Classification of Facial Images to Assist in the Diagnosis of Autism Spectrum Disorder

Gabriel C. Michelassi

University of Sao Paulo

Henrique S. Bortoletti

University of Sao Paulo

Tuany D. Pinheiro

University of Sao Paulo

Thiago Nobayashi

University of Sao Paulo

Fabio R. D. Barros

University of Sao Paulo

Rafael L. Testa

University of Sao Paulo

Andréia F. Silva

University of Sao Paulo

Mirian C. Revers

University of Sao Paulo

Joana Portolese

University of Sao Paulo

Helio Pedrini

State University of Campinas

Helena Brentani

University of Sao Paulo

Fatima L. S. Nunes

University of Sao Paulo

Ariane Machado-Lima (✉ ariane.machado@usp.br)

University of Sao Paulo

Research Article

Keywords: Autism Spectrum Disorder (ASD), typical development (TD), SVM classifier

Posted Date: April 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-448184/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Classification of facial images to assist in the diagnosis of Autism Spectrum Disorder

Gabriel C. Michelassi¹, Henrique S. Bortoletti¹, Tuany D. Pinheiro¹, Thiago Nobayashi¹, Fabio R. D. de Barros¹, Rafael L. Testa¹, Andréia F. Silva², Mirian C. Revers², Joana Portolese², Helio Pedrini³, Helena Brentani², Fatima L. S. Nunes¹, and Ariane Machado-Lima^{1,*}

¹School of Arts, Sciences and Humanities, University of Sao Paulo, R. Arlindo Bétio, 1000 - Ermelino Matarazzo, Sao Paulo-SP, 03828-000, Brazil

²Department of Psychiatry, School of Medicine, University of Sao Paulo, R. Dr. Ovídio Pires de Campos, 785, Sao Paulo-SP, 05403-903, Brazil

³Institute of Computing, University of Campinas, Av. Albert Einstein, 1251 - Cidade Universitária, Campinas-SP, 13083-852, Brazil

*ariane.machado@usp.br

ABSTRACT

Autism Spectrum Disorder (ASD) is a developmental disorder that persistently impairs communication and social interaction, with a high degree of clinical heterogeneity. The fact that the diagnosis is based on clinical observation which requires therapy expertise, interviews with parents and application of questionnaires, it contributes to delaying the diagnosis until school age. The study of anthropometric measures in individuals with ASD and individuals in typical development (TD) showed there are some differences between the two groups. However, there is a lack of computational tools to assist in the acquisition and analysis process of these measures. This paper proposes a computer-aided medical decision support system that, given a child's facial image captured by a digital camera, can discriminate between the two groups, ASD and TD, thus helping in the diagnosis. We defined a protocol for image acquisition and preprocessing, tested and compared methods for dataset balancing, dimension reduction and classification. The best results were obtained by a SVM classifier with 86.2% accuracy. Once the proposed model is based on facial images, it has the potential to facilitate early diagnosis of ASD.

Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental, polygenic and multifactorial disorder with evidence of genetic and environmental factors contributing to its etiology¹⁻⁴.

Individuals with ASD exhibit persistent impairments on the social communication and on the ability to develop, maintain and understand relationships, as well as restricted and repetitive patterns of behavior, interests, and activities. These symptoms are present from early childhood and can impair the person's daily life^{5,6}. The diagnosis is clinical. The instruments currently available for this diagnosis are mostly questionnaires and observation routines of the child's behavior, thus the diagnosis is often delayed until school age⁷.

The high degree of clinical heterogeneity in ASD further complicates the diagnosis. The clinical presentation can range from individuals with many stereotypes and nonverbal fluency to individuals with excellent fluency and difficulty to deal with everyday social situations and rigid behaviors. The presence of comorbidities contributes to this heterogeneity. About 70% of individuals with ASD have at least one comorbidity, while 48% have two or more. The most frequent comorbidities are Intellectual Disability, Attention Deficit Hyperactivity Disorder, epilepsy and sleep disorders⁸⁹.

In addition to clinical heterogeneity there is great genetic heterogeneity. About 10% of ASD patients are diagnosed with a genetic syndrome, in which the most common are fragile X syndrome ($\approx$ 1-2% of ASD cases), tuberous sclerosis ($\approx$ 1%), Rett syndrome ($\approx$ 0.5%) and neurofibromatosis (type 1, 1%)⁸⁹¹⁰.

Early diagnosis, followed by appropriate intervention, could enable the advantage of brain plasticity, with a higher probability of modifying ASD symptom trajectories. Appropriate prevention treatments for children in the risk group can be carried out before the manifestation of all disorder symptoms¹¹⁻¹³. Therefore, advances in the strategies for ASD diagnosis that could be applied in very young children represent enhanced treatment success possibilities, which is an important research topic.

Several studies have suggested that brain developmental problems can be reflected in the facial morphology¹⁴⁻¹⁸. Studies

of facial anthropometric features in individuals with ASD and individuals with typical development (TD) have shown there are significant differences between these two groups, such as the distance between the pupils, ear format, strabismus, head circumference^{15,19–21}. These results motivate the use of machine learning approaches to classify ASD versus TD individuals, based on anthropometric facial features, to facilitate diagnosing ASD.

An ASD computer-aided diagnosis system based on facial anthropometry should receive as input a facial image, extract the anthropometric features and perform the classification. The anthropometric features considered in this work are distances between pairs of points located on the face, known as landmarks. One example is the distance between the left and right corners of the mouth.

Studies using facial anthropometric pattern recognition to distinguish some syndromes indicate that the approach is both promising and amenable to improvement. For syndromes characterized by facial morphology clearly distinguishable from TD individuals, such as Down syndrome, accurate classifiers have been proposed^{22,23}. In²², the authors used geometric features based on facial anatomical landmarks, local texture features based on the Contourlet transform and local binary pattern to train a support vector machine (SVM) classifier, achieving 97.92% accuracy. In²³ the authors obtained 98.39% accuracy using facial fiducial point detection. A geometric descriptor was used to extract and represent a set of facial features and different classifiers were employed to generate the final results. For syndromes where the facial dysmorphism pattern is less evident, however, the accuracy may be lower. For example, in Cushing's syndrome, the achieved accuracy was of 62.8%²⁴.

A multiclass approach was applied to classify input images obtained using common digital cameras, among ten different syndromes: DiGeorge's Syndrome, Cri-du-chat, Cornelia de Lange, Fragile X, Mucopolysaccharidosis type 3, Noonan, Prader-Willi, Smith-Lemli-Opitz, Sotos and Williams-Beuren¹⁶. Two techniques were used to extract the landmarks from the images: by automatic and manual positioning of the points. Several classifiers were tested, the best result was 75.7% accuracy, achieved by Linear Discriminant Analysis (LDA) and manual positioning.

Most of the studies that have analyzed the anthropometric measurements of ASD patients used an electronic imaging device of stereophotogrammetry, which reconstructs a three-dimensional image from estimated coordinates using several two-dimensional (2D) images taken from different angles^{15,25,26}, without proposing or testing a classifier. The results presented in¹⁵ indicated 39 facial distances significantly different between ASD and TD boys. In²⁷, the authors performed cluster analysis to study ASD subgroups with similar morphological features and correlated behavioral and clinical traits. In²⁸ the authors performed an analysis regarding sexually dysmorphic features in adults with ASD. They showed that individuals with a more severe level of ASD have fewer discriminant features between the sexes. However, in order to disseminate the use of this system for ASD diagnosis, the cost of the entire system should be as low as possible.

Considering the perspectives and gaps in the literature, the objective of this paper is to propose and validate a low-cost computational method to classify individuals as ASD or TD, by processing 2D facial images. In addition to the proposed method, this paper offers three additional contributions: (1) a protocol defined to acquire and process the facial images; (2) a pipeline of image processing techniques to extract landmarks and compute distances between them and (3) the proposed ASD x TD classifier resulting from the evaluation of different machine learning algorithms.

Results

We created and tested several ASD versus TD classifiers varying the algorithm used for landmark detection, dataset preprocessing, dimension reduction and classifier induction. For this study we collected photographs from 41 ASD and 69 TD children and adolescents ranging from 5 to 18 years of age.

Table 1 shows, for each classifier induction algorithm, the estimated performance of the most accurate classifier from those obtained using that specific classifier induction algorithm and varying the algorithm for the other phases (landmark detection algorithm, dataset preprocessing strategies and dimension reduction algorithm). The performance values correspond to the average values of the 10-fold cross-validation. Figure 1 shows these classifiers' accuracies with the respective confidence interval.

The highest accuracy (86.2%) and highest precision (90.8%) were achieved by a SVM classifier using PCA for dimension reduction, Pearson's correlation filter with a 0.98 threshold applied in non-normalized features, SMOTE Tomek algorithm for class balancing and HOG algorithm for landmark detection.

The highest recall (81.2%) and highest f1-score (79.93%) were achieved by a Neural Network (NN) classifier, which also achieved the third best accuracy (80.9%). This classifier was trained using features selected by the PCA algorithm previously applying the Pearson's correlation filter with a 0.98, SMOTE SVM algorithm for class balance and CNN algorithm for landmark detection.

Table 1. Performance of the most accurate classifier for each classifier induction algorithm. The highest value for each performance metric is in bold. Cells with “-” value are for “not applied”.

Classifier	Test Settings				
	RF	SVM	NN	KNN	NB
Dimension Reduction	-	PCA	PCA	PCA	ReliefF
Correlation Filter	0.98	0.98	0.98	0.99	-
min_max Normalization	False	False	False	False	False
Class balance	SVM	Tomek	SVM	Borderline	KMeans
Landmark Detection	CNN	HOG	CNN	CLNF	HOG
Estimated Performance					
Accuracy	78.9%	86.2%	80.9%	83.1%	80.8%
Precision	73.1%	90.8%	81.0%	80.3%	81.3%
Recall	81.0%	73.5%	81.2%	75.5%	66.5%
F1-Score	75.2%	79.2%	79.9%	76.4%	72.0%

Figure 1. Classifier accuracies with the respective 95% confidence intervals

Discussion

This paper presented a low-cost method to assist in the diagnosis of ASD based on anthropometric facial features. Here, we propose the use of machine learning to train an ASD versus TD classifier based on 2D facial images captured by a digital photographic camera. Although we intend to validate the method with a greater number of individuals, the accuracy obtained indicates that the method has the potential to be applied in daily medical routines in the future.

In addition to the results exhibited, we presented three relevant contributions to improve research related to the ASD diagnosis, described in the Methods section. Firstly, we defined an image acquisition protocol that can be reused in future

studies. Secondly, we developed a pipeline of image processing to align, locate landmark, and extract features based on the distance of these points. Lastly, we composed a database of facial images from individuals with ASD and TD. The image collection was conducted during a period of six months, totalizing 41 ASD and 69 TD cases. This step was critical as we did not find public image databases with facial images related to this disorder.

Several classifiers were trained and tested varying the dataset preprocessing strategies, the dimension reduction and the classifier learning algorithms. The best result was achieved by the SVM classifier with 86.2% accuracy. Considering that the facial differences from individuals with ASD and TD appear to be subtle, this is a promising result compared to the 62.8% achieved in²⁴ for Cushing's syndrome. It indicates that it is possible to use anthropometric measures to assist in ASD diagnosis, even when a small sample size is available (41 ASD and 69 TD instances in this study) using traditional machine learning algorithms.

We intend, in the near future, to use this classifier in combination with other classifiers to analyze other aspects of the disorder, in order to create a multimodal classifier to achieve higher accuracy. We also intend to continue collecting images of cases and controls, especially from very young children. Our plan is to repeat the experiments with those images to verify if similar results could be obtained for young children. The intended goal is to assist in the advance of early ASD diagnosis, given that the earlier the treatment starts, the greater the success rate in diagnosing ASD.

Methods

Figure 2 shows the pipeline steps proposed in this work, which was entirely implemented in Python 3. The images were acquired following the protocol described in Section “Image acquisition protocol”. The image processing and feature extraction are described in Sections “Image preprocessing and band detection” and “Feature extraction”, respectively. Several classifiers were trained and tested using different strategies for dataset preprocessing, dimension reduction and induction algorithms, as described in Section “Classifier development”.

This project was approved by the Ethics Board of the School of Arts, Sciences and Humanities of University of Sao Paulo, protocol number 1.669.832 at August 8th, 2016. All experiments were performed in accordance with the Ethics Board requirements. The informed consent was obtained from all participants and/or their legal guardians.

Figure 2. Pipeline of the proposed method.

Image acquisition protocol

All images were captured with a Nikon Coolpix L120 compact camera using the protocol shown in Figure 3. In each acquisition, the camera was attached to a tripod fixed at a distance of 50 centimeters from the chair where the volunteer was accommodated. The chair was placed against the wall, and the volunteer was instructed to remain with his back fully resting on the chair. A white wall was chosen as the photo background, where a red band 50 centimeters by 5 centimeters was placed in order to be posteriorly used as a reference for normalizing the images and estimating distortion.

The volunteers were instructed to remain with a facial expression as neutral as possible and five front facial photographs were taken. A video was recorded when neutral expression facial image was not possible or if the face was inclined. Thus, a subsequent video analysis allowed selecting the best frame.

Images were captured from 41 ASD and 69 TD children and adolescents ranging from 5 to 18 years of age. The diagnosis of the ASD participants was confirmed by the Clinics Hospital of the University of São Paulo and the photographs were taken after the patient's follow-up visit. The TD group consisted of volunteers invited by email. All underwent screening by an infant

Figure 3. Image acquisition protocol. The images used are from²⁹ and³⁰ under free license.

psychiatrist specialized in ASD to confirm the absence of ASD and/or clinical evidence of genetic syndromes. All participants were authorized by the parents or legal guardians through an informed consent signed at the beginning of each data collection session.

Image preprocessing and band detection

The three first modules of the proposed pipeline (Figure 1), image preprocessing, band detection and feature extraction, were used to process the facial images in order to extract the features to be used in the classification module.

We acquired more than one frontal image from each volunteer. The best image was selected based on two criteria: facial neutrality and proportion of width and height of the reference band in order to avoid image distortion.

The image preprocessing was performed to rotate and crop all the images. The rotation intends to correct the lateral face inclination and was performed according to the following steps: eye region detection using *Haar cascades*; pupil localization; tracing a line passing over the pupil points; calculation of the angle θ formed by this line and the horizontal axis; image rotation according to θ . Finally, the images were cropped in order to remove the black background resulting from the rotation operation.

The steps for reference band detection were: (i) application of a bilateral filter for image smoothing, (ii) edge detection using the Canny method³¹, (iii) application of dilation and erosion operations to correct discontinuities on the edges, (iv) detection of the band contour using the method *findContours* from OpenCV library³² and (v) calculation of the width and height band.

Feature extraction

The next module is responsible to extract features from these images. These features are normalized distances between each pair of facial landmarks.

The first step consists in detecting facial points known as landmarks. The landmark detection was performed using the following algorithms: *Convolutional Neural Network* (CNN) and *Histogram of Oriented Gradients* (HOG) implemented by Dlib³³, *Deep Neural Network* (DNN) and *Haar cascades* implemented by OpenCv³⁴, and *Constrained Local Neural Fields* (CLNF) implemented by OpenFace³⁵. All these algorithms provide good precision to detect 68 points in the facial region (Figure 4) and all were tested to evaluate their impact on the final classification performance metrics.

The Euclidean distance between each pair of the detected landmarks will correspond to a feature. However, in order to deal with variations in the different images, the landmark distances are normalized by the reference band width w and height h . More specifically, for each pair of landmarks $p(x_1, y_1)$ and $q(x_2, y_2)$, each one represented by a coordinate (x, y) , the normalized Euclidean distance is calculated according to Equation 1:

$$d_n(p, q) = \sqrt{[(x_1 - x_2)/w]^2 + [(y_1 - y_2)/h]^2} \quad (1)$$

The set of all normalized distances forms the feature vector of an image, which was later used in the classification process.

Figure 4. All 68 detected points in the facial region. The image used is from CBCL Face database³⁶.

Classifier development

The last module of the proposed pipeline (Figure 1) consists in training and evaluating several classifiers varying the procedures for dataset preprocessing, dimension reduction and classifier induction. Here we describe these procedures and how the parameter tuning and performance evaluation were performed.

Dataset preprocessing

Three preprocessing tasks were applied on the dataset: initial feature filtering, Min-Max normalization and class balancing.

In the initial feature filtering, when two or more features presented Pearson correlation greater than a threshold value only one was kept. We tested threshold values of 0.99 and 0.98.

We also tested using the original or the normalized values of the features. To calculate the normalized values we applied the min-max normalization, which identifies the maximum and minimum values of each feature X in the dataset and transforms every feature value X_i into $X'_i = (X_i - X_{min}) / (X_{max} - X_{min})$.

For class balancing, four approaches were tested: i) undersampling, ii) oversampling with four variants of SMOTE algorithm³⁷ (Default, SVM, Borderline and KMeans Smote³⁸), iii) combination of under and over sampling with SMOTE Tomek, and iv) no balancing.

Dimension reduction

Dimension reduction methods are extremely important for classification tasks because they eliminate redundant features and can increase classifier performance.

Filter-type feature selection methods tend to have a lower computational cost than wrapper-type methods. In this paper, the implementation of the scikit-learn libraries³⁹ and skfeature⁴⁰ libraries were used and six different methods were applied:

- Principal Component Analysis (PCA), which performs feature extraction by transforming the initial set into a new set containing only the major components;
- Minimum Redundancy Maximum Relevance (mRMR), which performs the selection of features based on two criteria: one is the minimum criterion of redundancy between pairs of features and the other is the maximum relevance through the measurement of information;
- Correlation-based Feature Selection (CFS), which performs the selection of features considering that a suitable subset contains features highly correlated with the target class, but little correlated with each other;
- Fast Correlation-based Filter (FCBF), which performs the selection of features considering the main idea of the CFS method, but calculating the correlation measure based on the entropy concept of information theory;
- ReliefF, which performs the selection of features by means of a relevance ranking of each feature;
- Robust Feature Selection (RFS), which performs the feature selection through sparse dictionary learning.

Classifier induction

Six classification algorithms were evaluated in this study: K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Gaussian Naïve Bayes (NB), Neural Networks (NN) and Random Forests (RF).

RF was applied throughout the entire set of features because the feature selection process is embedded in the training algorithm. KNN, SVM, NB and NN were applied in several of the subset of features selected by the algorithms described in the previous section.

Parameter tuning and performance estimation

Parameter tuning of the classifier induction algorithms was performed using a grid search approach to select the set of parameters that gives the highest accuracy value, estimated by 10-fold cross-validation. In each fold of the cross-validation, the synthetic instances produced by the class balancing algorithm were excluded from the test sample.

Table 2 shows the parameters and their values that were tuned for each classifier. In the column “Tested values”, “Nr. features” means the number of features after dimension reduction. Default parameters were used for Gaussian Naive Bayes classifier. The parameter *hidden_layer_sizes* is specified as an ordered pair (x, y) where x is the number of hidden layers and y is the number of neurons per layer.

Table 2. Parameters and values tested for each classifier.

Classifier	Parameters	Tested values
RF	n_estimators	500, 1000, 1500, 2000
	min_samples_leaf	1, 2, 5, 10, 15, 20
	max_features	[0.1 * Nr. features, 6 * Nr. features]
SVM	kernel	linear, poly, rbf, sigmoid
	C	0.01, 0.1, 1, 5, 10, 50, 100
	gamma	'scale', 'auto', 0.00001, 0.0001, 0.001, 0.01, 0.1, 0.5, 1, 5, 10
	degree	2, 3, 5
NN	hidden_layer_sizes	(1, 100), (1, Nr. features), (3, Nr. features), (1, 50), (3, 50)
	alpha	0.0001, 0.001, 0.01
	activation	'identity', 'logistic', 'tanh', 'relu'
	solver	'lbfgs', 'sgd'
KNN	n_neighbors	1, 2, 5, 10, 15, 20, 25, 30, 35, 40
	weights	uniform, distance
	algorithm	ball_tree, kd_tree, brute

For each option of the dataset preprocessing tasks (threshold of initial feature filtering, use of original or normalized feature values, class balancing strategy), dimension reduction and classifier induction algorithms, the parameter values found in the parameter tuning were used to estimate the performance of the resulting classifier via 10-fold cross-validation. This performance estimation was executed to ensure that no synthetic data produced in the class balancing step is used in the test folds. The calculated measures were accuracy, precision, recall and f1-score. In addition, confidence intervals were calculated with 95% confidence.

References

1. Lord, C., Cook, E. H., Leventhal, B. L. & Amaral, D. G. Autism spectrum disorders. *Neuron* **28**, 355–363 (2000).
2. Johnson, C. P. & Myers, S. M. Identification and evaluation of children with autism spectrum disorders. *Pediatrics* **120**, 1183–1215 (2007).
3. Muhle, R., Trentacoste, S. V. & Rapin, I. The genetics of autism. *Pediatrics* **113**, e472–e486 (2004).
4. Deth, R., Muratore, C., Benzecry, J., Power-Charnitsky, V.-A. & Waly, M. How environmental and genetic factors combine to cause autism: A redox/methylation hypothesis. *Neurotoxicology* **29**, 190–201 (2008).
5. Association, A. P. et al. *Diagnostic and statistical manual of mental disorders (DSM-5®)* (American Psychiatric Pub, 2013).
6. Levy, S. E., Mandell, D. S. & Schultz, R. T. Autism. **374**, 1627–1638 (2009).

7. Mandell, D. S., Novak, M. M. & Zubritsky, C. D. Factors associated with age of diagnosis among children with autism spectrum disorders. *Pediatrics* **116**, 1480–1486 (2005).
8. Lai, M.-C. et al. Prevalence of co-occurring mental health diagnoses in the autism population: a systematic review and meta-analysis. *The Lancet Psychiatry* **6**, DOI: [10.1016/S2215-0366\(19\)30289-5](https://doi.org/10.1016/S2215-0366(19)30289-5) (2019).
9. Mannion, A., Leader, G. & Healy, O. An investigation of comorbid psychological disorders, sleep problems, gastrointestinal symptoms and epilepsy in children and adolescents with autism spectrum disorder. *Res. Autism Spectr. Disord.* **7**, 35–42, DOI: [10.1016/j.rasd.2012.05.002](https://doi.org/10.1016/j.rasd.2012.05.002) (2013).
10. Wiśniowiecka-Kowalnik, B. & Nowakowska, B. Genetics and epigenetics of autism spectrum disorder—current evidence in the field. *J. Appl. Genet.* **60**, DOI: [10.1007/s13353-018-00480-w](https://doi.org/10.1007/s13353-018-00480-w) (2019).
11. Dawson, G. Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Dev. Psychopathol.* **20**, 775–803 (2008).
12. Dawson, G. et al. Randomized, controlled trial of an intervention for toddlers with autism: the early start denver model. *Pediatrics* **125**, e17–e23 (2010).
13. Corsello, C. M. Early intervention in autism. *Infants & Young Child.* **18**, 74–85 (2005).
14. DeMyer, W., Zeman, W. & Palmer, C. G. The face predicts the brain: diagnostic significance of median facial anomalies for holoprosencephaly (arhinencephaly). *Pediatrics* **34**, 256–263 (1964).
15. Aldridge, K. et al. Facial phenotypes in subgroups of prepubertal boys with autism spectrum disorders are correlated with clinical phenotypes. *Mol. Autism* **2**, 15 (2011).
16. Boehringer, S. et al. Syndrome identification based on 2D analysis software. *Eur. J. Hum. Genet.* **14**, 1082–1089 (2006).
17. Hammond, P. et al. 3D analysis of facial morphology. *Am. J. Med. Genet. Part A* **126**, 339–348 (2004).
18. Zhao, Q. et al. Digital facial dysmorphology for genetic screening: Hierarchical constrained local model using ICA. *Med. Image Analysis* **18**, 699–710 (2014).
19. Miles, J., Hadden, L., Takahashi, T. & Hillman, R. Head circumference is an independent clinical finding associated with autism. *Am. J. Med. Genet.* **95**, 339–350 (2000).
20. Rodier, P. M., Bryson, S. E. & Welch, J. P. Minor malformations and physical measurements in autism: data from nova scotia. *Teratology* **55**, 319–325 (1997).
21. Ozgen, H. et al. Morphological features in children with autism spectrum disorders: a matched case–control study. *J. Autism Dev. Disord.* **41**, 23–31 (2011).
22. Zhao, Q. et al. Down syndrome detection from facial photographs using machine learning techniques. In *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 8670 (2013).
23. Cornejo, J., Pedrini, H., Machado-Lima, A. & Nunes, F. L. S. Down syndrome detection based on facial features using a geometric descriptor. *J. Med. Imaging* **4**, 1–6 (2017).
24. Popp, K. et al. Computer vision technology in the differential diagnosis of cushing’s syndrome. *Exp. Clin. Endocrinol. Diabetes* **127**, DOI: [10.1055/a-0887-4233](https://doi.org/10.1055/a-0887-4233) (2019).
25. Aldridge, K., Boyadjiev, S. A., Capone, G. T., DeLeon, V. B. & Richtsmeier, J. T. Precision and error of three-dimensional phenotypic measures acquired from 3dmd photogrammetric images. *Am. J. Med. Genet. Part A* **138**, 247–253 (2005).
26. Hammond, P. & Suttie, M. Large-scale objective phenotyping of 3D facial morphology. *Hum. Mutat.* **33**, 817–825 (2012).
27. Obafemi-Ajayi, T. et al. Facial structure analysis separates autism spectrum disorders into meaningful clinical subgroups. *J. Autism Dev. Disord.* **45**, 1302–1317 (2015).
28. Gilani, S. Z. et al. Sexually dimorphic facial features vary according to level of autistic-like traits in the general population. *J. Neurodev. Disord.* **7**, 14 (2015).
29. Jaimy. Stick figure line art. <https://bit.ly/3sCq9Ha> (2020). Online; Accessed 20 February 2021.
30. Nisthu. Camera silhouette. <https://bit.ly/3rB8s9E> (2020). Online; Accessed 20 February 2021.
31. Canny, J. A computational approach to edge detection. *IEEE Transactions on pattern analysis machine intelligence* 679–698 (1986).
32. Suzuki, S. Topological structural analysis of digitized binary images by border following. *Comput. Vision, Graph. Image Process.* **30**, 32–46 (1985).

33. King, D. E. Dlib-ml: A machine learning toolkit. *J. Mach. Learn. Res.* **10**, 1755–1758 (2009).
34. Bradski, G. Opencv library. *Dr. Dobb's J. Softw. Tools* (2000).
35. Baltrusaitis, T., Robinson, P. & Morency, L.-P. Constrained local neural fields for robust facial landmark detection in the wild. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 354–361 (2013).
36. Sung, K.-K. *Learning and Example Selection for Object and Pattern Recognition*. Ph.D. thesis, MIT, Artificial Intelligence Laboratory and Center for Biological and Computational Learning, Cambridge, MA (1996).
37. N. V. Chawla, L. O. H. W. P. K., K. W. Bowyer. Smote: Synthetic minority over-sampling technique. *J. Of Artif. Intell. Res.* **16**, 321–357 (2011).
38. Lemaître, G., Nogueira, F. & Aridas, C. K. Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *J. Mach. Learn. Res.* **18**, 1–5 (2017).
39. Pedregosa, F. et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
40. Li, J. et al. Feature selection: A data perspective. *arXiv:1601.07996* (2016).

Acknowledgements

We thank the patients and their families that allowed the execution of this research, as well as the team of the Autism Spectrum Program of the Clinics Hospital (PROTEA-HC). This study was funded by Brazilian National Council of Scientific and Technological Development, (CNPq) (grants 309330/2018-1, 157535/2017-7 and 309030/2019-6) and Scientific and Technological Initiation Program at University of Sao Paulo (PIBIC/PIBIT-CNPq/USP 2020/2021), the São Paulo Research Foundation (FAPESP) – National Institute of Science and Technology – Medicine Assisted by Scientific Computing (INCT-MACC) – grant 2014/50889-7, NAP eScience-PRP-USP and Sao Paulo Research Foundation (FAPESP) grants 2005/556288, #2017/12646-3, #2020/01992-0, Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), Dean's Office for Research of the University of São Paulo (grant 18.5.245.86.7) and the National Health Support Program for People with Disabilities (PRONAS/PCD) grant 25000.002484/2017-17.

Author contributions statement

GCM, TDP, HB, FLSN and AML defined the conceptualization of the model. TDP and TN captured the photographs. AFS, MCR, JP and HB performed the psychiatric screening of the participants. TDP, HSB, HP and FLSN defined and implemented the image processing pipeline. GCM, TDP, FRDB, RLT, HP and AML defined and implemented the machine learning pipeline. GCM, HB, FLSN and AML wrote the manuscript. HB, FLSN and AML were responsible for funding acquisition. All authors reviewed the manuscript.

Additional information

The authors declare that they have no conflict of interest.

Figures

Figure 1

Classifier accuracies with the respective 95% confidence intervals

Figure 2

Pipeline of the proposed method

Figure 3

Image acquisition protocol. The images used are from29 and30 under free license.

Figure 4

All 68 detected points in the facial region. The image used is from CBCL Face database³⁶.