Performance Evaluation of Machine Learning Techniques Applied to Magnetic Resonance Imaging of Individuals with Autism Spectrum Disorder

V. F. Carvalho, G. F Valadão, S. T. Faceroli, F. S Amaral and M. Rodrigues

Abstract

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder, affected by persistent deficits in communication and social interaction and by restricted and repetitive patterns of behavior, interests or activities. Its diagnosis is still a challenge due to the diversity between the manifestations of autistic symptoms, requiring interdisciplinary assessments. This work aims to investigate the performance of the application of techniques of extraction of characteristics and machine learning in magnetic resonance imaging (MRI), in the classification of individuals with ASD. In MRI, the techniques of features extraction were applied: histogram, histogram of oriented gradient and local binary pattern. These features were used to compose the input data of the Support Vector Machine and Artificial Neural Network algorithms. The best result shows an accuracy percentage of 89.66 and a false negative rate of 6.89%. The results obtained suggest that magnetic resonance analysis can contribute to the diagnosis of ASD from the advances in studies in the area.

Keywords

Autism spectrum disorder • Magnetic resonance imaging • Features extraction • Support vector machine • Artificial neural network

1 Introduction

The Autism Spectrum Disorder (ASD) is a disorder in the neurodevelopment, defined by persistent deficits in social communication and social interaction in multiple contexts. It shows restrict and repetitive patterns of behavior, interests or activities, with early symptoms in the period of development, causing losses at the social functioning of the individual's life [1]. The main features are related to language delay, difficulties at comprehension, echolalic speech, use of literal language with low or any social initiative. These symptoms begins since childhood, causing limitations and daily life losses [2].

Because it is a developmental disorder defined from a behavioral point of view, with multiple etiologies and many degrees of severity, its diagnosis is quite complex [3,4]. The criteria currently used to diagnose autism are those described in the Diagnostic and Statistical Manual of Mental Disorders (DSM) [5]. Currently it is used the DSM-V with some changes compared to the previous one [6].

The ASD precise etiopathogenesis is yet not proven, but some works suggest that structural cerebral regions may show alterations including the frontal lobes, amygdala, cerebellum [7], corpus callosum [8] and basal ganglia [9]. These results allow the possibility of helping the autism diagnosis using brains scans. Magnetic resonance imaging has been studied for some years with this objective, as shown by [10–12].

The magnetic resonance imaging is a versatile technique that obtains medical images that has the ability to demonstrate different brains structures and their minimal changes [13]. Recent studies suggest that magnetic resonance imaging (MRI), when analysed by machine learning techniques, could help at ASD diagnosis [14].

In order to contribute on the evaluation of application of MRI as an auxiliary scan in the diagnosis of autism, this paper presents a performance comparison among several image processing techniques, based on machine learning, applied to classification and identification of individuals with ASD.

2 Materials and Methods

A system was created in Python language, version 3.6, which follows the methodology presented in Fig. 1. At first, a database that could provide quality and reliable magnetic resonance imaging, both from patients diagnosed with

V. F. Carvalho · G. F. Valadão (✉) · S. T. Faceroli · F. S. Amaral · M. Rodrigues
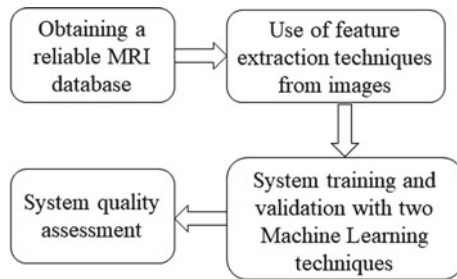Mechatronics Engineering, IF SUDESTE MG, Juiz de Fora, Brazil

Fig. 1 Methodology workflow

ASD and from patients in a neurotypic control group, was searched. After obtaining the images, its characteristics were extracted by means of three different techniques. The data was separated in 70% of patients for training and 30% patients for testing, in other words, 66 patients for training and 29 for testing for each slice, in this way there are not reused patients in train and test. The characteristics were used to train, with training data, two different types of predictive models of machine learning. In order to validate the systems, the models were applied to the images intended for validation. Finally, evaluation metrics were used to validate and compare these systems.

2.1 Database

MRI data analyzed in this work were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu).[1] The ADNI was launched in 2003 as a public-private partnership, led by principal investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD).

Autism Brain Imaging Data Exchange (ABIDE), that belongs to ADNI, were the database used in the development of this work. It has 1112 patients, which 539 diagnosed with ASD and 573 control patients [15].

---

[1]Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Each patient has several MRI, separated by three cuts (axial, coronal and sagital). The first step in this project was to separate which patients would be selected for the study. Due to the number of patients and amount of images per patient, it would be impracticable to use all image set. Besides, there are the registers of different number of images from some patients. In order to deal with these issues, the largest group of patients, from the original database, who has the same number of images registered for each one of the cuts, was considered. This group has 55 patients diagnosed with ASD and 40 control patients, which yields a 95-patient selected database.

All patients have 512 images in the axial cut, 480 in coronal cut and 160 in sagital cut, all these images were made available, by ABIDE, in its 2D form. However, the beginning and the end of each plane have images that do not have any brain information. Therefore, these portions of each plane were discarded. Regarding the axial plane, the selected slices were from 150 to 470, while slices from 30 to 400 were considered for the coronal plane. Sagital plane has the selected slices from 20 to 140.

2.2 Feature Extraction Techniques

Three feature extraction techniques were used: histogram, histogram of oriented gradients (HOG) and Local binary pattern (LBP). These extractions were made slice by slice, in other words, for each slice were used the presented techniques. This procedure has been done for all 840 utilized slice.

An image histogram describes the frequency of the intensity values that occur on it. In an 8-bit gray scale image, for example, the histogram result will have a frequency of 2^8 possible intensity values, ranging from 0, which would be equivalent to black, to 255, which would be equivalent to white. Therefore, a darker image would have a histogram concentrated at values closer to 0, while a lighter image would be at 255.

HOG is a simple and fast method, based on the idea that the appearance and shape of an object can be described by the directions of the edges or by the distribution of the local intensity gradients. This characterizer summarizes the distribution measurements in the regions of the image, being particularly useful for cases in which it is necessary to recognize the texture of objects [16].

LBP is a method of describing texture in the local neighborhood and can be considered as a gradient of binary direction. The LBP operator labels the pixels of an image by limiting the neighbors of each pixel by the central pixel value and displays the results binary [17].

2.3   Predictive Models of Machine Learning

To perform the classification, two techniques were used: the support vector machine (SVM) and the artificial neural network (ANN). Both techniques are algorithms that learn to identify patterns based on the information given to them and the labels of each piece of information and these techniques belong to supervised classification paradigm.

Support vector machines, proposed by Vapnik [18], is one of the most popular learning algorithms and can be used for both classification and regression problems from structured data. When used in regression problems the support vector technique is called Support Vector Regression (SVR). The SVM aims to find an efficient way to separate a high-dimensional space with hyperplanes. In this way, SVM training produces a function that minimizes training error while maximizes the margin that separates the data classes. The margin can be calculated as the perpendicular distance that separates the hyperplane and the generated hyperplanes from the nearest points [19]. In addition, these planes need not be only linear, they can vary based on the polynomial degrees of each one.

For ANN, the information is processed in computational cells, called artificial neurons, which relate the input data with the output tags. For the development of an ANN it is necessary to determine its architecture, that is, the number of neurons per layer and the number of layers in the network.

2.4   Training

The SVM training is based on hyperplanes. In this study, they were used based on their polynomial degrees. For each feature extraction technique, the degree varied from 1 to 15. For higher degree values, similar or worse results were observed. It was necessary to change the SVM kernel, to match the degrees variation. In other words, if the degree was 1, the kernel would be linear and, above that, the kernel would be polynomial. So, for each feature extraction technique, 15 results were generated for each slice. The other parameters were: decision function shape: one-vs-rest(ovr), gamma: scale.

Regarding the ANN training, the Multilayer Perceptron technique was used. The data were separated into 70% for training and 30% for testing. The structure of the network had 3 hidden layers and the number of neurons was varied from 3 to 5 per layer. In this way, it was possible to avoid overfitting in the neural network, because, with the increased neuron number, overfitting could happen. The other parameters were: solver: adam, alpha: 0.0001, learning rate: constant, learning rate initial: 0.001, tolerance: 0.0001, momentum: 0.9. The algorithm chosen was the back-propagation, which is an algorithm with supervised paradigm and is one of the most used and the most important in artificial neural networks. Its main advantage is that it works with multiple layers and solves nonlinearly separable problems. The activation function used for network training was the rectified linear unit (ReLU), as they are the most used and reduce the training time.

2.5   Metrics

In order to compare the results obtained by each training with the various characteristics, two metrics were used. The accuracy, which relates the number of correct answers in the prediction with the total number of images per slice, and the number of false negatives, which analyzes the number of patients diagnosed by ASD that the prediction considered as control. This metric was chosen because if some patients are not diagnosed with ASD but are false negatives they do not receive adequate assistance and may seek ineffective treatments or interrupt their search for treatments.

3   Results

The images obtained using the selected database are exemplified in Figs. 2 and 3, which represent the slice 366 of the axial plane for the group diagnosed with ASD and the control group, respectively.

Fig. 2 Slice 366 of the axial plane of a patient diagnosed with ASD
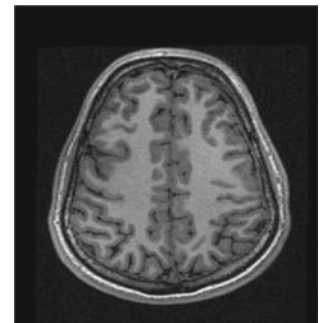

Fig. 3 Slice 366 of the axial plane of a control patient
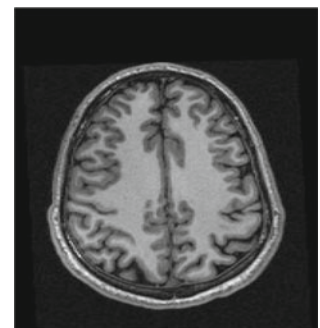
Table 1 presents the best results of the metrics used for the different simulated models. For each proposed extraction

technique, the two predictive models for each type of MRI plane were evaluated. These results were chosen as the best from all trained slices, utilizing each one of the extraction techniques. All the 840 slices were analyzed.

4   Discussion

According to the results summarized in Table 1, it is possible to verify that the HOG characteristic extraction technique was the one that presented, in the set of plans, the best results regarding the performance of the tested machine learning networks. Besides, the histogram technique also showed significant results.

In relation to the tested machine learning techniques, both presented a similar result regarding the precision of the algorithm. However, the artificial neural network had less false negatives compared to the SVM technique.

Evaluating the types of plane, made by MRI, the axial plane presented better results, both in relation to the accuracy of the network and the percentage of false negatives. In fact, the algorithm using the histogram technique with SVM showed 0% false negatives.

The best result obtained in the tests performed was using the axial plane in the MRI, extracting the characteristics by the HOG and using the SVM algorithm. The classifier had an accuracy of 89.66%, which indicates that the improvement of the method can lead to significant results for the use of magnetic resonance imaging exams to assist in the process of diagnosing people with ASD.

5   Conclusion

This scientific work proposed the analysis of the performance of techniques used in computer vision in MRI to classify individuals with ASD. The main contribution of the study is to evaluate the possibility of using magnetic resonance imaging to assist in the diagnosis of autism spectrum disorder, which is still a challenge due to the complexity of the disorder. Based on the results obtained, it is possible to observe that some proposed techniques had quite significant results, indicating that the magnetic resonance imaging of ASD patients may contain relevant information to aid the diagnosis.

The main contribution of this article, comparing with articles already done in this area, is the application of smart techniques for each slice in MRI. In this way, is possible to observe how the algorithms behave and each brain area, building an information system capable of evaluating the best techniques and the most relevant regions.

In this regard, future work can improve the techniques of image processing and machine learning, seeking to increase the accuracy of the classification. In addition, such research can provide very important indications such as the main brain locations of the characteristics of individuals with ASD and the variation of such characteristics with age and with stimuli. Besides that, it may be tested algorithms using Deep Learning. In this case, it is necessary to use data augmentation techniques for repository expansion, considering that the available databases for this type of study do not have a sufficient amount of data for using Deep Learning.

Table 1 Results

Extraction techniques	Predictive model	Plane	Slices	Accuracy (%)	False negatives (%)
HOG	SVM	Axial	366	89.66	6.89
		Coronal	398	86.21	13.79
		Sagital	120	75.86	20.68
	ANN	Axial	366	86.21	3.44
		Coronal	398	82.76	13.79
		Sagital	48	75.86	6.89
Histogram	SVM	Axial	177	86.21	0
		Coronal	398	86.21	13.79
		Sagital	120	75.86	20.68
	ANN	Axial	257	82.76	10.34
		Coronal	398	82.76	13.79
		Sagital	48	75.86	6.89
LBP	SVM	Axial	296	82.76	10.34
		Coronal	283	79.31	3.44
		Sagital	80	72.41	3.44
	ANN	Axial	191	79.31	17.24
		Coronal	287	82.76	10.34
		Sagital	63	76.31	17.24

References

1. Association American Psychiatric et al. (2013) Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub
2. Silva BS, Carrijo DT, Firmo JDR, Freire MQ, Pina MFÁ, Macedo J (2018) Dificuldade no diagnóstico precoce do transtorno do espectro autista e seu impacto no âmbito familiar. CIPEEX 2:1086–1098
3. Rutter M, Schopler E (1992) Classification of pervasive developmental disorders: some concepts and practical considerations. J Autism Dev Disord 22:459–482
4. Gadia CA, Tuchman R, Rotta NT (2004) Autismo e doenças invasivas de desenvolvimento. J Pediatr 80:83–94
5. American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, 5th edn
6. Lobar SL (2016) DSM-V changes for autism spectrum disorder (ASD): implications for diagnosis, management, and care coordination for children with ASDs. J Pediatr Health Care 30:359–365
7. Amaral DG, Schumann CM, Nordahl CW (2008) Neuroanatomy of autism. Trends Neurosci 31:137–145
8. Bellani M, Calderoni S, Muratori F, Brambilla P (2013) Brain anatomy of autism spectrum disorders I. Focus on corpus callosum. Epidemiol Psychiatr Sci 22:217–221
9. Calderoni S, Billeci L, Narzisi A, Brambilla P, Retico A, Muratori F (2016) Rehabilitative interventions and brain plasticity in autism spectrum disorders: focus on MRI-based studies. Front Neurosci 10:139
10. Berthier ML, Bayes A, Tolosa ES (1993) Magnetic resonance imaging in patients with concurrent Tourette's disorder and Asperger's syndrome. J Am Acad Child Adolesc Psychiatry 32:633–639
11. Piven J, Berthier ML, Starkstein SE, Nehme E, Pearlson G, Folstein S (1990) Magnetic resonance imaging evidence for a defect of cerebral cortical development in autism. Am J Psychiatry 147(6):734–739
12. Nowell MA, Hackney DB, Muraki AS, Coleman M (1990) Varied MR appearance of autism: fifty-three pediatric patients having the full autistic syndrome. Magn Reson Imag 8:811–816
13. Amaro JE, Yamashita H (2001) Aspectos básicos de tomografia computadorizada e ressonância magnética. Braz J Psychiatry 23:2–3
14. Pagnozzi AM, Conti E, Calderoni S, Fripp J, Rose SE (2018) A systematic review of structural MRI biomarkers in autism spectrum disorder: a machine learning perspective. Int J Dev Neurosci 71:68–82
15. Di Martino A, Yan C-G, Li Q et al (2014) The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry 19:659–667
16. Shu C, Ding X, Fang C (2011) Histogram of the oriented gradient for face recognition. Tsinghua Sci Technol 16:216–224
17. Zhang B, Gao Y, Zhao S, Liu J (2009) Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. IEEE Trans Image Process 19:533–544
18. Vapnik V (2013) The nature of statistical learning theory. Springer Science and Business Media
19. Joachims T (1999) Svm-light: support vector machine. University of Dortmund, p 19. http://svmlight.joachims.org/