Hybrid Feature Vector for Screening of Autistic People using Deep Learning

Suruchi Dedgaonkar[1]*   Rajneeshkaur Sachdeo[2]   Suneel Godbole[3]

*[1]MIT School of Engineering, MIT-ADT University
and Faculty at Vishwakarma Institute of Information Technology, Pune, India
[2]MIT School of Engineering, MIT-ADT University, Pune, India
[3]Developmental Pediatrician, Chiranjeev Clinic, Pune, India*
* Corresponding author's Email: suruchi.dedgaonkar@viit.ac.in

Abstract: Screening tools play a vital role in Autism Spectrum Disorder (ASD) and in the detection of related Sensory Processing Disorder. Since the plan of relevant therapy depends on early and accurate detection of these disorders, getting the appropriate answers to the questionnaire in a timely manner is of utmost importance. The automatic collection of the features related to behavioural parameters and the response to given stimuli designed by experts is possible with the recent technology. The proposed novel feature vector makes use of the health parameters obtained by wearable devices and emotion histogram based on facial expressions along with the inputs from the reduced questionnaire. The deep learning technique is used to create a model, which will assist the behavioural therapist to overcome problems of the current screening tools. For the purpose of experimentation, we have used Autism Spectrum Quotient AQ10 dataset using fully connected Deep Neural Network (DNN) to train the autism detection model and Facial Expression Recognition FER2013 dataset to train the emotion recognition model using Convolutional Neural Network (CNN). The sensory processing disorder detection (SPD) model is trained with the novel feature vector and performance of various classifiers is compared. The classification accuracy is 100% for the ASD detection, 86% for emotion recognition and 100% for SPD detection.

Keywords: Autism, Screening, Sensory processing disorder, Deep learning, CNN, Emotion.

1. Introduction

Autism Spectrum Disorder (ASD) is a neurological condition commonly known as Autism, which is detected in the early years of one's life and lasts throughout the individual's life. Every autistic child has different symptoms, health issues, and behavior. Indifference to pain, aversion to specific noises or textures, and lack of visual interest are all behavioral symptoms of ASD. Many autistic children suffer from Sensory Processing Disorder (SPD). There is no medical or pathological test available to detect autism and related disorders. [1, 2]

The detection of ASD and SPD is done with the help of screening tools. Screening tools are standard questionnaires used for the detection and assessment of ASD and related disorders. On-time diagnosis of the associated sensory disorders will help in the initial suitable therapy sessions, which will help in reducing

complications. [3] Routine screening is important for monitoring the development of children with autism. High-quality screening tools are costly and also need time and trained staff for obtaining accurate results. [4, 5]

SPD is a condition in which the brain is concerned about getting and answering the stimuli or information that comes in through the senses. Stimuli are used to excite the person, and the corresponding response is observed. A child's aberrant reaction to sensory inputs is connected to poor academic performance and certain behavioural and emotional characteristics. [6] Sensory processing problems can lead to difficulties with activity participation, academic achievement, and social participation among children with ASD. Sensory processing problems are categorized into sensory over-responsivity and sensory under-responsivity as compared to typical or expected, leading to the

children's physical and emotional health problems. [7, 8]

Many Asian and African countries have different cultures and socio-economic backgrounds where people have little knowledge of ASD and its symptoms. In the low and middle-income countries, easy approaches with less cost are required to detect disorders. The development of children is fast and overlaps the symptoms, making detection of ASD and SPD difficult. [9] For ethical, therapeutic and economic reasons, a thorough examination of the viability of sensory-based therapy for children with ASD is required.

The frameworks for sensory disorder diagnosis are not universally accepted and there is no claim for its accuracy. A clear scoring criterion is not available. The parents need to continuously observe and note their children's behavior and regularly communicate their observations to the therapists. The therapist must know and should have the expertise to relate these observations which proves beneficial in deciding the proper plan for therapy. [10]

The objectives of this paper are:

1. Detection of autism on AQ10 dataset using deep learning

2. Proposal of a new classification model for sensory processing disorder detection with the novel and hybrid feature vector consisting of the emotion histogram based on facial expression, sensor parameters related to stress and reduced sensory profile questionnaire answers.

The paper is structured as follows: in Section 2, various clinical tools for screening autism and sensory processing disorder are summarized. Section 3 explains the modelling techniques and the data used for experimentation purpose. In section 4, the proposed system is introduced to detect auditory and visual SPD. The experiments done are described in section 5 and the findings are outlined in section 6. Section 7 summarizes the conclusion and further research scope and the various issues and challenges faced during the research.

2. Related work

Autism Diagnostic Observation Schedule (ADOS) is one of the most widely used tools for behavioural diagnosis. ADOS is a time-consuming behavioral test and questionnaire. Both parents and physicians would devote a considerable amount of time to behavioral exams and questionnaires. The authors developed classifiers using machine learning from two modules of the ADOS, one of the most widely used behavioral experiments. Their findings reinforce the idea that using machine learning software to test fewer symptoms will achieve high levels of precision in autism risk prediction. The authors removed 8 parameters from ADOS and applied Machine Learning (ML) techniques. The findings have prompted efforts to enhance screening-based tools for ASD diagnosis and online health interventions so that patients can receive treatment sooner than is now possible. [3, 5, 11] According to scientists, the ADTree performs almost perfectly in ASD detection, with a data-driven approach and ML algorithms on a reduced range of questions from ADOS. [12]

In another study, Q-CHAT-10 and AQ-10 datasets were used for ASD classification using various ML classifiers. The classification performance for Support Vector Machine (SVM) on Q-CHAT-10 and Adaboost on adult AQ-10 was best. [13]

For SPD diagnostic and therapeutic planning, the sensory profile is used to assess the child's sensory processing pattern in everyday scenarios and to highlight the sensory system's influence on functional performance based on sensory reactions and actions. The frequency, severity, and patterns of the behavior are focused on this diagnosis. Children are asked different questions, and behavioral observations are noted to complete the sensory profile. This can be achieved in several sessions. By incorporating touch sensitivity, taste/smell sensitivity, motion sensitivity, vision, auditory sensitivity, etc., the sensory over-responsivity score (SOR) is derived. [2]

Sensory processing disorder detection using a short sensory profile (SSP) is the shorter version of the sensory profile. [14] A lower SOR score indicates more significant difficulties on SSP, i.e., greater sensory over-responsivity. Therapists use a variety of approaches to improve constructive or reduce poor behavior and track strengths and weaknesses on a regular basis. It will be evident whether the child is developing in this manner.

With recent technologies, it is now possible to automatically capture features related to behavioral parameters and responses to given stimuli. The motor impairments of autistic people can also be used to differentiate between ASD and non-ASD. Participants were instructed to watch and then repeat a series of hand gestures. The activity was monitored with the motion sensors to record speed, trajectory and amplitude. 20 kinematic parameters were retrieved which were used in statistical histogram analysis and classification using SVM and Naïve bias. [15]

Table 1. Summary of methods used for autism detection and SPD detection

Authors	Method	Description	Performance
T.Akter, M.Satu, M.Khan, M.Ali, S.Uddin, P.Lio, J.Quinn, and M.Moni, 2019 [13]	Autism detection on AQ10 data using Adaboost, FDA, C5.0, LDA, MDA, PDA, SVM and CART	Adaboost is identified as the best classifier for child and adult AQ10 dataset	Accuracy: 97.20%
D. Wall, J.Kosmicki, T.DeLuca, E.Harstad, and V.Fusaro , 2012 [3]	Autism detection using ADTree algorithm on revised ADOS	ADTree algorithm , Removed 8 questions from ADOS module1	Accuracy: 99%
B.Li, A.Sharma, J.Meng, S.Purushwalkam, and E.Gowen , 2017 [15]	Adult autism identification based on 20 kinematic characteristics in movement pattern using Naïve bias and SVM	stereotyped , repetitive motor movement captured and ML used	NB: Accuracy: 80%, SVM: accuracy : 86.7%,
C.Busso, Z.Deng, Z.Yildirim, M.Bulut, C.Lee, A.Kazemzadeh, S.Lee, U.Neumann and S.Narayanan, 2014 [16]	EEG monitoring + sensory simulation and Discrete Wavelet Transform (DWT)	EEG associated to physical tasks. Mood change is monitored for ASD.	Cost of hardware and neural functioning knowledge
S.Tomchek, and W.Dunn, 2007 [14]	Sensory Profile and Short Sensory Profile questionnaires used.	Sensory over-responsivity score (SOR) is derived by adding- Tactile, Taste/Smell, Movement, Visual/Auditory sensitivity.	Manual method. Depends completely on observations done by parents. It is time consuming.
S.Dedgaonkar, R.Sachdeo, and S.Godbole, 2021	Autism detection using DNN on AQ10 and SPD using Novel feature vector (Reduced 5 questions based on SSP+ PMF for emotions on stimuli+ Body temperature & heart rate)	Performance of ML algorithms is compared with DNN.	Autism detection-DNN accuracy: 1.0 and SPD detection, decision tree accuracy :1.0

Researchers have collected and monitored electroencephalography (EEG) signals associated with the particular activity. EEG is a non-invasive and safe method to record brain's electrical activity. [16] Table 1 summarizes some of the methods used for autism detection and SPD detection.

Issues such as social deficits and behavioral inflexibility are considered to be linked to unusual sensory sensitivity. The majority of assessment instruments used by therapists in India are manual and based on a static questionnaire. The cost of many screening sessions is prohibitively expensive for most Indian parents. Since children develop at a high rate, early diagnosis and treatment of these conditions are critical as the problems can worsen day-by-day if they are not addressed. The available diagnostic processes are not cost-effective and are time-consuming. Computational modelling of psychopathologies is one of the possible important actors in addressing heterogeneity and non-specificity and better understanding these diseases' cognitive processes.

The present study describes the auditory and visual sensory processing condition in a chosen embodiment with a reduced questionnaire using automated features gathered with sensors and facial expressions of the child. The information on the human face also proves to be an important attribute for emotion recognition. Therefore, the number of assessment hours, resources, and efforts by clinical practitioners focus on the research. [17-19]

3. Datasets

3.1 AQ10 dataset

AQ10 dataset follows the short questionnaire by National Institute for Health Science (NHS). The AQ10 adult dataset contains 704 instances, and the AQ10 child dataset includes 292 cases with 21 attributes each. This dataset is used to train the autism

Table 2. Significance of health parameters

Health parameter	Typical child	Over-responsive	Under-responsive
Body temperature (°F)	96.4 to 100.4	94.1 to 96.3	100.5 to 104
Heart rate (bpm)	81 to 100	50 to 80	121 to 220

detection module. Likewise, the AQ10 data for ten children are taken from the behavioral clinic for testing purposes. The questions are answered as TRUE/ FALSE, i.e., 1/0.

Little noises, focus on small visuals, tracking dialogues, activity sequencing, social or peer communication, discerning intentions, predicting emotions, and making friends are among issues included in the AQ10 children dataset. [12, 20]

3.2 Body sensors

The stress-related health parameters, heart rate, and body temperature are collected during the session while audio and visual stimuli are portrayed. A smart watch is used for monitoring the heart beat and infrared thermometer to capture the body temperature. Table 2 shows the significance of stress-related health parameters. [7, 21, 22] ASD is characterized by hyper- and hypo-reactivity to sensory stimuli including sensory-dominated i.e. over-responsive and sensory-ignoring i.e. under-responsive behaviours. [15]

3.3 FER 2013 dataset

The facial expression recognition dataset comprises 35,887 grayscale 48*48 pixel pictures of faces. Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral are the seven categories. There are 28,709 instances in the training set. The 'emotion' column includes a number code for the emotion in the image, ranging from 0 to 6. For each image, the 'pixels' column contains a string enclosed by quotes. The contents of this string are pixel values in row-major order, separated by spaces. The aim is to predict the emotion column. [17]

Following are the emotion labels in the FER2013 face dataset: 0: 'Angry'; 1: 'Disgust'; 2: 'Fear'; 3: 'Happy'; 4: 'Sad'; 5: 'Surprise'; 6: 'Neutral' images.

3.4 Reduced short sensory profile (RSSP)

The sensory profile by Winnie Dunn (1999) contains overall 125 questions from which 17 questions are related to auditory and visual processing. The short sensory profile (SSP) [14] includes 38 questions, out of which 11 are connected

to auditory and visual processing. Our research concentrates only on auditory and visual sensory processing, and we have reduced the number of questions from 17 in SSP to only to 5. The other questions are replaced with the automatic value captured using wearable/ infrared sensors data and emotion histogram data generated from facial expressions during the screening session when the stimuli are portrayed. The RSSP questions are rate from 1 to 5. 1 stands for never, and 5 stands for always. The questions are:

Q-1 Has trouble completing the task in case of noise around like radio, fan, refrigerator?
Q-2 Appears not to hear what you say (ignores you)?
Q-3 Enjoys strange noises and also makes noises?
Q-4 Has a hard time finding objects in competing for the background?
Q-5 Looks at things or people very closely or intensively?

3.5 Stimuli

The audio and visual stimuli are the short videos having audios, which are created based on the therapist's inputs. The stimuli are portrayed to the child during screening. The positive stimuli comprise happy images like cartoons, soft toys, sunlight, and the pleasing music of chirping of the birds, nursery rhymes, etc. The negative stimuli contain dark images, horrifying animals, etc., disturbing sounds like a vacuum cleaner, dog barking, ambulance siren, old fan noise, etc. The dataset used for the experimentation is mentioned in Table 3.

4. Modelling techniques

ASD detection is a binary classification problem with expected output class True and False. SPD classification is a ternary classification problem with

Table 3. Datasets used for experimentation

Dataset	Use	Details
AQ-10-child and AQ-10-adult	Autism detection	ASD Screening dataset by Fadi Fayez Thabtah
FER 2013 dataset	Training Emotion recognition model	Kaggle
Body sensor data	To identify stress level	Smart watch and infrared sensor
RSSP information	Contributes to audio/ visual sensory processing disorder detection	Based on inputs from SSP and developmental therapist

expected output class as over-responsive, under-responsive and typical. The algorithms used for the classification are:

4.1 Logistic regression

Data is modelled using the sigmoid function. The sigmoid function produces a 'S' shaped curve that may be used to convert any real-valued integer to a value between 0 and 1. If the curve reaches positive infinity, the projected value of y will be 1 and if the curve reaches negative infinity, the projected value of y will be 0. If the sigmoid function returns a value more than 0.5, we may categorise the result as 1 or TRUE, and if it returns a value less than 0.5, we may categorise it as 0 or FALSE. [23]

$$f(x) = p = \frac{1}{1+e^{-(x)}} \qquad (1)$$

x (x_1, x_2…, x_n) is independent variable and f(x) or p is the probability.

4.2 Linear discriminant analysis (LDA)

LDA creates a new axis by combining two features to reduce variance and increase class distance between the two variables. LDA is primarily concerned with projecting features from higher-dimensional space to lower-dimensional space. It assumes unimodal Gaussian likelihoods and thus parametric. [23]

4.3 k-nearest neighbor (k-NN)

During the training phase, this method uses the complete dataset. It scans the complete training dataset for the k-most comparable examples whenever a prediction is needed for an unknown data item. Finally, the prediction is based on the data with the most comparable instance. It's a straightforward algorithm that stores all existing examples and categorises new ones using a similarity metric like Euclidean distance. The distance values are sorted in ascending order. The top k- rows of the sorted array are then chosen. The most common class of these rows is used to assign a class to the test point. [23-26]

$$\text{Euclidean distance: } \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \qquad (2)$$

4.4 Decision tree

Decision trees produce transparent decisions which can be understood by people clearly. Various decision tree techniques include ID3, C4.5, Classification and Regression Tree (CART), Chi-square Automatic Interaction Detector (CHAID), and regression trees. CART employs a novel statistic called Gini index to construct decision points for classification tasks, whereas ID3 uses information gain and C4.5 utilises gain ratio for splitting. By default, CART performs binary splits. The limitation of the decision tree is that complex trees can be generated for simple problems. [24, 25]

$$\text{I\_Gini}=1-(\text{the probability of target "No"})^2-(\text{the probability of target "Yes"})^2 \qquad (3)$$

4.5 Gaussian naïve Bayes

The Bayes theorem calculates the probability of an event occurring given the probability of a previous event. When the features contain continuous values, it is employed. All of the characteristics are assumed to have a Gaussian distribution, or a normal distribution. Given the class variable y and the dependent feature vector x_1 through x_n, Bayes' theorem establishes the following relationship. [24, 25]

$$P(y|x_1, , x_n) = \frac{P(x_1|y)P(x_2|y)…P(x_n|y)P(y)}{P(x_1)P(x_2)….P(x_n)} \qquad (4)$$

4.6 Support vector machine (SVM)

Multiple continuous and categorical variables can be handled using SVM. In an ideal hyperplane, an SVM model is a multidimensional spatial representation of discrete classes. SVM generates the hyperplane repeatedly in order to reduce the error. SVM's purpose is to partition datasets into classes such that a maximum marginal hyperplane may be found. SVM is good for issues with multiple dimensions that aren't linearly separable. Several key parameters need to be set correctly viz. SVM type, SVM kernel type, gamma, degree, nu etc. [24-26]

4.7 Artificial neural network (ANN)

ANN sequentially creates high-level features using successive layers. Deep learning is a subset of a wider family of ANN-based machine learning algorithms. [27, 28] To perform the tasks like classification and detection, the use of representation learning, which is a technique that allows machines to discover relationships from raw data, is beneficial. It's the capacity to discover hidden relationships in data without imposing any fixed ones. ANNs can model non-linear and complex relationships and work with incomplete knowledge. They are fault tolerant and have parallel processing capability. ANNs are

networks of artificial neurons that accept input, merge it with their internal state and an optional threshold using activation function, and return using an output function. Weight is assigned to each connection that represents its relative importance.

Deep learning is an area of machine learning algorithm that uses many layers to extract higher-level functionality from raw input. [29, 30] Lower layers of image processing, for example, may identify edges, whereas higher layers may identify human-relevant concepts like numbers, letters, or faces. [31]

An artificial neuron is a function f_j of the input x $= x_1, \ldots, x_n$ weighted by a vector of connection weights $w_j = (w_{j,1}, \ldots, w_{j,d})$, completed by a neuron bias b_j. It is associated with an activation function ($\emptyset$). Activation functions (AFs) are used in DL architectures to perform various computations between the hidden layers and the output layers. These AFs are transition functions added to the linear model outputs to produce transformed non-linear results that can be processed further. [31, 32]

$$y_j = f_j(x) = \emptyset\left((w_j, x) + b_j\right) \qquad (5)$$

The logistic function, also known as the sigmoid activation function, has long been a standard activation function for neural networks. The function's input is converted to a value between 0.0 and 1.0. Inputs that are significantly greater than 1.0 are converted to 1.0, and inputs considerably smaller than 0.0 are snapped to 0.0. The sigmoid function is represented with:

$$\emptyset(x) = \frac{1}{1+e^{-x}} \qquad (6)$$

The softmax (optimization) function is mainly used to adapt the performance of neural networks between 0 and 1. It is used to denote the network output's "probability" of certainty. The normalization is determined by dividing the evaluated output's exponential value by the number of all potential outputs' exponential values.

$$\sigma\left(\vec{z}\right)_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \qquad (7)$$

4.8 Convolutional neural network (CNN)

CNN is a popular AI technique for image recognition, which utilizes a receptive field to extract every component of the input image. There are three types of layers in the CNN architecture: convolution, pooling, and fully connected. With multiplication or another dot product, the sequence of convolutional
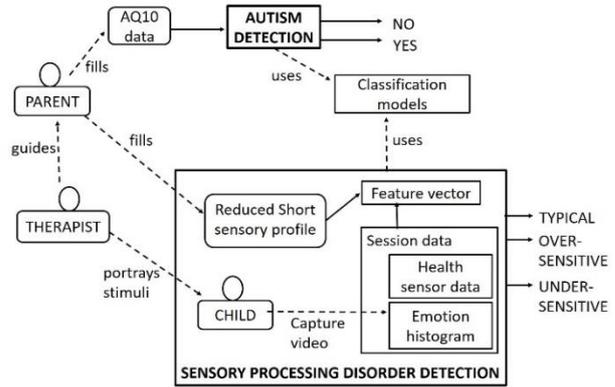

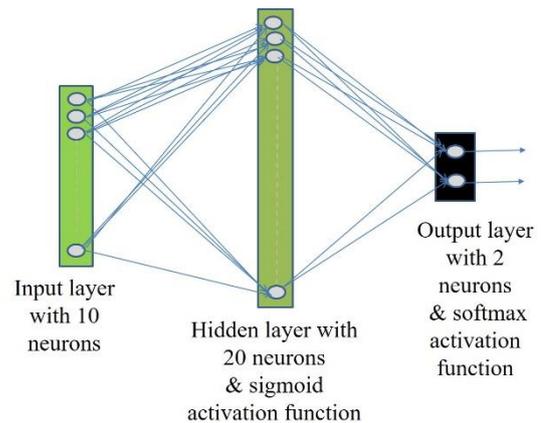Figure. 1 Block diagram of the proposed system

Figure. 2 Fully connected DNN model applied for Autism Detection on AQ10 dataset

layers convolves. A Rectified Linear Unit (ReLU) layer is typically used as the activation function, followed by pooling layers, fully connected layers, and normalisation layers. [27, 29, 30] Weight sharing decreases the number of free parameters, cutting memory needs and allowing the training of bigger, more powerful networks. The details of the CNN model used in this research are discussed in the proposed system. Large amount of training data is required to avoid overfitting in CNN. [31]

5. The proposed system

The proposed system, as illustrated in Figure. 1, consists of the following modules. Autism detection is done with the help of AQ10 dataset.

5.1 Autism detection module

The behavioral therapists create the AQ10 dataset after parent interaction and child observation. Machine learning algorithms are useful if the dataset is available. Fig. 2 shows the dense model applied for Autism Detection on the AQ10 dataset.

5.2 Stimuli for SPD

Stimuli are input to senses of children which cause the response. Stimuli are portrayed to children to check their reactions for the same. Stimuli are created by considering the questions in the sensory profile by Winnie Dunn and consulting with behavioral therapists.

The visual stimuli collected are very dark background images and bright images. These images are a collection of various sharp and extreme colors. Many autistic patients get triggered by such colors or images or cannot perceive the images or colors like ordinary people.

Similarly, the audio stimuli can be created with loud noises like vacuum cleaner, dog barking, hairdryer, car driving, old fan, and siren. It also includes strange noises, the sound of a radio tuning and refrigerator sound. These are our typical sounds in daily life with various frequencies and pitches that the autistic patient can be sensitive to or behave weirdly.

5.3 Health sensor data

A combination of wearable sensors is used to collect the values of health parameters of the child at the very instant when he reacts to the stimuli. The heart rate and temperature of the body are recorded when different stimuli are given to the autistic person to analyze the body parameter changes. Many children with autism oppose to wear any accessories like a pulse meter. Use fabric sensors or smart watches to collect body parameters would be recommended. Use of infrared thermometers to remotely collect temperature is possible too.

5.4 Emotion histogram generation module

The training to the facial expression recognition module is done using FER 2013 dataset. The facial images in the dataset are stored as 48x48 pixels along with the emotion code. [33, 34]
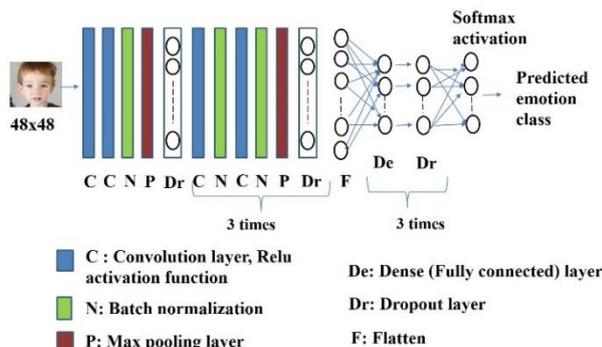

Figure. 3 CNN model used for facial emotion recognition

Table 4. Network architecture used for CNN

Layer No.	Dataset	Source	Total parameters
1	Convolution_1	(None, 46, 46, 32)	320
2	Convolution _2	(None, 46, 46, 32)	9248
3	Batch_normalization_1	(None, 46, 46, 32)	128
4	Max_pooling_1	(None, 23, 23, 32)	0
5	Dropout_1	(None, 23, 23, 32)	0
6	Convolution _3	(None, 23, 23, 64)	18496
7	Batch_normalization_2	(None, 23, 23, 64)	256
8	Max_pooling_2	(None, 11, 11, 64)	0
9	Dropout_2	(None, 11, 11, 64)	0
10	Convolution _4	(None, 11, 11, 128)	73856
11	Batch_normalization_4	(None, 11, 11, 128)	512
12	Max_pooling_3	(None, 5, 5, 128)	0
13	Dropout_3	(None, 5, 5, 128)	0
14	Convolution _5	(None, 5, 5, 256)	295168
15	Batch_normalization_6	(None, 5, 5, 256)	1024
16	Max_pooling_4	(None, 2, 2, 256)	0
17	Dropout_4 (Dropout)	(None, 2, 2, 256)	0
18	Flatten_1	(None, 1024)	0
19	Dense_1	(None, 256)	262400
20	Dropout_5	(None, 256)	0
21	Dense_2	(None, 128)	32896
22	Dense_3	(None, 64)	903
Total parameters: 6,95,207 Trainable parameters: 6,94,247 Non-trainable parameters: 960			

The CNN model for face emotion recognition is shown in Fig. 3, TensorFlow was used to implement the CNN in a clustered CPU processing environment. The end-to-end algorithm was written in Python 3.8. The CNN was made up of five 2D convolution layers with a 3x3 kernel and four max-pooling layers with a stride of 2. To avoid overfitting, each layer was regularised with a dropout of 0.5. Each of the convolutional layers was activated with ReLU. The parameters were mixed to integrate the feature vectors in the following three layers, which were fully connected. To produce the expected emotion output, the last fully connected layer was employed with soft-max activation.
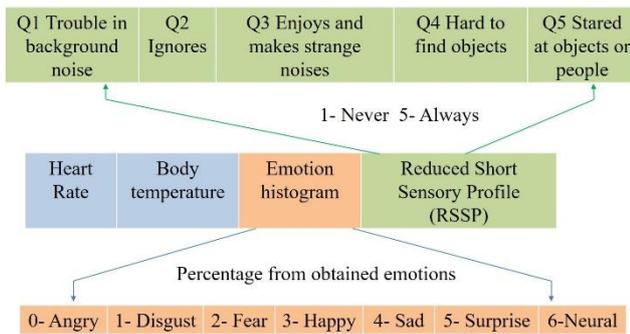
Figure. 4 Feature vector for sensory processing disorder detection

Figure. 5 System architecture

In machine learning, optimizers are used to fine-tune the parameters of a neural network in order to reduce the cost function. Due of the unbalanced training dataset, a weighted cross-entropy with suitably adjusted weights for the positive class was utilised as the loss function. We utilised the Adam optimizer, which has the benefit of keeping an exponentially declining the average of previous gradients, much like momentum. Table 4 contains the specifics of the network architecture employed in the model. [17, 22, 35, 36]

5.5 Feature vector

The combined dataset is created, which contains body temperature, pulse rate, emotion histogram and RSSP. Fig. 4 represents the novel hybrid feature vector for sensory processing disorder detection.

5.6 Pattern recognition

Machine learning and deep learning models trained with the feature vector of normal children of same age is used to find the pattern of emotions for particular stimuli.

5.7 System architecture

The detailed system architecture is shown in Fig. 5.

Step1: The identification of child with autism is made based on the AQ10 trained deep learning model.

Step2: Standard audio and visual stimuli based on sensory profile questionnaire and expert suggestions are portrayed.

Step3: Capture the video of facial expression during the session. This video is input to the facial emotion recognition model trained with the FER2013 dataset.

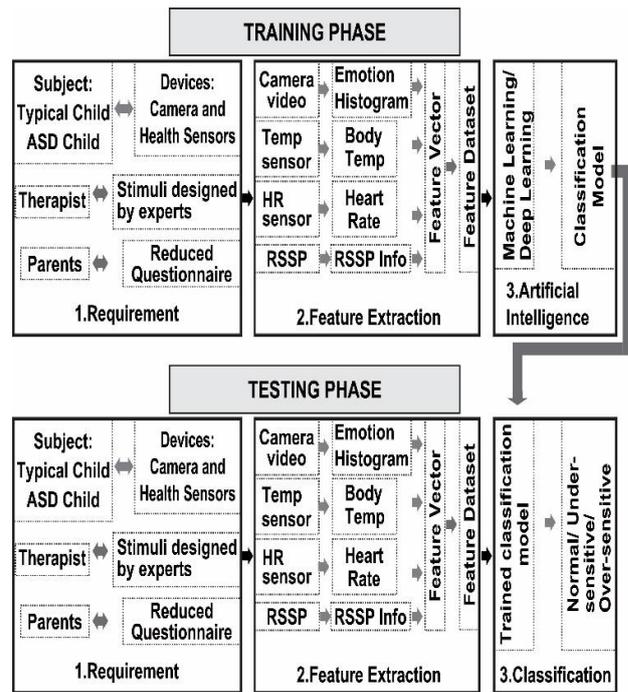Step4: Capture the stress-related health parameters during the session using a heart rate sensor and thermometer.

Step4: The system architecture as specified in Figure. 5 shows that the generation of the benchmark dataset is done according to the feature vector shown in the figure.

Step5: A classification model is created by applying various Machine Learning (ML) and Deep Learning (DL) algorithms. Thus, the trained model is generated for identifying the auditory or visual sensory disorder.

Step6: Classification is done on the test data in a real-time environment.

6. Results

6.1 Autism detection

To classify autistic and non-autistic people, we have applied various ML techniques to the AQ10 dataset. For the adult AQ10 and child AQ10 datasets, the accuracy and standard deviation areas are listed in Table 5. Figure. 6 and 7 show the box plots of the total classification accuracy for the six classifiers studied. The k-fold cross-validation for value of k as 10 is used to improve the algorithm performance. There were statistically significant differences between distributions of the total accuracy. It is observed that SVM gives better classification accuracy with a small standard deviation for both datasets.

A deep neural network using the dense model of fully connected layers with sigmoid activation and softmax optimization was applied on the same AQ10 datasets. A comparison is made with SVM.

Table 5. Results of ML classifiers for autism detection

Algorithm	adult AQ10 dataset Accuracy		child AQ10 dataset Accuracy	
	Mean	Standard deviation	Mean	Standard Deviation
Linear regression	0.980802	0.020057	0.945952	0.067797
LDA	0.953927	0.030454	0.950952	0.066045
k-nearest neighbour	0.971844	0.026162	0.911429	0.066503
Dicision tree (CART)	0.951396	0.031805	0.863333	0.093745
Gaussian naive bias	0.965433	0.028141	0.882143	0.095691
SVM	0.974391	0.028089	0.955952	0.045578

Figure. 6 Performance of ML classifiers for Adult AQ10 dataset

Figure. 7 Performance of ML classifiers for Child AQ10 dataset

Table 6. Results of SVM and fully connected deep learning model for autism detection

Dataset	Classifier	Class	Accuracy	Precision	f1-score	support
Adult AQ10 dataset	SVM	0	0.97	0.97	0.99	230
		1	0.97	1.00	0.97	106
	Deep learning	0	1.00	1.00	1.00	230
		1	1.00	1.00	1.00	106
Child AQ10 dataset	SVM	0	0.96	1.00	0.97	48
		1	0.96	0.93	0.96	40
	Deep learning	0	1.00	1.00	1.00	48
		1	1.00	1.00	1.00	40

Table 6 shows the performance of both models. The obtained validation accuracy of 1 was obtained. Also, the test data of 10 children was input for testing, and we received the test accuracy of 1. The accuracy obtained by Tania Akter and others [13] was 0.97 for the same AQ10 dataset. This DNN classifier for autism detection is one of the research contribution.

6.2 Sensory processing disorder detection

The short sensory profile questionnaire data and some observations given by the developmental therapist for six children S1, S2, S3, S4, S5 and S6 for auditory and visual processing are shown in the Table 7. It is observed that the profile is completely based on the observations done by the parents. But families are mostly engaged in the daily activities which makes it difficult to note down and convey the observations correctly at the time of screening. Also, a lot of processing needs to be done after the profile is filled properly so that appropriate therapy sessions are initiated. Screening is done multiple times during the therapy sessions to check whether the therapy is helping the child or not. This needs expertise, time and efforts of the developmental therapist. Also, the parents or the caregivers must be continuously engaged to observe the child's behaviour in every situation so that the sensory profile score will be perfect.

The proposed system represented in Fig. 1 and Fig. 5 will support the developmental therapists with reduced set of 5 questions and automatic collection of parameters related to the behaviour of the child. The video captured while portraying particular audio/ visual stimuli is processed, and an emotion histogram is generated by identifying the emotion using facial expression and health parameters related to stress.
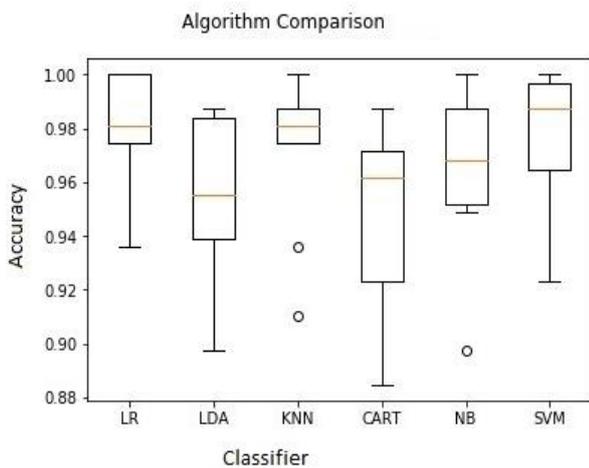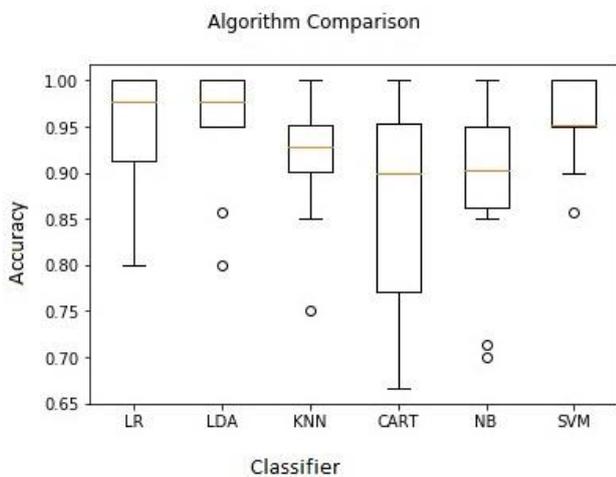
Table 7. Emotion count from the video captured during screening

Sample	Duration (sec)	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral	Total
S1	33	4	0	5	115	21	48	55	248
S1	65	1	0	2	52	89	4	175	323
S1	71	31	0	1	151	34	1	102	320
S2	165	10	0	12	10	61	7	60	160
S2	130	19	0	8	4	14	7	192	244
S3	186	4	0	5	9	27	2	66	113

Table 8. Performance analysis of emotion recognition model

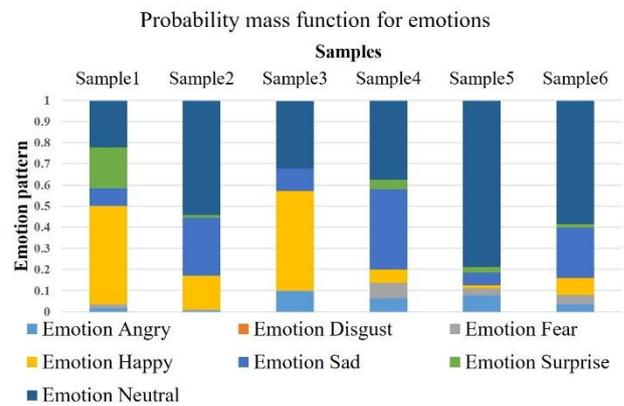| Emotion		Precision	Recall	f1-score	Support
0	Angry	0.74	0.83	0.78	165
1	Disgust	0.2	0.3	0.2	41
2	Fear	0.81	0.51	0.63	123
3	Happy	0.88	0.98	0.93	669
4	Sad	0.90	0.90	0.90	516
5	Surprise	0.72	0.81	0.77	159
6	Neutral	0.96	0.89	0.92	323
Test accuracy=0.86		0.75	0.75	0.73	1996
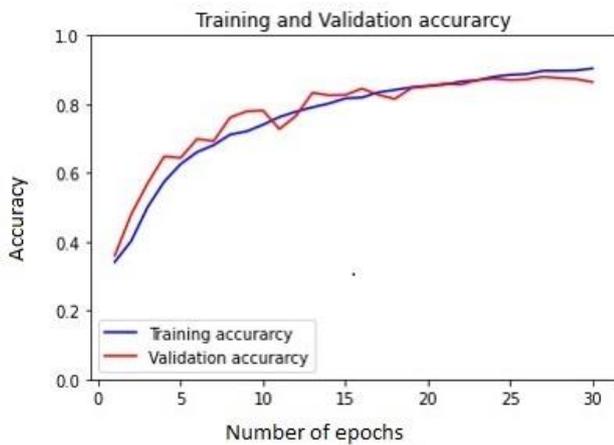
Figure. 8 Validation accuracy for CNN model

Figure. 9 Emotion pattern for 6 samples

The emotion count for three children is collected. The CNN model is trained using the FER2013 dataset and tested on the eight videos of 3 clinical samples. The training and validation accuracy and loss of the CNN model for 30 epochs are observed. Validation Accuracy is 86% after 30 epochs. Also, on the clinical data, Test Accuracy is 86%. The videos were captured during the screening session as a response to the portrayed audio and visual stimuli and the data of ASD child faces was created. We trained the CNN model with this dataset of ASD faces.

The CNN model is trained with the split of 90% train data and 10% validation data. The training and validation accuracy are shown the Fig. 8.

Table 8 shows the performance analysis obtained by the facial emotion recognition module using the trained CNN model. The model's performance concerning the predictive precision, recall, f1-score, and support are shown in the table. High precision to correctly identify the children with a sensory disorder and high recall is expected not to miss any case in the sensory disorder detection. Thus, the f1-score, which is the harmonic mean of precision and recall, is essential as it considers both the metrics by giving the same weightage.

The histograms for emotions are generated in various scenarios when audio and visual stimuli are portrayed. The percentage of emotions from the obtained emotion count is calculated. The mood swings can be monitored using these histograms easily as compared to EEG method in [16].

As shown in Fig. 9, the obtained emotion pattern will give clear intimation to the therapist regarding the expected emotions and observed emotions for a particular type of stimulus. The features collected using different sensors and videos captured during the session while stimuli designed by experts are portrayed will help the therapist get a clear knowledge about the child's behavior in various situations. This will help them to start the appropriate therapy. The obtained results show that the deep learning and CNN models are useful for autism detection and getting knowledge about the child's behaviour in various day-to-day situations.

Table 9. Results of ML classifiers for sensory processing disorder detection

Algorithm	Accuracy	Precision	Recall	RoC value
Linear regression	1.0	1.0	1.0	0.96
LDA	1.0	1.0	1.0	1.0
k-nearest neighbour	0.87	0.84	0.87	0.91
Dicision	1.0	1.0	1.0	1.0
Gaussian naive bias	1.0	1.0	1.0	1.0
SVM	1.0	1.0	1.0	1.0
DNN	0.94	0.97	0.96	1.0
ANN	0.97	0.95	0.97	1.0

To classify the typical, over-responsive and under-responsive children during SPD detection, we have applied various ML techniques to the novel feature vector dataset according to format given in Fig. 4. This feature vector dataset used for SPD detection is another contribution of this research work as the data is not readily available. The accuracy, precision, recall and RoC value for different classifiers applied on the data listed in Table 9. The k-fold cross-validation for value of k as 10 is used to improve the algorithm performance. It is observed that decision tree gives better classification accuracy. But, as the number of training samples increase, DNN may be more useful.

7. Conclusion

The ASD and SPD detection is currently based on the questionnaire. In Asian countries like India, the community is diverse in regional language, financial, and educational diversity. Thus, these questions need to be translated by the therapist and is time-consuming. Also, there is a lack of awareness of the developmental screening sessions and the need to maintain timely reporting. In the current pandemic situation, distant and fast screening of ASD people will help to reduce the risk of infection. There is a need for a tool that can make the screening faster. [36]

The author's contribution is to propose a novel feature vector for SPD detection. The questionnaire number can be reduced drastically and processed automatically using deep learning instead of manual score calculation on the lengthy questionnaire. The sensory profile questionnaire is also extensive and time-consuming, from which few parameters have been collected automatically. The use of video processing techniques, smart watch, body sensors for automatic feature collection, and the reduced

questionnaire is unique. The use of a dense, deep learning model has given 100% accuracy on the AQ10 dataset. The proposed system for sensory processing disorder detection creates the feature vector with inputs from RSSP questionnaire, heart rate, body temperature, and facial expressions obtained with CNN. The emotion histograms generated during the session are useful for the therapist to relate the expressions with the respective stimuli. The machine learning and deep learning algorithms applied to this feature vector create an accurate model for auditory and visual sensory disorder detection.

Thus, with this research, the authors have contributed ASD detection model using DNN on AQ10 dataset, emotion detection model for autistic children using CNN on FER2013 dataset, emotion histogram to track sensory processing pattern of children using CNN model and SPD classification model using novel hybrid feature vector.

There is scope for improvement of the proposed system by enhancing the stimuli and wearable sensor parameters. The use of two audio and visual stimuli is done for capturing the screening data. Multiple stimuli can be designed by in-depth analysis of the questions and inputs from various experts. Use of advanced smart watches is suggested for capturing more stress related parameters like SPO2, ECG and sweat. The emotion detection model is trained with FER2013 dataset which is biased and not specific ASD dataset resulting in less accuracy. There are many opportunities for creation of such dataset for better accuracy of emotion recognition model. [8, 15] Also, multiple cameras can be simultaneously used to capture proper face in the video. The fabric sensors or infrared sensors are available for getting more stress-related parameters. The current work concentrates only on auditory and visual sensory disorder detection. Other types of sensory disorders can also be automatically detected. This tool will prove valuable for developmental therapists, ASD children, and parents by saving time and effort without compromising detection accuracy.

Conflicts of Interest

Author Contributions

References

[1] H. Kuhaneck and P. Britner, "A Preliminary Investigation of the Relationship between Sensory Processing and Social Play in Autism Spectrum Disorder", *OTJR: Occupation, Participation and Health*, Vol. 33, No. 3, pp. 159-167, 2013.

[2] S. Tomchek, R. Huebner, and W. Dunn, "Patterns of sensory processing in children with an autism spectrum disorder", *Research in Autism Spectrum Disorders*, Vol. 8, No. 9, pp. 1214-1224, 2014.

[3] D. Wall, J. Kosmicki, T. Deluca, E. Harstad, and V. Fusaro, "Use of machine learning to shorten observation-based screening and diagnosis of autism", *Translational Psychiatry*, Vol. 2, No. 4, pp. e100-e100, 2012.

[4] D. Bone, S. Bishop, M. Black, M. Goodwin, C. Lord, and S. Narayanan, "Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion", *Journal of Child Psychology and Psychiatry*, Vol. 57, No. 8, pp. 927-937, 2016.

[5] F. Thabtah, "Autism Spectrum Disorder Screening", In: *Proc. of the 1st International Conf. on Medical and Health Informatics*, pp. 1-6, 2017.

[6] Y. Sinha, N. Silove, A. Hayen, and K. Williams, "Auditory integration training and other sound therapies for autism spectrum disorders (ASD)", *Cochrane Database of Systematic Reviews*, 2011.

[7] E. Benssassi, J. Gomez, L. Boyd, G. Hayes, and J. Ye, "Wearable Assistive Technologies for Autism: Opportunities and Challenges", *IEEE Pervasive Computing*, Vol. 17, No. 2, pp. 11-21, 2018.

[8] M. Kientz and W. Dunn, "A Comparison of the Performance of Children With and Without Autism on the Sensory Profile", *The American Journal of Occupational Therapy*, Vol. 51, No. 7, pp. 530-537, 1997.

[9] M. Marlow, C. Servili, and M. Tomlinson, "A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low- and middle-income countries", *Autism Research*, Vol. 12, No. 2, pp. 176-199, 2019.

[10] E. Barton, B. Reichow, A. Schnitz, I. Smith, and D. Sherlock, "A systematic review of sensory-based treatments for children with disabilities", *Research in Developmental Disabilities*, Vol. 37, pp. 64-80, 2015.

[11] J. Kosmicki, V. Sochat, M. Duda, and D. Wall, "Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning", *Translational Psychiatry*, Vol. 5, No. 2, pp. e514-e514, 2015.

[12] F. Abdeljaber, "Detecting autistic traits using computational intelligence & machine learning techniques", *PhD diss., University of Huddersfield*, 2019.

[13] T. Akter, M. Satu, M. Khan, M. Ali, S. Uddin, P. Lio, J. Quinn, and M. Moni, "Machine learning-based models for early stage detection of autism spectrum disorders", *IEEE Access*, Vol. 7, pp. 166509-166527, 2019.

[14] S. Tomchek and W. Dunn, "Sensory Processing in Children With and Without Autism: A Comparative Study Using the Short Sensory Profile", *The American Journal of Occupational Therapy*, Vol. 61, No. 2, pp. 190-200, 2007.

[15] B. Li, A. Sharma, J. Meng, S. Purushwalkam, and E. Gowen, "Applying machine learning to identify autistic adults using imitation: An exploratory study", *PLOS ONE*, Vol. 12, No. 8, p. e0182652, 2017.

[16] C. Busso, Z. Deng, Z. Yildirim, M. Bulut, C. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, "Analysis of emotion recognition using facial expressions, speech and multimodal information", In: *Proc. of the 6th International conf. on Multimodal Interfaces*, pp. 205-211, 2004.

[17] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey", *IEEE Transactions on Affective Computing*, 2020.

[18] A. Samal and P. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: a survey", *Pattern Recognition*, Vol. 25, No. 1, pp. 65-77, 1992.

[19] F. Thabtah, "An accessible and efficient autism screening method for behavioural data and predictive analyses", *Health Informatics Journal*, Vol. 25, No. 4, pp. 1739-1755, 2019.

[20] F. Thabtah, "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward", *Informatics for Health and Social Care*, Vol. 44, No. 3, pp. 278-297, 2019.

[21] K. Li, S. Lou, H. Tsai, and R. Shih, "he Effects of Applying Game-Based Learning to Webcam Motion Sensor Games for Autistic Students' Sensory Integration Training", *Turkish Online Journal of Educational Technology-TOJET*, Vol. 11, No. 4, pp. 451-459, 2012.

[22] N. Rad, S. Kia, C. Zarbo, T. V. Laarhoven, G. Jurman, P. Venuti, E. Marchiori, and C. Furlanello, "Deep learning for automatic stereotypical motor movement detection using wearable sensors in autism spectrum disorders", *Signal Processing*, Vol. 144, pp. 180-191, 2018.

[23] M. Ala'raj, M. Majdalawieh, and M. Abbod, "Improving binary classification using filtering based on k-NN proximity graphs", *Journal of Big Data*, Vol. 7, No. 1, pp. 1-18, 2020.

[24] "K-Nearest Neighbor in Machine Learning", Knowledgehut.com, 2021.

[25] "Issues mainkoon81/Study-09-MachineLearning-D", Github, 2021.

[26] H. Raeisi Shahraki, S. Pourahmad, and N. Zare, "K Important Neighbors: A Novel Approach to Binary Classification in High Dimensional Data", *BioMed Research International*, pp. 1-9, 2017.

[27] N. Jadhav and R. Sugandhi, "Survey on human behavior recognition using affective computing", In: *Proc. of IEEE Global Conference on Wireless Computing and Networking*, pp. 98-103, 2018.

[28] P. Lanillos, D. Oliva, A. Philippsen, Y. Yamashita, Y. Nagai, and G. Cheng, "A review on neural network models of schizophrenia and autism spectrum disorder", *Neural Networks*, Vol. 122, pp. 338-363, 2020.

[29] N. Hammerla, S. Halloran, and T. Plötz, "Deep, convolutional, and recurrent models for human activity recognition using wearables", *ArXiv Preprint ArXiv:1604.08880*, 2016.

[30] V. Patnam, F. George, K. George, and A. Verma, "Deep learning based recognition of meltdown in autistic kids", In: *Proc. of 2017 IEEE International Conference on Healthcare Informatics*, pp. 391-396, 2017.

[31] A. Khan, A. Sohail, U. Zahoora, and A. Qureshi, "A survey of the recent architectures of deep convolutional neural networks", *Artificial Intelligence Review*, Vol. 53, No. 8, pp. 5455-5516, 2020.

[32] L. Kaur and V. Khullar, "A Review on using Artificial Neural Network in Diagnosis of Autism Spectrum Disorder", *i-Manager's Journal on Computer Science*, Vol. 5, No. 1, p. 38, 2017.

[33] X. Huang, J. Kortelainen, G. Zhao, X. Li, A. Moilanen, T. Seppänen, and M. Pietikäinen, "Multi-modal emotion analysis from facial expressions and electroencephalogram", *Computer Vision and Image Understanding*, Vol. 147, pp. 114-124, 2016.

[34] M. Soleymani, S. A. Esfeden, Y. Fu, and M. Pantic, "Analysis of EEG Signals and Facial Expressions for Continuous Emotion Detection", *IEEE Transactions on Affective Computing*, Vol. 7, No. 1, pp. 17-28, 2016.

[35] N. Rad, A. Bizzego, S. Kia, G. Jurman, P. Venuti, and C. Furlanello, "Convolutional neural network for stereotypical motor movement detection in autism", *ArXiv Preprint ArXiv:1511.01865*, 2015.

[36] S. Saha, A. Pagnozzi, P. Bourgeat, J. George, D. Bradford, P. Colditz, R. Boyd, S. Rose, J. Fripp, and K. Pannek, "Predicting motor outcome in preterm infants from very early brain diffusion MRI using a deep learning convolutional neural network (CNN) model", *NeuroImage*, Vol. 215, p. 116807, 2020.