Journal of Medical Systems (2024) 48:15
https://doi.org/10.1007/s10916-023-02032-0

ORIGINAL PAPER

Multiple Classification of Brain MRI Autism Spectrum Disorder by Age and Gender Using Deep Learning

Hidir Selcuk Nogay[1] · **Hojjat Adeli**[2]

Received: 29 August 2023 / Accepted: 31 December 2023 / Published online: 22 January 2024
© The Author(s) 2024

Abstract

The fact that the rapid and definitive diagnosis of autism cannot be made today and that autism cannot be treated provides an impetus to look into novel technological solutions. To contribute to the resolution of this problem through multiple classifications by considering age and gender factors, in this study, two quadruple and one octal classifications were performed using a deep learning (DL) approach. Gender in one of the four classifications and age groups in the other were considered. In the octal classification, classes were created considering gender and age groups. In addition to the diagnosis of ASD (Autism Spectrum Disorders), another goal of this study is to find out the contribution of gender and age factors to the diagnosis of ASD by making multiple classifications based on age and gender for the first time. Brain structural MRI (sMRI) scans of participators with ASD and TD (Typical Development) were pre-processed in the system originally designed for this purpose. Using the Canny Edge Detection (CED) algorithm, the sMRI image data was cropped in the data pre-processing stage, and the data set was enlarged five times with the data augmentation (DA) techniques. The most optimal convolutional neural network (CNN) models were developed using the grid search optimization (GSO) algorism. The proposed DL prediction system was tested with the five-fold cross-validation technique. Three CNN models were designed to be used in the system. The first of these models is the quadruple classification model created by taking gender into account (model 1), the second is the quadruple classification model created by taking into account age (model 2), and the third is the eightfold classification model created by taking into account both gender and age (model 3).). The accuracy rates obtained for all three designed models are 80.94, 85.42 and 67.94, respectively. These obtained accuracy rates were compared with pre-trained models by using the transfer learning approach. As a result, it was revealed that age and gender factors were effective in the diagnosis of ASD with the system developed for ASD multiple classifications, and higher accuracy rates were achieved compared to pre-trained models.

Keywords ASD · Multiple classification · CNN · CED · Data augmentation · GSO · sMRI

Introduction

ASD is a neurodevelopmental disease that occurs in early childhood and is characterized by communication disorders and difficulties in socialization in children [1, 2]. There has been an increase in the incidence of Autism Spectrum Disorder over the years, and while one in every 150 children in America was autistic in 2000, it is reported that one in every 54 children has autism in 2020 [3, 4].

Despite an extensive range of signs of ASD [5], a complication that prolongs the diagnosis process is the high rate of comorbidity. The comorbidity problem in children with ASD means an extra disability like a vision problem or another health problem [6]. A study revealed that 88.5% of children diagnosed with autism had at least one of the neurodevelopmental disorders such as attention deficit hyperactivity disorder (ADHD), intellectual disability and developmental coordination disorder [7]. The incidence of autism is higher in boys than in girls [8]. Although the reason for this is not clear, hypotheses such as Extreme Male Brain, Female

✉ Hojjat Adeli
adeli.1@osu.edu

Hidir Selcuk Nogay
hsnogay@uludag.edu.tr

[1] Electrical and Energy Department, Bursa Uludag University, Bursa, Turkey

[2] Departments of Biomedical Informatics and Neuroscience, College of Medicine, The Ohio State University Neurology, 370 W. 9th Avenue, Columbus, OH 43210, USA

Protective Effect, and Female Autism Phenotype are being studied [9]. The lack of a known cure for autism, the long diagnosis and treatment process [10], and the high degree of comorbidity all indicate that more scientific work is needed on autism [11]. There is an important need to study the influence of age and gender factors on ASD diagnosis and to evaluate the possibility that multiple classifications, including age and gender, may contribute to the rapid early diagnosis of ASD. Recent genetic works show that ASD occurs differently between males and females and between youths and adults [12]. Artificial intelligence and machine learning (ML) techniques [13, 14] such as DL provide fresh opportunities to discover biomarkers for diagnosis of ASD taking into account factors like age and gender that affect ASD, to shorten the diagnostic process of ASD, to avoid subjective opinions of different doctors and possibly reach a definitive diagnosis [15–17].

DL techniques have found extensive application in medical and neurological fields such as seizure detection [18], seizure prediction [19–21], epilepsy diagnosis and classification [22, 23], autism [24–27], optimization of neuroprosthetic vision [28], post-stroke rehabilitation with motor imagery [29], sentiment analysis [30], emotion recognition [31, 32], patient-specific quality assurance [33], classification of the intracranial electrocorticogram [34], brain-computer interface (BCI) for discriminating hand motion planning [35], dyslexia biomarker detection [36–38], and many other fields such as mobile robots [39], drone-based water rescue and surveillance [40], and structural health monitoring in recent years [41–43].

The design and effectiveness of a DL method for diagnosing ASD varies according to the data set. The data set can be numeric or two-dimensional graphical, or visual data. Numerical data can be behavioral [44, 45], eye-following [46], or fingerprint data [47–49], converted into numerical data by pre-processing. Optical data are brain structural magnetic resonance scanning images (sMRI) or brain functional magnetic resonance scanning images (fMRI). Using numerical or visual data to train an ML algorithm for ASD diagnosis is ordinarily possible by determining the distinguishing features or using an automated feature extraction technique [50–52]. These features may be structural gray matter (GM) values acquired from cortical thickness (CT) [53–55], GM density (GMd) values from voxel-based morphometry (VBM) [56], diffusion-weighted imaging (DWI) [fractional anisotropy (FA)] in white matter (WM)) microorganism changes [57], connectivity matrices [58], parameters from network analysis [59–61], and resting/duty state fMRI information [62, 63]. However, if a type of DL known as convolutional neural network (CNN) is utilized, direct classification is performed because feature extraction is done automatically. This is known as end-to-end deep learning [64]. For this reason, the CNN method is employed in this research as the most suitable method for rapid diagnosis of ASD.

In the study, the influence of a certain age range and gender on the diagnosis of ASD is examined by performing multiple classifications of ASD based on age and gender. A DL system has been introduced that can diagnose ASD for certain age ranges and gender. The advantages and differences of the current research compared to previously-reported research on ASD diagnosis, binary classification, and/or multiple classification works can be listed as follows. First, multiple classifications, including age and gender, were performed in this study, and to the best of the authors' knowledge, this has never been done before. Second, compared to other works that employ a DA method, the number of image data in this study is huge and acquired from different brain regions. This is advantageous in terms of the generalizability of the models. Third, CNN was designed from scratch and utilized as a system element in this study. Thus, feature extraction is done automatically. Fourth, using a transfer learning (TL) method, today's popular pre-trained models were trained and tested with the same data set.

The following sections are organized as follows. In the next section, works on ASD classification using brain MRI images, which also considered other factors like age and gender, are discussed. The third section explains the techniques and materials utilized in the study. In the fourth section, metrics used to evaluate the performance of the study are presented. The fifth section reports the numerical experimental results acquired from the study. The paper ends with discussions and a conclusion.

Related works

Although multiple classifications are more informative for ASD diagnosis works using brain sMRI scans, researchers have not studied them due to their complexity and difficulty in achieving high accuracy rates. As a result, the authors could not find a CNN model trained with brain sMRI images that could perform quadruple and octal classification, including gender and age factors, for ASD diagnosis. Therefore, in this study, multiple classifications were made through binary pairings such as F-ASD and F-TD, M-ASD and M-TD (F represents female, M represents male, and TD represents typical development) using brain MRI images. The quadruple classification was made using only gender, another quadruple classification using only age range, and the octal classification using both gender and age factors.

In a multiple binary classification study of ASD conducted by [12], they created separate groups like 'adolescents-F (< 18years)' - 'adolescents-M(< 18years)' and 'adults-F(> 18years)' - 'adult-M(> 18years)' [12]. They investigated which group could be diagnosed with ASD with the highest accuracy rate by making separate binary

classifications using by Extended Metacognitive Radial Basis Function Neural Classifier (EMcRBFN) method, which is trained and tested by sMRI data. They found that ASD can be detected more accurately in women (81%) than in men (60%). In [65], it investigated the impact of gender factors on the diagnosis of ASD in multiple binary classifications. In their study with the Support Vector Machine (SVM) method, they obtained an accurate prediction rate of 69% for the ASD-F (female) group and 66% for the ASD-M (male) group. However, the data set was limited to the 18–49 age group [65]. In [66], it employed DL trained with brain fMRI scans and performed binary ASD classification reporting a 70% accuracy [66]. In another DL study dealing with the age factor, they were able to diagnose ASD in the 2-year-old group with an accuracy rate of 76.24% using the "Multi-Channel Convolutional Neural Network" (MC-CNN) [67]. In [68], they performed a binary classification of ASD and reported an accuracy rate of 65.69% deep belief network (DBN) model [68]. In [69], they diagnosed ASD with 90.39% accuracy in binary classification using a DL algorithm trained with brain sMRI scans of participators whose mean age was 15 [69].

Materials and methodology

Dataset

The ABIDE database, an international professional database made available on March 27, 2017, was used to train and test the models in the study [70]. Detailed information about ABIDE can be obtained from http://fcon_1000.projects.nitrc.org/indi/abide/. The data in the ABIDE database consist of data collected from 29 different sites shown in Table 1. T1 weighted sequence and sMRI of 2248 participators, 1072 ASD and 1176 TD gathered from 29 locations from ABIDE, constituted first repository of the study called Data1. All images in the repository were scanned for clarity one by one. After the sharpness scan, a sum of 1831 image data, 938 ASD and 893 TD, were used as Data1 in this study. No coloration or any application affecting discrimination was made on any image data. For the three multi-classification CNN models utilized in the study, the data were distributed for all three models, as shown in Table 2, before pre-processing.

Table 1 Distribution of the data utilized in this research

SITE	ASD		TD		ALL
	Ages				
	5–17	18–65	5–17	18–65	
STANDFORD	20	0	20	0	40
KKI	77	0	188	0	265
KUL	5	23	0	0	28
LEUVEN	16	13	21	14	64
UCD	19	0	14	0	33
OHSU	51	0	70	0	121
MAXMUN	9	15	6	26	56
UCLA	81	0	68	0	149
BNI	2	25	2	27	56
CALTECH	1	18	2	17	38
EMC	27	0	27	0	54
GU	51	0	54	0	105
IP	17	4	10	23	54
NYU	134	20	104	31	289
PITT	18	12	15	12	57
SDSU	46	0	47	0	93
TRINITY	16	8	15	10	49
UM	80	0	89	3	172
UPSM	17	1	15	2	35
YALE	28	0	28	0	56
SU	21	0	21	0	42
OLIN	14	6	10	6	36
ETH	4	7	3	21	35
TCD	18	3	16	5	42
IU	2	17	0	19	38
ONRC	5	18	1	28	52
USM	36	38	26	32	132
CMU	0	14	0	13	27
SBL	0	15	0	15	30
ALL Sites	815	257	872	304	2248

Data pre-processing

The raw data were subjected to a three-step pre-processing. In the first pre-processing step, the unclear images were eliminated. Figure 1 shows examples of vague images.

In the second pre-processing step, the Canny Edge Detection (CED) algorithm subjected each image data to edge detection. The new image acquired after CED processing is cropped from the determined edges, and the lost area is minimized. Figure 2 describes the second step of data pre-processing, including the CED processing.

In the third step of the pre-processing, each image was subjected to DA by right-left flip, 90° right rotation, 180° right rotation, and 5% salting, magnifying the dataset five-fold. Figure 3 shows the application of the planned DA technique for only one sample image.

Proposed CNN models

Optimal hyper-parameter selection

Three DL models were designed as part of the system developed in this study. For each model, the GSO algorithm was utilized to decide the most optimal

Table 2 Summary of the data sets

Datasets	Class Number	Groups	Size	Total Size	Gender	Age range
Data1	1	ASD + f	127	1831	Female	-
	2	ASD + m	811		Male	-
	3	TD + f	209		Female	
	4	TD + m	893		Male	
Data2	1	ASD 5–17	648	1831	-	5–17
	2	ASD 18–65	290		-	18–65
	3	TD 5–17	594		-	5–17
	4	TD 18–65	299		-	18–65
Data3	1	ASD 5–17 f	96	1831	Female	5–17
	2	ASD 5–17 m	552		Male	5–17
	3	ASD 18–65 f	31		Female	18–65
	4	ASD 18–65 m	259		Male	18–65
	5	TD 5–17 f	157		Female	5–17
	6	TD 5–17 m	437		Male	5–17
	7	TD 18–65 f	52		Female	18–65
	8	TD 18–65 m	247		Male	18–65

hyperparameters among the limit values determined in Table 3 [71]. After pre-processing the data set, each model was tested with a randomly selected 20% of the ready-to-utilize data. From top to bottom, the first five rows of hyperparameters in Table 3 are about the architecture of the CNN models, and the next five are about fine-tuning each architecture. In Fig. 4, the system designed in the study is described schematically.

Convolution and pooling

The convolution operation is the processing of acquiring the **B** output matrix as a result of filtering the **A** image matrix entering a CNN model with the **K** filter matrix, as shown in Eq. (1). The resulting output matrix **B** is smaller than the input matrix **A**. During the filtering processing, the filter matrix **K** on matrix **A** can be shifted as much as the shift step (*stride*). In some strategies, resizing the output matrix to the same size as matrix **A** may be desirable. In this case, it can be brought to the same size as the **A** matrix by filling the blank parts of the **B** matrix with the number zero. This processing is called *padding*. In addition, before the convolution operation, pooling, which is a sub-operation of the convolution operation, is performed to reduce overfitting. In the pooling processing, the input matrix of the pooling layer is filtered by the selected filter matrix on the principle of mean or maximum values [72]. With Eq. 2, the size of the output matrix is obtained as a result of the filtering used in both pooling and convolution operations [72].

$$B_{ij} = (A * K)_{ij} = \sum_{f=0}^{n_k-1} \sum_{h=0}^{n_k-1} A_{i+f,j+h} K_{i+f,j+h} \tag{1}$$

$$n_B = \lfloor \frac{n_A + 2p - n_K}{s} + 1 \rfloor \tag{2}$$

Softmax and classification

The cross-entropy loss is calculated during the classification process. Softmax function is the layer before the classification layer. Multiclassification is performed as probabilistic in the Softmax layer. The softmax function for the multiple classifications is expressed as follows [73].

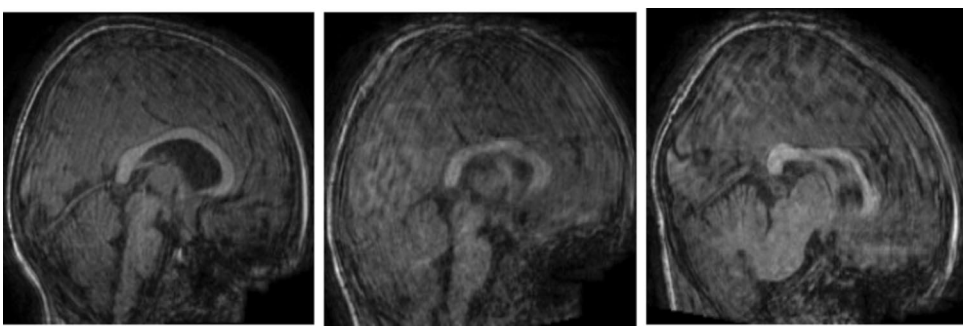

Fig. 1 Examples of eliminated sample images

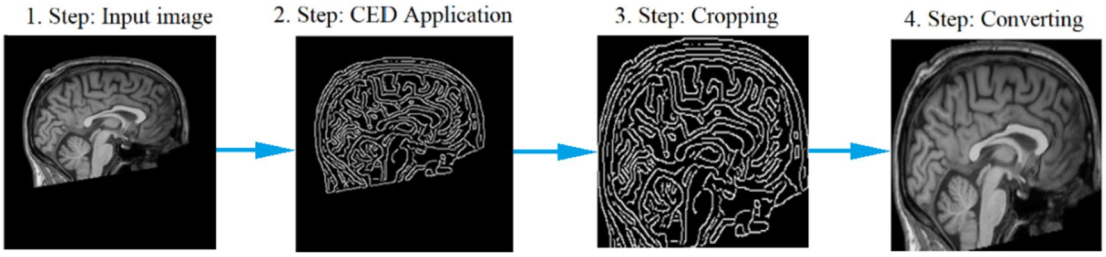

Fig. 2 The second step of the data pre-processing

$$y_r(x) = \frac{\exp\left[a_r(x)\right]}{\sum_{j=1}^{k}\exp\left[a_j(x)\right]} \tag{3}$$

In Eq. (3); $0 \le y_r \le 1$, $\sum_{j=1}^{k} y_j = 1$, and a_r is the conditional probability of the given r class sample [73].

Designing processing of the proposed models

Three multiple classifications were performed using the system designed within the scope of this study. First, the acquired brain sMRI image data were pre-processed. After pre-processing, the data were divided into three separate data sets, taking into account age, gender, and both. Grid search optimization (GSO) algorithm was utilized to design the CNN models to be trained with these three separate data sets from scratch to achieve optimal hyperparameters and the highest accuracy rate. The flow diagram of the designed system is illustrated in Fig. 5.

In the study, the dimensions of the input data matrix is chosen as 224 × 224 regardless of any criterion. Input image sizes are not contained in the GSO algorithm. The study utilized Model 1, Model 2, and Model 3 CNN model names for Data 1, Data 2, and Data 3, respectively. Table 4 shows the hyperparameters decided due to the GSO for each model and the structures of the CNN models thus designed. Figure 6 shows the architectural scheme in which Model 3 is utilized in the developed system.

Performance metrics

Utilizing the loss function shown in Eq. (4), the network continues to be trained throughout the training processing of the network until the loss values calculated for each iteration reach their minimum value.

$$Loss = -\sum_{i=1}^{N}\sum_{j=1}^{K} t_{ij} \ln y_{ij} \tag{4}$$

In Eq. (4), we have N samples represented by t_{ij}, where each sample I belongs to one of the K classes, and the corresponding output y_{ij} is assigned to sample j of class I [73]. Equation (5) shows the accuracy rate as another performance criterion [73].

$$Accuracy = \frac{Total\ correct\ prediction\ labels}{Total\ Equation\ Number\ of\ real\ labels} \times 100 \tag{5}$$

Experimental results

In this study, a system was developed that can contribute to ASD automatic diagnosis. CNN models, a part of this system, are designed to have the most optimal parameters through the GSO algorithm. Within this system, three CNN models were designed, and multiple classifications were performed to view the role of gender-age factors

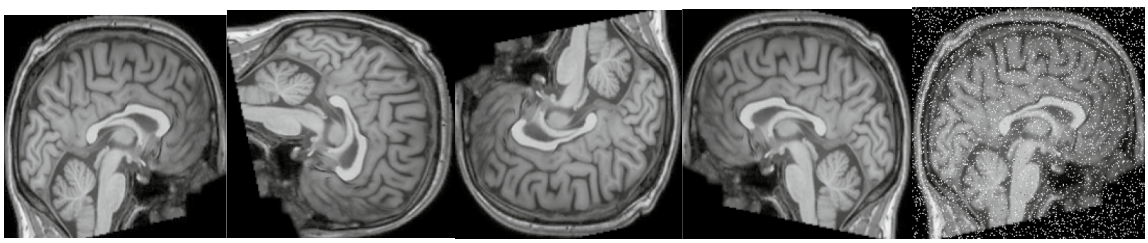

Fig. 3 **a** The raw image, **b** rotating by 90 degrees, **c** rotating by 180 degrees. **d** right/left flip, **e** 5% salting to the image

Table 3 Hyper-parameters and value ranges

	Hyper-parameters to optimize	Value ranges
1	Number of Convoluiton layer	[1, 2, 3, 4, 5, 6, 7, 8]
2	Number of Maxpooling layer	[1, 2, 3, 4, 5, 6, 7, 8]
3	Number of FC layers	[1, 2, 3, 4]
4	Number of filters	[16, 24, 32, 48, 64, 96]
5	Filter sizes for conv and pooling	[2, 3, 4, 5, 6, 7]
6	Padding	[0, 1, Same]
7	Stride	[1, 2, 3]
8	L2 regularization	[0.0001, 0.0005, 0.001, 0.005]
9	Momentum	[0.70, 0.75, 0.80, 0.85, 0.9, 0.95]
10	Mini-batch size	[8, 16, 32, 64, 128]
11	Learning rate	[0.0001, 0.0003, 0.0005, 0.001, 0.003, 0.005]
12	Activation function	ReLu, Leaky Relu, ELU, SELU

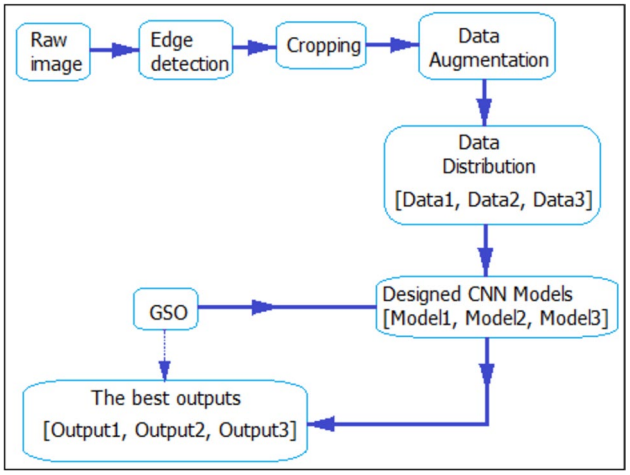

Fig. 5 Flow chart of the proposed system

in diagnosing ASD. Gender with Model 1, age with Model 2, and both with Model 3 were highlighted. In addition, the developed system was compared to four pre-trained networks using TL. The accuracy and loss curves acquired for all three models utilized in the designed system in Fig. 7, the confusion matrices in Fig. 8, and the comparison of the results with the pre-trained networks in Table 5 are presented. According to the results given in Table 5, the accuracy rate obtained in the quadruple classification made with Model 1, which highlights the gender factor, is 80.94% and is higher than all pre-trained models designed for the same purpose. It is seen that an accuracy rate of 85.42% was achieved in the quadruple classification made with Model 2, which highlights the age factor. It is seen that the accuracy rates obtained with pre-trained models designed for the same purpose are higher than all of them. This result leads us to think that the age factor has a greater impact on the diagnosis of ASD than the gender factor. An accuracy rate of 67.94% was achieved with the octahedral classification model made with Model 3, which takes both age and gender factors into consideration. Although it seems to be less successful

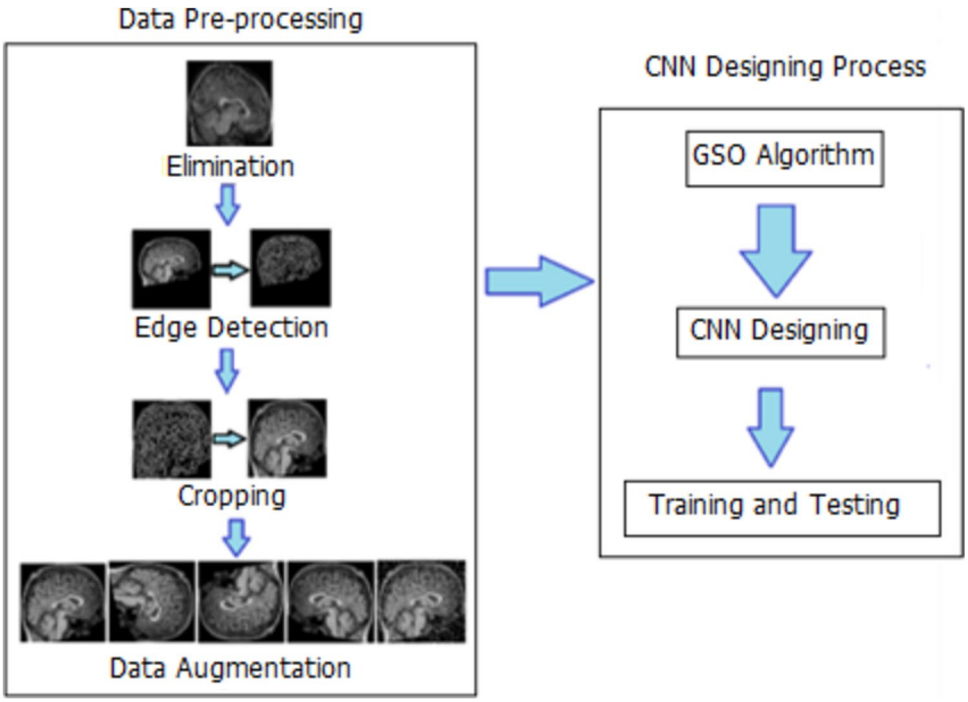

Fig. 4 Schematic representation of the proposed system

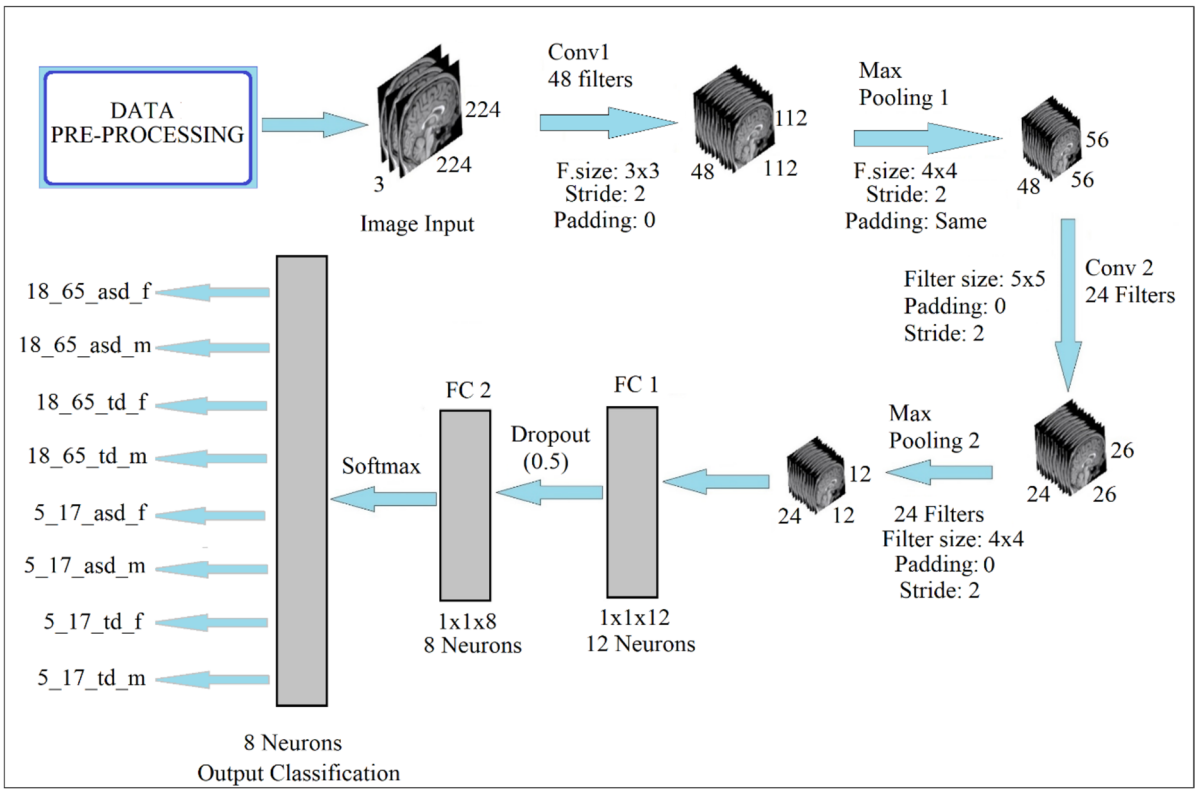

Fig. 6 The architecture of the designed Model 3

than Model 1 and Model 2, this result is quite high for the eight-fold classification. When compared to other eight-class classification pre-trained models implemented for the same purpose, it is seen that the highest result is obtained with Model 3. Similar comments can be made by examining missing values. With this system designed

Table 4 Optimal hyper-parameters of the proposed CNN models

		Value		
	Hyper-parameters	**Model 1 (Gender)**	**Model 2 (Age)**	**Model 3 (Both)**
1	Number of Convoluiton layer	7	6	2
2	Number of Maxpooling layer	7	6	2
3	Number of FC layers	2	2	2
4	Number of filters [Conv1, Pool1, Conv2, Pool2, Conv3, Pool3, Conv4, Pool4, Conv 5, Pool 5, Conv 6, Pool6, Conv7, Pool7…]	[48, 48, 24, 24, 24, 24, 24, 24, 16 16,16, 16, 16, 16]	[48, 48, 32, 32, 24, 24, 24, 24, 16, 16, 16, 16]	[48, 48, 24, 24]
5	Filter sizes [Conv1, Pool1, Conv2, Pool2, Conv3, Pool3, Conv4, Pool4, Conv 5, Pool 5, Conv 6, Pool6, Conv7, Pool7 …]	[3, 4, 3, 3, 5, 4, 5, 5, 3, 3, 3, 3, 5, 4]	[4, 3, 4, 3, 5, 4, 3, 5, 4, 4, 4, 4]	[2, 4, 5, 4]
6	Padding [Conv1, Pool1, Conv2, Pool2, Conv3, Pool3, Conv4, Pool4, Conv 5, Pool 5, Conv 6, Pool6, Conv7, Pool7 …]	[0, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1]	[0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0]	[0, same, 0, 0]
7	Stride [Conv1, Pool1, Conv2, Pool2, Conv3, Pool3, Conv4, Pool4, Conv 5, Pool 5, Conv 6, Pool6, Conv7, Pool7 …]	[1, 1, 2, 1, 1, 1, 1, 1, 2, 2, 2, 1, 2, 1]	[1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2]	[2, 2, 2, 2]
8	L2 regularization	0.0001	0.0001	0.0001
9	Momentum	0.9000	0.9000	0.9000
10	Mini-batch size	32	32	32
11	Learning rate	0.0001	0.0001	0.0002
12	Activation function	ReLu	ReLu	ReLu

Table 5 Test results for all three classifications and comparison with pre-trained models

CNN Models	Model 1 (Gender)		Model 2 (Age)		Model 3 (Both)	
	Accuracy (%)	Loss	Accuracy (%)	Loss	Accuracy (%)	Loss
Alexnet	78.13	0.9126	81.09	0.6631	62.73	1.8453
Googlenet	77.68	1.2609	78.59	1.0999	59.32	1.9831
Resnet18	73.12	1.2083	82.46	0.8289	66.36	1.3508
Squeezenet	72.67	1.3960	79.27	1.0946	64.55	1.7776
Proposed	**80.94**	0.4893	**85.42**	0.3785	**67.94**	0.8418

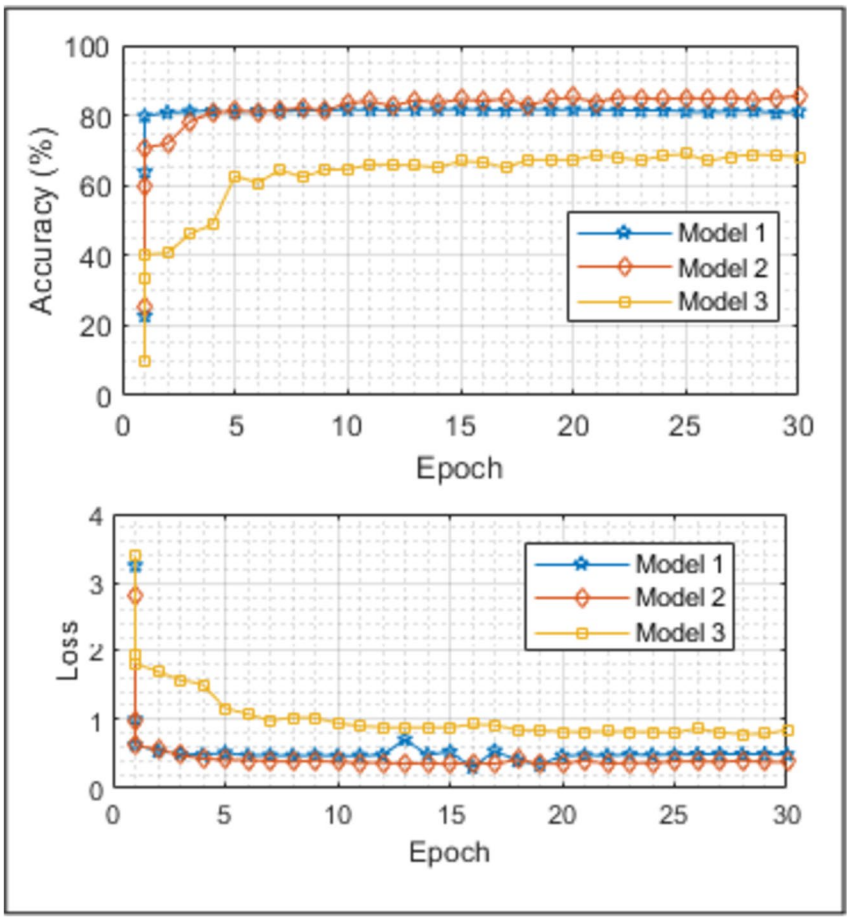

Fig. 7 Accuracy and Loss curves for all three classifications

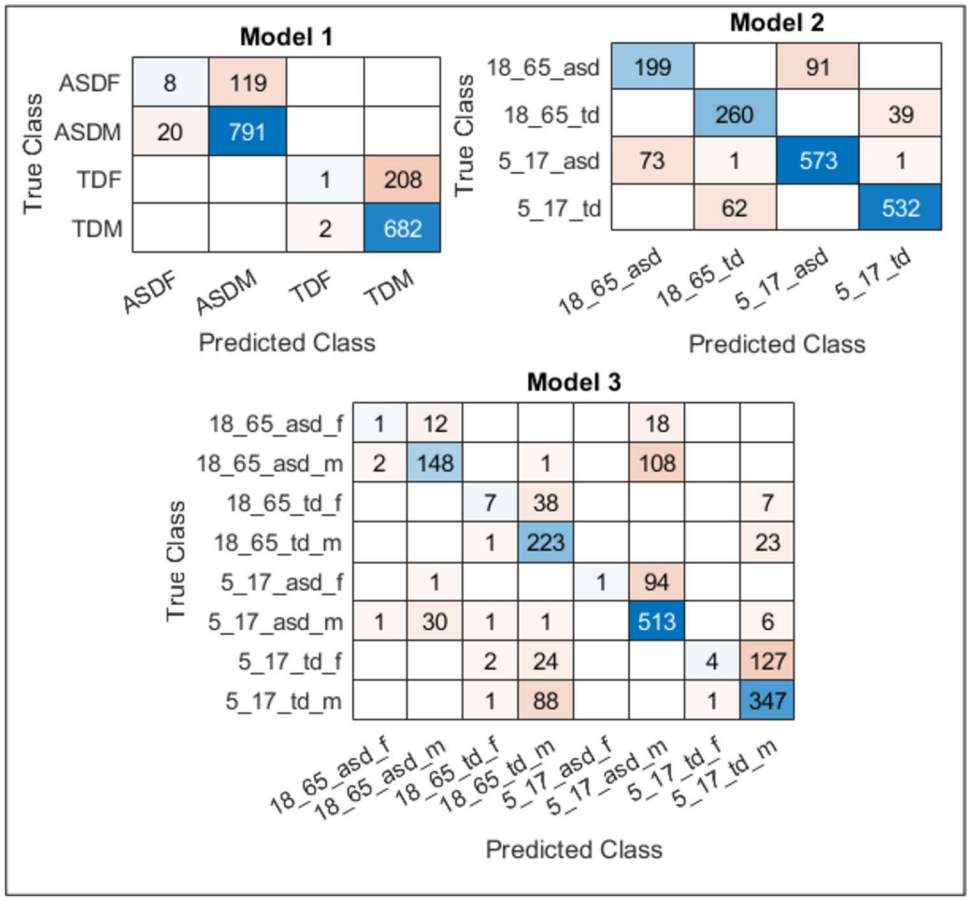

Fig. 8 Confusion matrices for all three models

for ASD classification and diagnosis, it is seen that the effect of gender and age factors in multiple classification emerges. The results showed that all three networks outperformed the pre-trained models. In the diagnosis of ASD, the influence of the age factor seems to be more than the gender factor. With this system designed for ASD classification and diagnosis, it is seen that the influence of gender and age factor in multi-classification is revealed.

Conclusions and discussion

All studies on automatic diagnosis of ASD with artificial intelligence are binary classification studies. An octal classification study that takes both age and gender factors into consideration cannot be found in the literature. In this study, a deep learning system that is different from what has been done so far and is a first, as far as we know, makes quadruple and eight-fold classifications by taking age and gender factors into account and uses sMRI brain images. In this study, an estimation and classification system was designed, which, as far as we know, is different from what has been done so far, and which is a first, takes into account age and gender factors and utilizes sMRI brain images. The success and reliability of the designed system were provided by comparing it with the Alexnet, Googlenet, Resnet-18, and Squeezenet popular pre-trained networks. The model developed in this research performs better than these pre-trained models. In addition, the designed system has the feature of generalizability since the data set was acquired from the ABIDE database created by acquiring from 29 different locations, and the data set was enlarged five times by DA techniques. As a result, the accuracy rates acquired as a result of the test performed with all three CNN models designed to be utilized within the system show that the designed system has robust dynamics enough to give the highest accuracy rates.

In the future, more successful applications are planned by adding advanced (ML) algorithms like enhanced probabilistic neural network (EPNN) [74] and neural dynamic classification (NDC) algorithm [75] to a system that examines age and gender factors.

Appendix 1

SITE	DEFINITON
STANDFORD	Stanford University
KKI	Kennedy Krieger Institute
KUL	Katholieke Universiteit Leuven
LEUVEN	University of Leuven
UCD	University of California Davis
OHSU	Oregon Health and Science University
MAXMUN	Ludwig Maximilians University Munich
UCLA	University of California, Los Angeles
BNI	Barrow Neurological Institute
CALTECH	California Institute of Technology
EMC	Erasmus University Medical Center Rotterdam
GU	Georgetown University
IP	Institut Pasteur and Robert Debré Hospital
NYU	New York University Langone Medical Center
PITT	University of Pittsburgh
SDSU	San Diego State University
TRINITY	Trinity Centre for Health Sciences
UM	University of Michigan
UPSM	University of Pittsburgh School of Medicine
YALE	Yale Child Study Center
SU	Stanford University (ABIDE II)
OLIN	Olin Neuropsychiatry Research Center (ABIDE I)
ETH	Eidgenössische Technische Hochschule Zürich
TCD	Trinity College Dublin's School of Medicine
IU	Indiana University
ONRC	Olin Neuropsychiatry Research Center (ABIDE II)
USM	University of Utah School of Medicine
CMU	Carnegie Mellon University
SBL	Social Brain Lab. BCN NIC UMC Groningen and Netherlands Institute for Neurosciences

Author Contributions H.A. guided the overall direction of the research, and participated in the development of the algorithms and writing of the article.H.N. implemented the algorithms, obtained results, and prepared the first draft of the paper.

Funding None.

Data Availability The ABIDE database, a publicly available database was used in this research.

Declarations

Ethical Approval Not applicable.

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Preet K, Shoba S, Shekhar PS, Satish CG, and John VSK (2017) Lost time: Need for more awareness in early intervention of autism spectrum disorder. Asian Journal of Psychiatry 25:13–15
2. Constantino JN, Charman T (2016) Diagnosis of autism spectrum disorder: reconciling the syndrome, its diverse origins, and variation in expression www.thelancet.com/neurology 15
3. Maenner, M.J., Shaw, K.A., Bakian, A.V., Bilder, D.A., Durkin, M.S., Esler, A.N., Furnier, S.M., Hallas, L., Hall-Lande, J., Hudson, A., Hughes, M.M., Patrick, M.E., Pierce, K.J., Poynter, J.N., Salinas, A., Shenouda, J., Vehorn, A.C., Warren, Z., Constantino, J.N., Dirienzo, M., Fitzgerald, R.T., Grzybowski, A., Spivey, M.H., Pettygrove, S., Zahorodny, W., Ali, A.R., Andrews, J.G., Baroud, T., Gutierrez, J., Hewitt, A.S., Lee, L.C., Lopez, M.L., Mancilla, K.C., McArthur, D., Schwenk, Y.D., Washington, A., Williams, S., & Cogswell, M.E. (2021). Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018. MMWR Surveillance Summaries 70: 1–16.
4. Senn, M. CDC Estimate on Autism Prevalence Increases by Nearly 10%, to 1 in 54 Children in the U.S. (2020) Available online at: https://www.autismspeaks.org/press-release/cdc-estimate-autism-prevalence-increasesnearly-10-percent-1-54-children-us. Accessed April 26, 2020.
5. Jacob, S., Wolff, J.J., Steinbach, M.S., Doyle, C.B., Kumar, V., and Elison, J.T. (2019). Neurodevelopmental heterogeneity and computational approaches for understanding autism. Translational Psychiatry 9.
6. Matson, J. L., Rieske R. D., Williams, L.W (2013) The relationship between autism spectrum disorders and attention-deficit/hyperactivity disorder: an overview. Research in Developmental Disabilities 34: 2475–2484.
7. Saito, M., Hirota, T., Sakamoto, Y., Adachi, M., Takahashi, M., Osato-Kaneda, A., Kim, Y.S., Leventhal, B.L., Shui, A.M., Kato, S., and Nakamura, K. (2020) Prevalence and cumulative incidence of autism spectrum disorders and the patterns of co-occurring neurodevelopmental disorders in a total population sample of 5-year-old children. Molecular Autism 11, 2020.

8. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2002 Principal Investigators (2007) Prevalence of autism spectrum disorders—Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2002. MMWR Surveill Summ 56: No. SS–1
9. Napolitano A, Schiavi S, La Rosa P, Rossi-Espagnet MC, Petrillo S, Bottino F, Tagliente E, Longo D, Lupi E, Casula L, Valeri G, Piemonte F, Trezza V and Vicari S. Sex (2022) Differences in Autism Spectrum Disorder: Diagnostic, Neurobiological, and Behavioral Features. Front. Psychiatry 13: 889636. doi: https://doi.org/10.3389/fpsyt.2022.889636
10. Horlin C., Falkmer M., Parsons R., Albrecht MA, Falkmer T (2014) The cost of autism spectrum disorders. PloS One 9 e106552
11. Klin A., Klaiman, C., Jones, W (2015) Reducing age of autism diagnosis: developmental social neuroscience meets public health challenge. Revista de Neurologia 60(Suppl 1): 3–11
12. Subbaraju V, Sundaram S, Narasimham S, Suresh, MB (2015) Accurate detection of autism spectrum disorder from structural MRI using extended metacognitive radial basis function network. Expert Systems with Applications 42: 8775–8790, 2015.
13. Lemm S, Blankertz B, Dickhaus T, Müller KR (2011) Introduction to machine learning for brain imaging. NeuroImage 56: 387–399
14. Hastie T, Tibshirani R, Friedman J (2009) The Elements of Statistical Learning. 2nd edition, Springer, New York.
15. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC Pickles A., Rutter M (2000) The autism diagnostic observation schedule-generic a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 30: 205–223
16. Lord C, Rutter M, Le Couteur A (1994) Autism diagnostic interview-revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 24: 659–685
17. Lord C, Petkova E, Hus V, Gan W, Lu F, Martin DM Ousley O, Guy L, Bernier R, Gerdts J. et al (2012) A multisite study of the clinical diagnosis of different autism spectrum disorders. Archives of General Psychiatry 69: 306–313
18. O'Shea A, Ahmed R, Lightbody G, Mathieson SR, Pavlidis E, Lloyd RO Pisani F, Marnane WP, Boylan GB, & Temko A (2020) Deep Learning for EEG Seizure Detection in Preterm Infants. International journal of neural systems 2150008
19. Peng P, Xie L, & Wei H (2021) A Deep Fourier Neural Network for Seizure Prediction Using Convolutional Neural Network and Ratios of Spectral Power. International journal of neural systems 2150022
20. Yu Z, Albera L, Bouquin-Jeannès RL, Kachenoura A, Karfoul A, Yang C, & Shu H (2022) Epileptic Seizure Prediction Using Deep Neural Networks Via Transfer Learning and Multi-Feature Fusion. International journal of neural systems 2250032
21. Bhattacharya A, Baweja T, & Karri SP (2021) Epileptic Seizure Prediction Using Deep Transformer Model. International journal of neural systems 2150058
22. Thangavel P, Thomas J, Peh WY, Jing J, Yuvaraj R, Cash SS, Chaudhari R, Karia S, Rathakrishnan R, Saini V, Shah N, Srivastava R, Tan Y, Westover B, & Dauwels J (2021) Time-Frequency Decomposition of Scalp Electroencephalograms Improves Deep Learning-Based Epilepsy Diagnosis. International journal of neural systems 2150032
23. Ozdemir M, Cura OK, and Akan A (2021) Epileptic EEG Classification by Using Time-Frequency Images for Deep Learning. International journal of neural systems 2150026.
24. Bone D, Goodwin MS, Black MP, Lee C, Audhkhasi K, and Narayanan S (2015) Applying Machine Learning to Facilitate Autism Diagnostics: Pitfalls and promises. Journal of Autism Developmental Disorders 45(5):1121–1136. https://doi.org/10.1007/s10803-014-2268-6
25. Ardakani HA, Taghizadeh M, and Shayegh F (2022) Diagnosis of Autism Disorder Based on Deep Network Trained by Augmented EEG Signals. International journal of neural systems 2250046.
26. Li S, Tang Z, Jin N, Yang Q, Liu G, Liu T, Hu J, Liu S, Wang P, Hao J, Zhang Z, Zhang X, Li J, Wang X, Li Z, Wang Y, Yang B, and Ma L (2022) Uncovering Brain Differences in Preschoolers and Young Adolescents with Autism Spectrum Disorder Using Deep Learning International journal of neural systems 2250044.
27. Rahman MM, Usman OL, Muniyandi RC, Sahran S, Mohamed S, & Razak RA (2020) A Review of Machine Learning Methods of Feature Selection and Classification for Autism Spectrum Disorder. Brain Sciences 10
28. Küçükoglu B, Rueckauer B, Ahmad N, de Ruyter van Steveninck J, Güçlü U, & van Gerven MA (2022) Optimization of Neuroprosthetic Vision via End-to-end Deep Reinforcement Learning. bioRxiv 32(11): 2250052 (16 pages)
29. Xu F, Dong G, Li J, Yang Q, Wang L, Zhao Y, Yan Y, Zhao J, Pang S, Guo D, Zhang Y, & Leng J (2022) Deep Convolution Generative Adversarial Network-Based Electroencephalogram Data Augmentation for Post-Stroke Rehabilitation with Motor Imagery. International journal of neural systems 32(9): 2250039, (15 pages)
30. Alexandridis G, Aliprantis J, Michalakis K, Korovesis K, Tsantilas P, & Caridakis G (2021) A Knowledge-Based Deep Learning Architecture for Aspect-Based Sentiment Analysis. International journal of neural systems 31(10): 2150046 (18 pages)
31. Olamat A, Ozel P, and Atasever S (2022) Deep Learning Methods for Multi-Channel EEG-based Emotion Recognition. International Journal of Neural Systems 32(5): 2250021 (17 pages)
32. Lope JD, & Graña M, (2022) A Hybrid Time-Distributed Deep Neural Architecture for Speech Emotion Recognition. International journal of neural systems 32(6): 2250024 (14 pages)
33. Hu T, Xie L, Zhang L, Li G, & Yi Z (2021) Deep Multimodal Neural Network Based on Data-Feature Fusion for Patient-Specific Quality Assurance. International journal of neural systems 32 (1): 2150055
34. Hashimoto H, Kameda S, Maezawa H, Oshino S, Tani N, Khoo HM, Yanagisawa T, Yoshimine T, Kishima H, & Hirata M (2020) A Swallowing Decoder Based on Deep Transfer Learning: AlexNet Classification of the Intracranial Electrocorticogram. International journal of neural systems 31(11): 2150056 (16 pages)
35. Ieracitano C, Morabito FC, Hussain A, & Mammone N (2021) A Hybrid-Domain Deep Learning-Based BCI For Discriminating Hand Motion Planning From EEG Sources. International journal of neural systems 31(9): 2150038 (20 pages)
36. Usman OL, Muniyandi RC, Omar KB, & Mohamad M (2021) Advance Machine Learning Methods for Dyslexia Biomarker Detection: A Review of Implementation Details and Challenges. IEEE Access 9 36879–36897
37. Usman OL, & Muniyandi RC (2020) CryptoDL: Predicting Dyslexia Biomarkers from Encrypted Neuroimaging Dataset Using Energy-Efficient Residue Number System and Deep Convolutional Neural Network. Symmetry 12, 836
38. Usman OL, Muniyandi RC, Omar KB, and Mohamad M (2021) Gaussian smoothing and modified histogram normalization methods to improve neural-biomarker interpretations for dyslexia classification mechanism. PLoS ONE 16.
39. Macias-Garcia E, Galeana-Perez D, Medrano-Hermosillo J, and Bayro-Corrochano E (2021) Multi-stage Deep Learning Perception System for Mobile Robots. Integrated Computer-Aided Engineering 28(2): 191–205
40. Gasienica-Józkowy J, Knapik M, and Cyganek B (2021) An ensemble deep learning method with optimized weights for

drone-based water rescue and surveillance. Integrated Computer-Aided Engineering 28(3): 221–235
41. Jiang K, Han Q, and Du X (2022) Lost Data Neural Semantic Recovery Framework for Structural Health Monitoring Based on Deep Learning. Computer-Aided Civil and Infrastructure Engineering 37(9): 1160–1187
42. Pan X, and Yang TT (2022) Image-based monitoring of bolt loosening through deep-learning-based integrated detection and tracking. Computer-Aided Civil and Infrastructure Engineering 37(10): 1207–1222
43. Chun PJ, Yamane T, and Maemura Y (2022) A deep learning based image captioning method to automatically generate comprehensive explanations of bridge damage. Computer-Aided Civil and Infrastructure Engineering 37(11): 1387–1401
44. Rad NM, and Furlanello C (2016) Applying Deep Learning to Stereotypical Motor Movement Detection in Autism Spectrum Disorders. 2016 IEEE 16th International Conference on Data Mining Workshops 16: 2375–9259, DOI https://doi.org/10.1109/ICDMW.2016.184
45. Rad NM, Kia MS, Zarbo C, Laarhoven TV, Jurman G, Venuti P, Marchiori E, Furlanello C (2018) Deep Learning for Automatic Stereotypical Motor Movement Detection using Wearable Sensors in Autism Spectrum Disorder. Signal Processing 144: 180–191
46. Wang S, Jiang M, Duchesne XMM, Laugeson EAA, Kennedy DPP, Adolphs R, and Zhao Q (2015) Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. Neuron 88(3): 604–616
47. Milièiæ J, Petkoviæ BZ, Boikov J (2003) Dermatoglyphs of Digito-Palmar Complex in Autistic Disorder: Family Analysis. Clinical Sciences 44(4): 469–476
48. Kazemi M, Bordbar MRF, Shahri NM (2017) Comparative Dermatoglyphic Study between Autistic Patients and Normal People in Iran. Iran J Med Sci 42(4)
49. Stošljeviü M, and Adamoviü M, (2013) Dermatoglyphic characteristics of digito-palmar complex in autistic boys in Serbia. Vojnosanit Pregl 70(4): 386–390
50. Peng G, Nourani M, Harvey J, & Dave H, (2021) Personalized EEG Feature Selection for Low-Complexity Seizure Monitoring. International journal of neural systems 31(8): 2150018 (16 pages)
51. Xue Y, Zhu H, and Neri F (2022) A Self-adaptive Multi-objective Genetic Algorithm for Feature Selection in Classification. Integrated Computer-Aided Engineering 29(1): 3–21
52. Tian Y, Ding X, Lin YF, Ma S, and Li L (2022) Automatic Feature Type Selection Network in Digital Photogrammetry of Piping. Computer-Aided Civil and Infrastructure Engineering 37(10): 1335–1348
53. Ad-Dab'bagh Y, Lyttelton O, Muehlboeck JS, et al (2006) The CIVET image processing environment: a fully automated comprehensive pipeline for anatomical neuroimaging research. In: Proceedings of the 12th Annual Meeting of the Organization for Human Brain Mapping Florence Italy 2266
54. Fischl B, Dale AM (2000) Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci 97: 11050–11055
55. Yahata N, Kasai K, and Kawato M (2017) Computational neuroscience approach to biomarkers and treatments for mental disorders. Psychiatry and Clinical Neurosciences 71: 215–237
56. Ashburner J, and Friston KJ (2000) Voxel-based morphometry—the techniques. NeuroImage 11: 805–821
57. Mandl RC, Schnack HG, Zwiers MP, van der Schaaf A, Kahn RS, Pol HEH (2008) Functional diffusion tensor imaging: measuring task-related fractional anisotropy changes in the human brain along white matter tracts. PloS One 3 e3631
58. Iturria-Medina Y, Canales-Rodriguez EJ, Melie-Garcia L, Valdes-Hernandez PA, Martinez-Montes E, Alemán-Gómez Y, Sánchez-Bornot JM (2007) Characterizing brain anatomical connections using diffusion weighted MRI and graph theory. NeuroImage 36: 645–660
59. Iturria-Medina Y (2013) Anatomical brain networks on the prediction of abnormal brain states. Brain connectivity 3: 1–21
60. Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. NeuroImage 52: 1059–1069
61. Zeighami Y, Ulla M, Iturria-Medina Y, Dadar M, Zhang Y, Larcher KMH, Fonov V, Evans AC, Collins DL, Dagher A (2015) Network structure of brain atrophy in de novo Parkinson's disease. Elife 4 e08440. https://doi.org/10.7554/eLife.08440
62. Pereira F, Mitchell T, Botvinick M (2009) Machine learning classifiers and fMRI: a tutorial overview. NeuroImage 45: 199–209
63. Mirzaei G, and Adeli H (2016) Resting State Functional Magnetic Resonance Image Processing Techniques in Stroke Studies. Reviews in the Neurosciences 27(8): 871–885
64. Hassanpour A, Moradikia M, Adeli H, Khayami SR, and Shamsinejad P (2019) A Novel End-to-End Deep Learning Scheme For Classifying Multiclass Motor Imagery EEG Signals. Expert Systems 36(6).
65. Gorriz JM, Ram´ırez J, Segovia F, Mart´ınez FJ, Lai MC, Lombardo MV, Baron-Cohen S, and Suckling J (2019) A Machine Learning Approach to Reveal the NeuroPhenotypes of Autisms. Int J Neural Sys 29(7)
66. Heinsfeld AS, Franco AR, Craddock C, Buchweitz A, & Meneguzzi F (2018) Identification of autism spectrum disorder using deep learning and the ABIDE dataset. NeuroImage: Clinical 17: 16–23. https://doi.org/10.1016/j.nicl.2017.08.017
67. Li G, Liu M, Sun Q, Shen D, Wang L (2018) Early Diagnosis of Autism Disease by Multi-channel CNNs. Mach Learn Med Imaging 11046, 303–309. https://doi.org/10.1007/978-3-030-00919-9_35
68. Aghdam MA, Sharifi A, & Pedram MM (2018) Combination of rs-fMRI and sMRI Data to Discriminate Autism Spectrum Disorders in Young Children Using Deep Belief Network. Journal of Digital Imaging 31: 895–903
69. Kong Y, Gao J, Xu Y, Pan Y, Wang J, Liu J (2019) Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier. Neurocomputing 324: 63–68
70. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, Anderson JS, Assaf M, Bookheimer SY, Dapretto M et al. (2014) The Autism Brain Imaging Data Exchange: towards large-scale evaluation of the intrinsic brain architecture in autism. Molecular Psychiatry 19(6): 659–667 https://doi.org/10.1038/mp.2013.78
71. Irmak E, (2021) Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework. Iranian Journal of Science and Technology Transactions of Electrical Engineering 45:1015–1036
72. Michelucci U (2019) Advanced Applied Deep Learning: Convolutional Neural Networks and Object Detection, 1st ed. Apress
73. Bishop CM (2006) Pattern Recognition and Machine Learning. Springer New York NY
74. Ahmadlou M, & Adeli H (2006) Enhanced probabilistic neural network with local decision circles: A robust classifier. Integr Comput Aided Eng 17: 197–210
75. Rafiei MH and Adeli H (2017) A New Neural Dynamic Classification Algorithm," IEEE Transactions on Neural Networks and Learning Systems, 28(12): 3074–3083 https://doi.org/10.1109/TNNLS.2017.2682102

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.