u^j respectively. The BS-MVO pseudo code is in Algorithm 1. The flowchart of proposed BS-MVO is shown in Figure 2

Algorithm 1: BS-MVO

Start
Parameter and population initialization
Fitness calculation
if rand ≥ 0.5
$wc_j = -\delta^u \cdot \boldsymbol{c} \cdot \text{sign}\left(g\left(y_{\text{right}}\right) - g\left(y_{\text{left}}\right)\right)$
else

$$wc_u^j = \begin{cases} WC_j + TDR \times \left((ud_j - ld_j) \times q4 + ld_j\right) & q3 < 0.5 \\ & q2 < WEP \\ wc_u^j WC_j - TDR \times \left((ud_j - ld_j) \times q4 + ld_j\right) & q3 \geq 0.5 \\ & q2 \geq WEP \end{cases}$$

Return best solution
Stop

6. RESULTS

A Experimental Setup

The proposed EEG signal classification-based ASD was implemented and analyzed. The dataset regarding the proposed work is gathered from https://www.kaggle.com/datasets/disbeat/bciaut-p300/data.The suggested ASD model was compared with different heuristic-based algorithms like GA McCall (2005), WOA Mirjalili and Lewis (2016), MVO Mirjalili et al. (2015), and BSO Chen et al. (2018) concerning analysis such as accuracy, sensitivity, TPR, specificity, and MCC to describe the betterment of the introduced method.

B Accuracy Analysis

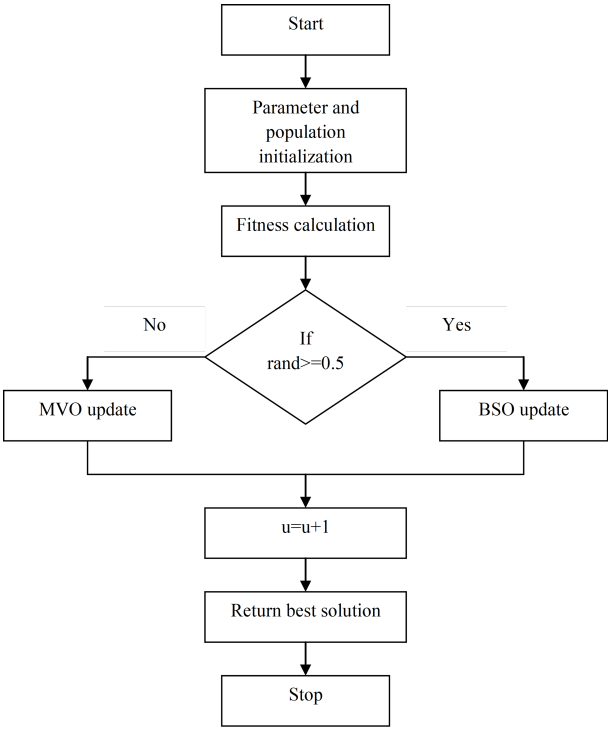

Fig. 2. Flowchart of proposed BS-MVO

The accuracy analysis Sekar et al. (2021) of different heuristic-based algorithms for the introduced ASD model is shown here. It can be seen that BS-MVO outperforms other methods at various learning percentages. Hence, it can be revealed clearly that the proposed BS-MVO is better than the other methods for the developed ASD model.

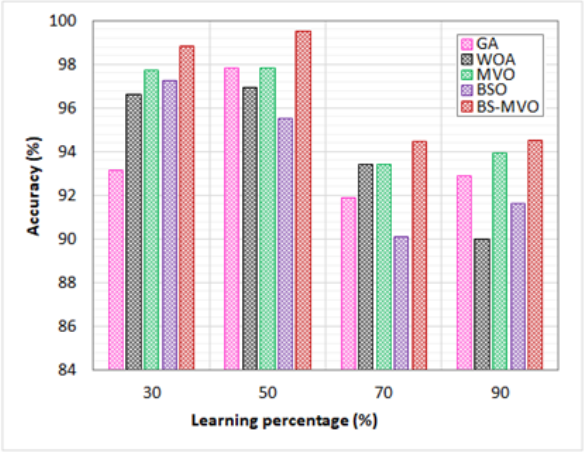

Fig. 3. Accuracy analysis

C Sensitivity Analysis

The sensitivity Jayachitra and Prasanth (2020) of several heuristic-oriented methods for the proposed ASD model is examined in this paper. At various learning percentages, it can be demonstrated that BS-MVO beats other approaches. As a result, it is obvious that the suggested BS-MVO is superior to the other approaches in terms of the created ASD model for sensitivity.

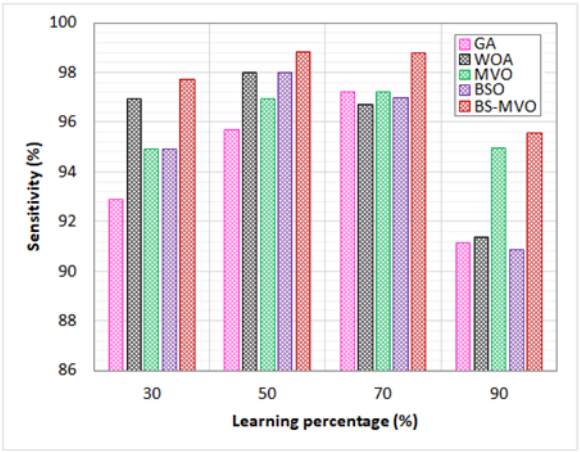

Fig. 4. Sensitivity Analysis

D TPR Analysis

This research investigates the TPR of multiple heuristic-oriented strategies for the suggested ASD model. It can be proved that BS-MVO outperforms other techniques at various learning percentages. As a consequence, the proposed BS-MVO is clearly superior to the other techniques in terms of the ASD model for TPR generated.

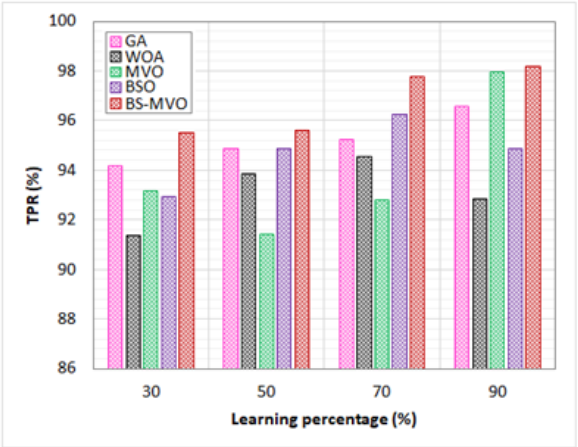

Fig. 5. TPR Analysis

E Specificity Analysis

The specificity Kumar and Boda (2021) of several heuristic-oriented techniques for the proposed ASD model is investigated in this study. At various learning percentages, it can be demonstrated that BS-MVO outperforms other approaches. As a result, in terms of the ASD model for specificity produced, the suggested BS-MVO obviously outperforms the other strategies.

F MCC Analysis This study investigates the MCC of numerous heuristic-oriented strategies for the suggested ASD model. It has been established that BS-MVO outperforms other techniques at various learning percentages. As a consequence, the recommended BS-MVO clearly beats the other techniques in terms of the ASD model for MCC created.

Table 1. ACCURACY ANALYSIS

Methods	Learning 30%	Learning 50%	Learning 70%	Learning 90%
GA McCall (2005)	93.16	97.85	91.87	92.90
WOA Mirjalili and Lewis (2016)	96.62	96.96	93.40	89.99
MVO Mirjalili et al. (2015)	97.73	97.84	93.43	93.91
BSO Chen et al. (2018)	97.24	95.52	90.10	91.60
BS-MVO	98.81	99.53	94.47	94.49

Table 2. SENSITIVITY ANALYSIS

Methods	Learning 30%	Learning 50%	Learning 70%	Learning 90%
GA McCall (2005)	92.89	95.71	97.21	91.12
WOA Mirjalili and Lewis (2016)	96.96	97.98	96.70	91.39
MVO Mirjalili et al. (2015)	94.93	96.95	97.23	94.95
BSO Chen et al. (2018)	94.90	97.98	96.97	90.85
Proposed Optimized BS-MVO	97.72	98.81	98.80	95.56

Table 3. TPR ANALYSIS

Methods	Learning 30%	Learning 50%	Learning 70%	Learning 90%
GA McCall (2005)	94.18	94.86	95.23	96.59
WOA Mirjalili and Lewis (2016)	91.36	93.84	94.55	92.83
MVO Mirjalili et al. (2015)	93.18	91.43	92.80	97.95
BSO Chen et al. (2018)	92.93	94.88	96.25	94.88
BS-MVO	95.51	95.59	97.77	98.18

Table 4. SPECIFICITY ANALYSIS

Methods	Learning 30%	Learning 50%	Learning 70%	Learning 90%
GA McCall (2005)	90.83	93.65	91.12	90.40
WOA Mirjalili and Lewis (2016)	93.39	93.89	92.64	92.18
MVO Mirjalili et al. (2015)	92.37	95.94	94.17	93.19
BSO Chen et al. (2018)	93.94	90.12	95.46	91.43
BS-MVO	94.41	96.67	96.12	95.11

Table 5. MCC ANALYSIS

Methods	Learning 30%	Learning 50%	Learning 70%	Learning 90%
GA McCall (2005)	91.78	97.62	91.80	92.40
WOA Mirjalili and Lewis (2016)	96.95	97.63	93.86	92.47
MVO Mirjalili et al. (2015)	96.97	97.64	90.81	91.46
BSO Chen et al. (2018)	97.65	98.65	94.22	93.50
BS-MVO	98.80	98.90	95.14	94.99

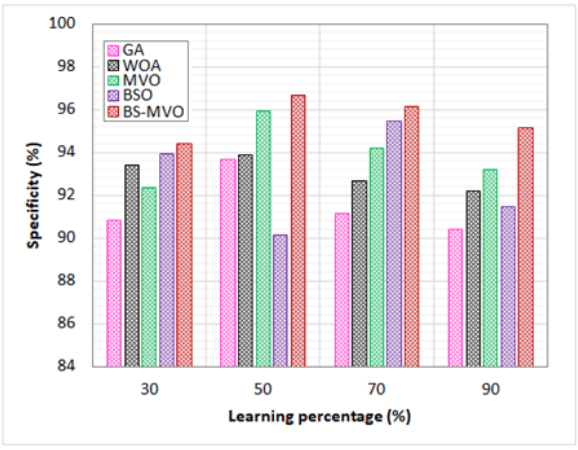

Fig. 6. Specificity Analysis

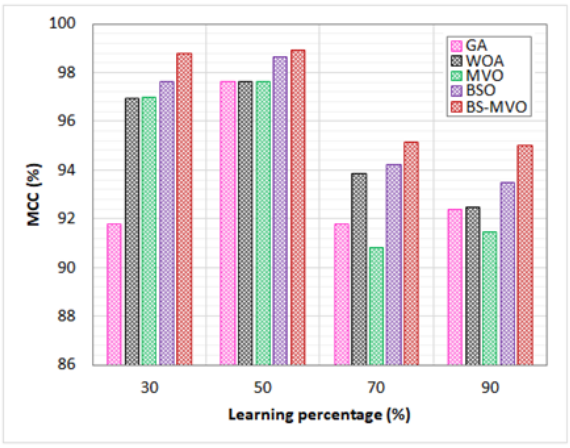

Fig. 7. MCC Analysis

7. CONCLUSION

This study aimed to use EEG signal processing and analysis to detect ASD. The ASD dataset was used to create the dataset. The pre-processing was then completed using median filtering and normalization. Then, the STFT technique was used to convert the signals to images. The GLCM was employed to extract features from altered images. PCA was used to choose relevant characteristics from the recovered features since they are lengthy. Based on these carefully selected charac-

teristics, a hybrid deep learning model was used to categorize the data (ASD or not). In this scenario, the fully connected layer of CNN was replaced with SVM. By combining BSO with MVO, the novel BS-MVO optimized CNN learning rate and SVM iteration with the goal of accuracy maximization as the fitness or objective function, resulting in enhanced CNN. This upgraded CNN was used to classify the final output. The proposed method achieved results with an MCC outcome of 94.99%, a specificity of 95.11%, and an accuracy of 94.49%. Simulation results showed that it was superior to established techniques on various criteria. However, the proposed method has dimension limitations in the collective dataset, so it is highly complex. In the future, ensemble learning models will be used to reduce the computational complexity to detect ASD with a low false rate.

REFERENCES

Abdolzadegan, D., Moattar, M.H., and Ghoshuni, M. (2020). A robust method for early diagnosis of autism spectrum disorder from eeg signals based on feature selection and dbscan method. *Journal of Applied Biomedicine*, 40(1), 482–493. doi: 10.1016/j.bbe.2020.01.008.

Ahmadlou, M., Adeli, H., and Adeli, A. (2012a). Fuzzy synchronization likelihood-wavelet methodology for diagnosis of autism spectrum disorder. *Journal of Neuroscience Methods*, 211(2), 203–209. doi: 10.1016/j.jneumeth.2012.08.020.

Ahmadlou, M., Adeli, H., and Adeli, A. (2012b). Improved visibility graph fractality with application for the diagnosis of autism spectrum disorder. *Physica A: Statistical Mechanics and its Applications*, 391(20), 4720–4726. doi:10.1016/j.physa.2012.04.025.

Albregtsen, F. (2008). Statistical texture measures computed from gray level co-occurrence matrices. Technical report. Technical report or lecture note.

Alhaddad, M.J., Kamel, M.I., Malibary, H.M., Alsaggaf, E.A., Thabit, K., Dahlwi, F., Hadi, A.A., and KAU, K.A.U. (2012). Diagnosis autism by fisher linear discriminant analysis (flda) via eeg. *International Journal of Bio-Science and Bio-Technology*, 4(2), 45–47.

Alsaggaf, E.A. and Kamel, M.I. (2014). Using eegs to diagnose autism disorder by classification algorithm. *Life Science Journal*, 11(6), 305–308. Department of Computer Science, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia.

Alturki, F.A., Aljalal, M., Abdurraqeeb, A.M., Alsharabi, K., and Al-Shamma'a, A.A. (2021). Common spatial pattern technique with eeg signals for diagnosis of autism and epilepsy disorders. *IEEE Access*, 9, 24334–24349. doi:10.1109/ACCESS.2021.3056619.

Billeci, L., Tonacci, A., Tartarisco, G., Narzisi, A., Di Palma, S., Corda, D., Baldus, G., Cruciani, F., Anzalone, S.M., Calderoni, S., Pioggia, G., and Muratori, F. (2016). An integrated approach for the monitoring of brain and autonomic response of children with autism spectrum disorders during treatment by wearable technologies. *Frontiers in Neuroscience*, 10, 276. doi:10.3389/fnins.2016.00276.

Bosl, W., Tierney, A., Tager-Flusberg, H., and Nelson, C. (2011). Eeg complexity as a biomarker for autism spectrum disorder risk. *BMC Medicine*, 9(1), 18. doi: 10.1186/1741-7015-9-18.

Capal, J.K., Carosella, C., Corbin, E., Horn, P.S., Caine, R., and Manning-Courtney, P. (2018). Eeg endophenotypes in autism spectrum disorder. *Epilepsy & Behavior*, 88, 341–348. doi:10.1016/j.yebeh.2018.09.036.

Chen, T., Zhu, Y., and Teng, J. (2018). Beetle swarm optimization for solving investment portfolio problems. *The Journal of Engineering*, 2018(16), 1600–1605.

Chen, T. and Xia, L. (1994). *Digital Image Processing*. Posts & Telecommunications Press, Beijing.

Cociu, B.A., Das, S., Billeci, L., Jamal, W., Maharatna, K., Calderoni, S., Narzisi, A., and Muratori, F. (2017). Multimodal functional and structural brain connectivity analysis in autism: A preliminary integrated approach with eeg, fmri, and dti. *IEEE Transactions on Cognitive and Developmental Systems*, 10(2), 213–226. doi: 10.1109/TCDS.2017.2680408.

Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., and Mayberg, H.S. (2011). A whole brain fmri atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, 33(8), 1914–1928. doi: 10.1002/hbm.21333.

Şen, B., Peker, M., Çavuşoğlu, A., and Çelebi, F.V. (2014). A comparative study on classification of sleep stage based on eeg signals using feature selection and classification algorithms. *Journal of Medical Systems*, 38(3), 18. doi:10.1007/s10916-014-0018-0.

Dawson, G., Jones, E.J., Merkle, K., Venema, K., Lowy, R., Faja, S., Kamara, D., Murias, M., Greenson, J., Winter, J., Smith, M., Rogers, S.J., and Webb, S.J. (2012). Early behavioral intervention is associated with normalized brain activity in young children with autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(11), 1150–1159. doi: 10.1016/j.jaac.2012.08.018.

Demuru, M. and Fraschini, M. (2020). Eeg fingerprinting: Subject-specific signature based on the aperiodic component of power spectrum. *Computers in Biology and Medicine*, 120, 103748. doi: 10.1016/j.compbiomed.2020.103748.

Eberhart, R. and Kennedy, J. (1995). A new optimizer using particle swarm theory. In *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, 39–43. IEEE.

Eslami, T., Mirjalili, V., Fong, A., Laird, A.R., and Saeed, F. (2019). Asd-diagnet: A hybrid learning approach for detection of autism spectrum disorder using fmri data. *Frontiers in Neuroinformatics*, 13, 70. doi: 10.3389/fninf.2019.00070.

Hadoush, H., Alafeef, M., and Abdulhay, E. (2019). Automated identification for autism severity level: Eeg analysis using empirical mode decomposition and second order difference plot. *Behavioural Brain Research*, 362, 240–248. doi:10.1016/j.bbr.2019.01.018.

Ibrahim, S., Djemal, R., and Alsuwailem, A. (2017). Electroencephalography (eeg) signal processing for epilepsy and autism spectrum disorder diagnosis. *Journal of Applied Biomedicine*, 38(1), 16–26. doi: 10.1016/j.bbe.2017.08.006.

Jamal, W., Das, S., Oprescu, I., Maharatna, K., Apicella, F., and Sicca, F. (2014). Classification of autism spectrum disorder using supervised learning of brain connectivity measures extracted from synchrostates. *Journal of Neural Engineering*, 11(4), 046019. doi:10.1088/1741-2560/11/4/046019.

Jayachitra, S. and Prasanth, A. (2020). Multi-feature analysis for automated brain stroke classification using weighted gaussian naïve bayes classifier. *Journal of Circuits, Systems and Computers*, 30(10), 2150178. doi:10.1142/S0218126621501784.

Jiang, X. and Li, S. (2017). Bas: Beetle antennae search algorithm for optimization problems. Technical report. Unpublished manuscript or technical report.

Kana, R.K., Libero, L.E., and Moore, M.S. (2011). Disrupted cortical connectivity theory as an explanatory model for autism spectrum disorders. *Physics of Life Reviews*, 8(4), 410–437. doi:10.1016/j.plrev.2011.10.001.

Kang, J., Han, X., Song, J., Niu, Z., and Li, X. (2020). The identification of children with autism spectrum disorder by svm approach on eeg and eye-tracking data. *Computers in Biology and Medicine*, 120, 103722. doi:10.1016/j.compbiomed.2020.103722.

Kumar, K.A. and Boda, R. (2021). A multi-objective randomly updated beetle swarm and multi-verse optimization for brain tumor segmentation and classification. *The Computer Journal*, 65(4), 1029–1052. doi:10.1093/comjnl/bxab171.

Lavanga, M., De Ridder, J., Kotulska, K., Moavero, R., Curatolo, P., Weschke, B., Riney, K., Feucht, M., Krsek, P., Nabbout, R., Jansen, A.C., Wojdan, K., Domanska-Pakieła, D., Kaczorowska-Frontczak, M., Hertzberg, C., Ferrier, C.H., Samueli, S., Jahodova, A., Aronica, E., and Caicedo, A. (2021). Results of quantitative eeg analysis are associated with autism spectrum disorder and development abnormalities in infants with tuberous sclerosis complex. *Biomedical Signal Processing and Control*, 68, 102658. doi:10.1016/j.bspc.2021.102658.

Matthew, R.P., Mica, E.J., Meinhold, W., Loeza, J.A., Tomizuka, M., and Bajcsy, R. (2015). Initial investigation into the effect of an active/passive exoskeleton on hammer curl performance in healthy subjects. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 3607–3610. IEEE. doi:10.1109/EMBC.2015.7319173.

McCall, J. (2005). Genetic algorithms for modelling and optimisation. *Journal of Computational and Applied Mathematics*, 184(1), 205–222. doi:10.1016/j.cam.2004.07.034.

Mirjalili, S. and Lewis, A. (2016). The whale optimization algorithm. *Advances in Engineering Software*, 95, 51–67. doi:10.1016/j.advengsoft.2016.01.008.

Mirjalili, S., Mirjalili, S.M., and Hatamlou, A. (2015). Multi-verse optimizer: A nature-inspired algorithm for global optimization. *Neural Computing and Applications*, 27(2), 495–513. doi:10.1007/s00521-015-1870-7.

Patro, S.G.K. and Sahu, K.K. (2015). Normalization: A preprocessing stage. *International Advanced Research Journal in Science, Engineering and Technology (IARJSET)*, 2(3), 20–22. doi:10.17148/iarjset.2015.2305.

Puyati, W. and Walairacht, A. (2008). Efficiency improvement for unconstrained face recognition by weightening probability values of modular pca and wavelet pca. In *2008 International Conference on Advanced Communication Technology (ICACT)*, 1449–1453. IEEE. doi:10.1109/ICACT.2008.4494037.

Sekar, J., Aruchamy, P., Abdul, H.S.L., Mohammed, A.S., and Khamuruddeen, S. (2021). An efficient clinical support system for heart disease prediction using tanfis classifier. *Computational Intelligence*, 38(2), 610–640. doi:10.1111/coin.12487.

Selesnick, I.W. (2009). Short-time fourier transform and its inverse. Technical report. Unpublished lecture notes or technical report.

Sheikhani, A., Behnam, H., Mohammadi, M.R., Noroozian, M., and Golabi, P. (2008). Connectivity analysis of quantitative electroencephalogram background activity in autism disorders with short time fourier transform and coherence values. In *Proceedings of the Congress on Image and Signal Processing (CISP)*, 207–212. IEEE. doi:10.1109/CISP.2008.595.

Sheikhani, A., Behnam, H., Mohammadi, M.R., Noroozian, M., and Mohammadi, M. (2010). Detection of abnormalities for diagnosing of children with autism disorders using of quantitative electroencephalography analysis. *Journal of Medical Systems*, 36(2), 957–963. doi:10.1007/s10916-010-9560-6.

Shillington, A. and Capal, J.K. (2020). Genetic testing in patients with nonsyndromic autism spectrum disorder and eeg abnormalities with or without epilepsy: Is exome trio-based testing the best clinical approach? *Epilepsy & Behavior*, 114, 107564. doi:10.1016/j.yebeh.2020.107564.

Srivastava, D.K. and Bhambhu, L. (2005). Data classification using support vector machine. *Journal of Theoretical and Applied Information Technology*, 12(1). URL `https://www.jatit.org/volumes/Vol12No1/1Vol12No1.pdf`.

Sultana, F., Sufian, A., and Dutta, P. (2018). Advancements in image classification using convolutional neural network. In *2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, 122–129. IEEE. doi:10.1109/ICRCICN.2018.8718718.

Vettori, S., Dzhelyova, M., Van Der Donck, S., Jacques, C., Van Wesemael, T., Steyaert, J., Rossion, B., and Boets, B. (2019). Combined frequency-tagging eeg and eye tracking reveal reduced social bias in boys with autism spectrum disorder. *Cortex*, 125, 135–148. doi:10.1016/j.cortex.2019.12.013.

Vissers, M.E., Cohen, M.X., and Geurts, H.M. (2011). Brain connectivity and high functioning autism: A promising path of research that needs refined models, methodological convergence, and stronger behavioral links. *Neuroscience & Biobehavioral Reviews*, 36(1), 604–625. doi:10.1016/j.neubiorev.2011.09.003.

Zhu, Y. and Huang, C. (2012). An improved median filtering algorithm for image noise reduction. *Physics Procedia*, 25, 609–616. doi:10.1016/j.phpro.2012.03.133.

Abbreviation	Description
GLCM	Grey Level Co-occurrence Matrix
ASD	Autism Spectrum Disorder
GA	Genetic Algorithm
STFT	Short Time Fourier Transform
MCC	Matthews Correlation Coefficient
BS-MVO	Beetle Swarm-based Multi Verse Optimization
DTI	Diffusion Tensor Imaging
EEG	Electro Encephalo Gram
DSM	Diagnostic and Statistical Manual
CARS	Childhood Autism Rating Scale
BSO	Beetle Swarm Optimization
PCA	Principal Component Analysis
fMRI	functional Magnetic Resonance Imaging
ATEC	Autism Treatment Evaluation Checklist
ICA	Independent Component Analysis
CSP	Common Spatial Pattern
MVO	Multi Verse Optimization
EMD	Empirical Mode Decomposition
LDA	Linear Discriminant Analysis
dDTF	direct Directed Transfer Function
CCHMC	Cincinnati Children's Hospital Medical Center
SODP	Second-Order Difference Plotting
KNN	k-Nearest Neighbour
TSC	Tuberous Sclerosis Complex
MFA	Multi FrActality
TD	Typically Developing
ANN	Artificial Neural Network
MSE	Multi Scale Entropy
IMFs	Intrinsic Mode Functions
MSEL	Mullen Scales of Early Learning
FFT	Fast Fourier Transform
PSO	Particle Swarm Optimization
AOI	Areas Of Interest
CTM	Central Tendency Measure
WOA	Whale Optimization Algorithm
mRMR	minimum-Redundancy Maximum-Relevancy
ASM	Angular Second moment
CNN	Convolutional Neural Network
VABS	Vineland Adaptive Behavior Scales
MI	Mutual Information
IDM	Inverse Difference Moment
SRM	Structural Risk Minimization
PLS	Preschool Language Scales
SVM	Support Vector Machine
IG	Information Gain
BAS	Beetle Antennae Search
TPR	True Positive Rate

Table 6. Nomenclature