Original Paper

Deep Learning–Based Autism Spectrum Disorder Detection Using Emotion Features From Video Recordings: Model Development and Validation

Essam Sleiman[1]; Onur Cezmi Mutlu[2], BSc, MSc; Saimourya Surabhi[1], BSc, MSc; Arman Husic[1], BSc; Aaron Kline[1], BS; Peter Washington[3], PhD; Dennis Paul Wall[1,4], PhD

[1]Department of Pediatrics (Systems Medicine), Stanford University, Palo Alto, CA, United States

[2]Department of Electrical Engineering, Stanford University, Stanford, CA, United States

[3]Department of Information and Computer Sciences, University of Hawai‘i at Mānoa, Honolulu, HI, United States

[4]Department of Biomedical Data Science, Stanford, CA, United States

Corresponding Author:
Dennis Paul Wall, PhD
Department of Pediatrics (Systems Medicine)
Stanford University
3145 Porter Drive
Palo Alto, CA, 94304
United States
Phone: 1 650 666 7676
Email: dpwall@stanford.edu

Abstract

Background: Autism spectrum disorder (ASD) is a prevalent neurodevelopmental disorder encountered by 1 in 44 children in the United States. Patients with autism face difficulty effectively communicating with peers, articulating feelings and emotions, and controlling behaviors. Rapid and early diagnosis leads to improved treatment outcomes; however, current medical techniques often take years. Difficult, time-consuming diagnoses accompanied by the recent advances in computer vision have led to a surge of interest from researchers in developing models to streamline the autism diagnosis process.

Objective: We aimed to assess the viability of at-home autism diagnosis by leveraging video data collected from a mobile game app to train a computer vision model using insights extracted from the change in emotion-related expression features over time.

Methods: With our *GuessWhat* game-based mobile app, we collected a video data set of 74 children with ASD and neurotypical diagnoses actively playing in a natural home environment. To investigate and quantify the significance of facial emotion features for ASD detection, we developed a deep learning–based autism classifier with the following two components: a convolutional neural network (CNN) attached to a long short-term memory (LSTM). We pretrained our CNN backbone in the following two ways: (1) on ImageNet or (2) on a compilation of facial expression recognition data sets to analyze autism detection performance improvements using emotion features. The output of each CNN is fed into an LSTM model to diagnose ASD from video data.

Results: Our top-performing architecture employing a CNN backbone trained on the facial expression recognition data sets and a unidirectional LSTM achieved a top accuracy of 91.2%, precision of 91.2%, recall of 89.8%, and an F_1-score of 90.6%.

Conclusions: We discovered the change in an individual’s facial expression features over time as a relevant marker for autism detection. Extracting emotional features from each frame resulted in a 27.8% F_1-score improvement when compared to ImageNet weights. Our study demonstrates the capability of mobile apps to collect a natural, diverse data set for an improved autism diagnosis. These results demonstrate that deep learning and computer vision–based methods are instrumental for automated autism diagnosis from at-home recorded videos using unspecialized equipment.

(JMIR Biomed Eng 2022;7(2):e39982) doi: 10.2196/39982

KEYWORDS

autism; emotion; video; computer vision; machine learning; mobile app; digital data; health artificial intelligence; AI; detection; pattern recognition; vision; mobile diagnostics; game application; mobile health; deep learning; neural network; facial expression; facial recognition

Introduction

Background

According to the Centers for Disease Control and Prevention, autism spectrum disorder (ASD) is a lifelong developmental disability formulated in late infancy or early childhood [1,2]. Since the year 2000, ASD has become further prevalent among children, increasing from 1 in 150 children diagnosed to 1 in 44 [1]. Even though autism can be detected in a child as early as 18 months [3], the global average diagnosis age is between 38 and 120 months, with more recent estimates suggesting average age of 60 months [4]. Studies show early treatment of ASD leads to enhanced cognitive development and communication later in life [5]. However, the average time span between initial behavioral concerns and an ASD diagnosis is 2 to 4 years [6]. Private services are taken advantage of by financially able parents, moderately speeding up their child's diagnosis process [7]; however, underprivileged patients cannot enjoy such amenities [8-10]. The treatment enhancements of early detection, the extensive current diagnosis process, and the lack of proper resources, particularly for disadvantaged communities, have induced rapid interest by the research community in early ASD diagnosis [11-24].

One emphasis has been on identifying behavioral markers for autism detection. Studies show that continuous communication barriers and restricted, repetitive behaviors are some symptoms of ASD. Furthermore, 50% of patients with ASD have intellectual impairments [25]. Patients with autism often lack social skills, have unusual moods or emotional reactions, and struggle to show proper facial expressions such as happy, sad, angry, and surprised [26]. Improved emotion features detection has been demonstrated to improve digital diagnostics and therapies [11,27-32].

As a result of increasing medical treatment costs, digital technologies have become a cost-effective alternative tool for managing, preventing, and diagnosing medical complications with similar performance to the current standard of care [11,33]. Mobile diagnosis can be a viable method for autism detection in the future. Specifically, mobile apps can effectively crowdsource data and labels [34-37], which are fundamental for training artificial intelligence models to detect ASD.

Prior Work

Autism classification data sets are gathered in the form of multiple modalities, including image, audio, and video. Given the properties of the video modality, both spatial (per frame) and temporal (across frames) analysis are vital for ASD classification. Prior research has investigated applying various machine learning models to detect autism using behaviors such as facial expressions and engagement. Cook et al use information such as the change of one's body movement over time in video extracted from a CNN to distinguish the two groups [38]. In a similar setup, head movement rate has also been explored to detect autism in toddlers [39].

Deveau et al [40] use game play metrics from a gamified mobile therapeutic to classify autism with a 0.745 area under the receiver operator characteristic curve. Additional research has investigated using a CNN to extract indicative features from audio signals to classify ASD or neurotypical (NT) with a 79% accuracy [12]. Hauser et al [41] used an age-informed approach to achieve 82.9% accuracy detecting ASD from audio snippets.

Recently, using facial expression signals such as emotion have also proven extremely effective for autism detection [42]. Owada et al [43] suggested facial expressions as a promising marker for core ASD social symptoms. Subsequent studies demonstrated patients with ASD make fewer emotional expressions and facial movements [11,44-46] compared with NT patients. Li et al [28] indicated facial attributes are statistically significant for ASD classification achieving an accuracy of 76% using a CNN. Furthermore, recurrent neural networks were found to be powerful at detecting ASD through gesture analysis from sequential data [27], achieving an accuracy of about 80%. Emotion analysis over frames in a video also proved to be effective for classifying ASD by taking the mean and standard deviation of a concatenation of CNN-produced emotion vectors [28].

Our Contributions

Prior work has not taken advantage of modern deep learning approaches for analyzing the change in a child's emotional features over time for autism diagnosis. Furthermore, previous research also lacks naturalistic, at-home data for emotion-based autism diagnosis training. Typical experiments employ bulky camera and audio systems [47,48], potentially leading to biased participant responses. Accordingly, our lab developed *GuessWhat* [49,50], a charades-like mobile app game children play with their parents. We gathered a video data set of emotion-motivated gameplay of patients with ASD and NT participants while playing *GuessWhat*. Children and parents are recorded with an institutional review board–approved consent process where the parents and children review each video before being presented with the option to upload their gameplay video to our Health Insurance Portability and Accountability Act (HIPAA)–compliant database. Parents have full control to delete each gameplay video after its recording.

In contrast to prior work, we propose a modern deep learning-based classifier to detect children with ASD from at-home recorded videos of natural parent-child interactions with assistance from our *GuessWhat* app. We extracted emotion-rich vectors from the video frames with a CNN and passed these vector frames into a long short-term memory (LSTM) to process and learn sequential dependencies of a child's emotion change over time. We analyzed varying video preprocessing methods and compared 2 methods for pretraining our emotion-extractor CNN. We achieved top accuracy of 91.2% and F_1-score of 90.6% with cross-validation. To our knowledge, our contribution is unique compared with previous work.

Methods

We used two data sets in our experiments: (1) a compilation of Facial Expression Recognition (FER) data sets and (2) the *GuessWhat* data set.

Ethics Approval

The study protocol was approved by the Stanford University Institutional Review Board (eProtocol number: 39562) and conducted in accordance with the Declaration of Helsinki. All participants provided informed consent electronically before any research procedures took place.

FER Data Sets

Our FER data sets consist of a combination of subsets of 18 publicly available emotion data sets totaling 33,074 images detailed in Table 1. We intentionally selected a variety of data sets to ensure diversity in gender, race, and age. Given these data sets come from different sources, we resized all images to 224 x 224 pixels. We used RetinaFace [51] to crop each image's face ensuring our model only learns from facial features and not from irrelevant background information. These 18 data sets are publicly available and as a result have already been preprocessed to some extent by their authors. Therefore, we faced little to no challenges in further preprocessing this data set for our needs. Each image is labeled one of the seven following Ekman emotions: sad, anger, fear, disgust, happy, surprise, and neutral.

Table 1. Compilation of 18 diverse public facial emotion recognition data sets with number of images, mean age, ethnicity, and gender split.

Data set	Number of images	Age (mean)	Ethnicity	Female, n/N (%)	Sad	Surprise	Fear	Neutral	Anger	Happy
ADFES [52]	194	21.5	55% European and 45% Mediterranean	87/194 (45)	43	21	22	22	44	42
AffectNet [53]	2937	23.73	N/A[a]	N/A	74	104	27	483	101	2148
CAFE [54]	1055	5.3	17% African; 17% Asian; 50% White; and 15% Latino	611/1055 (58)	96	93	125	207	349	185
CK+ [55]	432	34	81% White and 13% African	298/432 (69)	28	83	25	123	104	69
Dartmouth [56]	5480	9.84	White	2740/5480 (50)	722	729	727	693	1515	1094
emotionet [57]	1139	23	White, Asian, African American, and Hispanic	603/1139 (53)	133	67	49	N/A	130	760
eu [58]	111	40	N/A	65/111 (59)	15	16	15	17	33	15
faces [59]	2052	N/A	White	N/A	342	N/A	342	342	684	342
JAFFE [60]	213	N/A	Asian	213/213 (100)	31	30	32	30	59	31
KDEF [61]	2934	25	Latino	1467/2934 (50)	419	419	420	418	838	420
mefed [62]	4024	22.5	N/A	2012/4024 (50)	575	575	575	574	1150	575
mug [63]	401	27	White	160/401 (40)	48	66	47	25	128	87
NIMH [64]	525	25.5	N/A	346/525 (66)	104	N/A	102	111	102	106
NimStim [65]	586	19.4	58% White; 23% African American; 14% Asian; and 5% Latino	205/586 (35)	83	45	80	84	168	126
Oulu-CASIA [66]	1120	40.5	N/A	291/1120 (26)	160	160	160	160	320	160
RaFD [67]	4221	21.2	White	2152/4221 (51)	603	603	603	603	1206	603
tif [68]	111	0.66	White	61/111 (55)	21	12	5	26	16	31
Tsinghua-FED [69]	761	23.82	Asian	380/761 (50)	109	109	108	109	217	109

[a]N/A: not applicable.

GuessWhat

Our lab developed the *GuessWhat* mobile app, a charades-like game-based therapy for children with neurodevelopmental disorders. Through this app, we collected our second data set, the *GuessWhat* data set, which contains recorded videos of children aged 2-8 years in an at-home, natural setting acting out one of a range of appropriate and therapeutically relevant prompts shown on a mobile phone held against or at a parent's head or held by hand (Figure 1). The children are expected to respond to various prompts such as emojis and facial expressions (eg, a sad face), or more complex prompts such as sports games, occupations, and animals with support and mediation from the parent. Following the game, the parent can then choose whether to share the 90-second recording with the *GuessWhat* research team [49,70-73].

Owing to their nature, videos taken at home from a mobile phone camera require further data cleaning. One of the prominent problems is the number of people in a given video. Occasionally, we observe a parent or another child appearing in the same clip as the subject and sometimes even completely missing the subject. To mitigate this issue, we chose to use a popular computer vision algorithm named RetinaFace [51] to detect and crop the faces in all frames and discard the ones including multiple or no faces. Through this we also aimed to clean the data from irrelevant background information in the frames. After this data cleaning process, we are left with a total of 74 videos of children with ASD and 52 videos of NT children, each 90 seconds long (Table 2).

The videos have a default 30 frames-per-second (FPS) and with the given 90 seconds duration, the number of frames to be processed by the predictive model reaches a large value. Considering the known problems that sequential models experience under long sequences we chose to subsample and split these videos to allow proper training and predictive behavior. We first reduce the effective FPS to 4, 8, or 16 by subsampling. Then, using the new subsampled video, we create smaller clips by grouping every 4, 8, or 16 consecutive frames. Each of these small clips become one data point in our training and test routine. Lastly, we split our data set into 5 folds after leaving 10% of the data out as the test set. We ensured none of the children present in the train set appears in the test set, and the test set had an even distribution of patients with ASD or NT . Each video has one of the following two labels: ASD or NT.

Figure 1. GuessWhat app. (A) Child acts out a prompt on a mobile phone while playing with the parent. (B) GuessWhat video frames are split into two groups of Train and Test. Images and illustrations used in the figure are from Vecteezy.com.

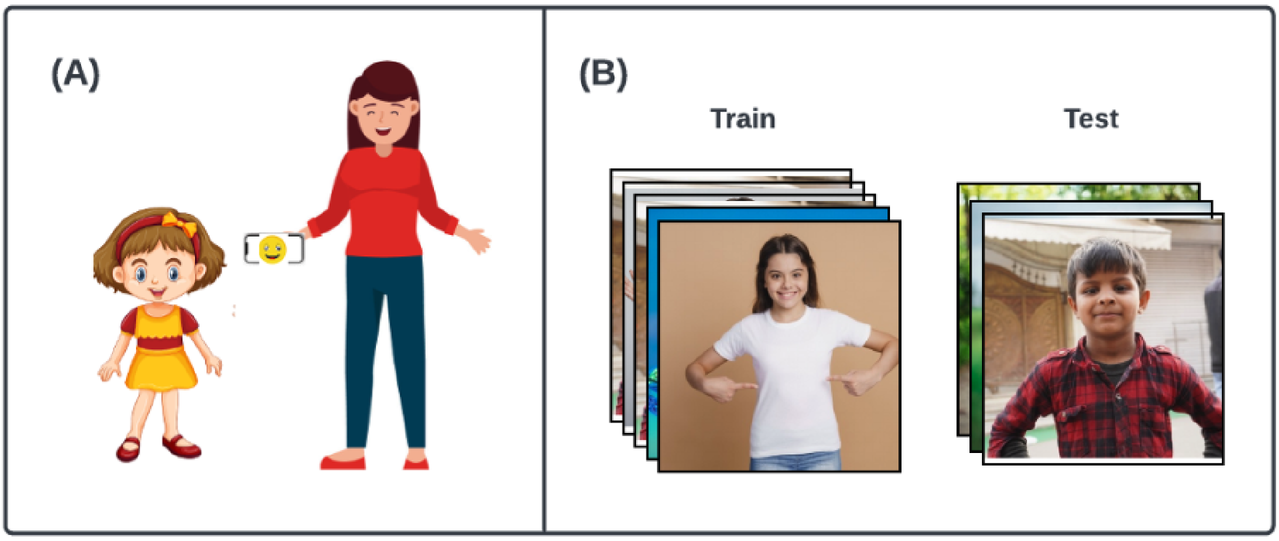


Table 2. Gender and age breakdown of GuessWhat data set across autism spectrum disorder (ASD) and neurotypical (NT) split.

Variables	ASD	NT
Age (years), mean	7.611	5.134
Sex, n (%)		
Male	14 (19.4)	18 (34.6)
Female	58 (80.6)	34 (65.4)

Training and Feature Extraction

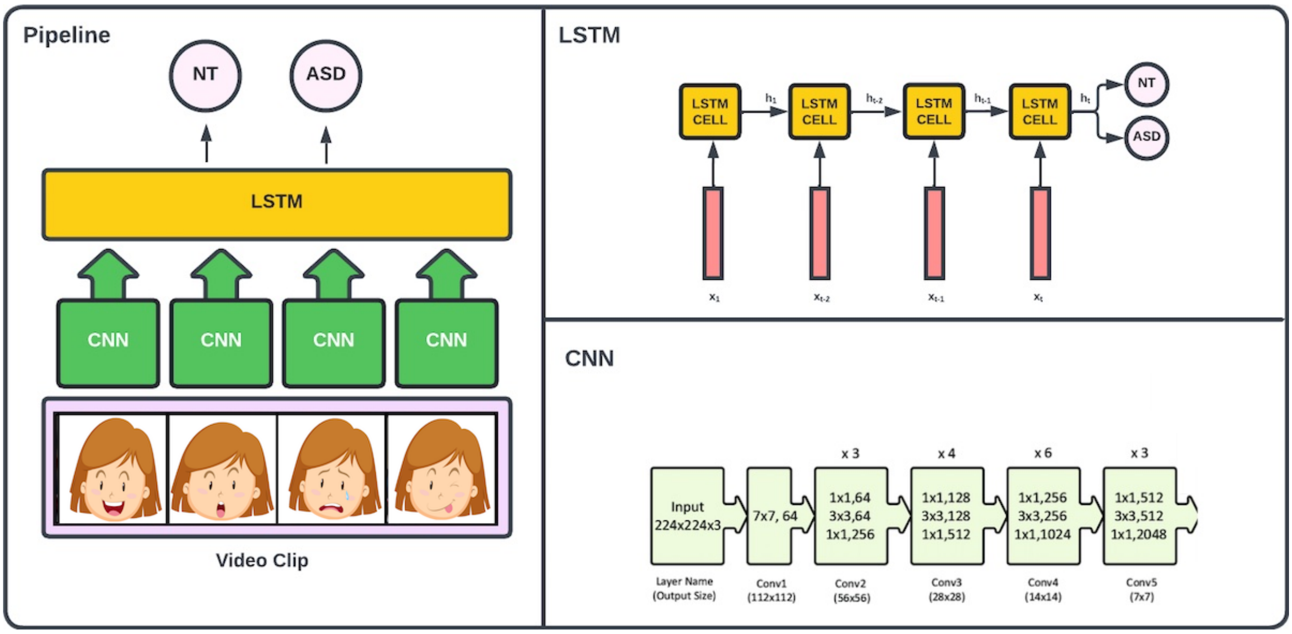
Our pipeline leverages the CNN + LSTM architecture, where first the CNN extracts facial emotion feature vectors from the cropped face in each frame. The LSTM then learns the temporal relationship between these feature vectors in a clip for ASD detection (Figure 2).

We used a ResNet-50 CNN model pretrained on ImageNet [74] as an emotion-feature extractor and further trained it on our FER data sets. Comparisons across different feature encoders and pretraining strategies can be found in the experiments section.

For the training of feature extractor, we used Adam optimizer [75] with a learning rate of 10^{-5} and a batch size of 256. For better generalization and robustness, we incorporated several image augmentation methods during our training such as color jitter, random rotation, and random horizontal and vertical flip.

Once the CNN backbone was trained, we removed the fully connected layers and froze the convolutional layers, which allowed us to extract a 512-dimensional vector of each face, retaining the underlying facial expression information in the case of the FER-trained CNN. We fed a sequence of frames extracted from each of our short GuessWhat clips into our CNN. The output embeddings from the frames of each clip passed into the CNN were then fed sequentially into an LSTM. The final hidden state of the LSTM was fed into a multilayer perceptron to predict ASD or NT. Our LSTM is optimized on the *GuessWhat* data set using Adam, with a learning rate of 10^{-2}, a batch size of 200, and an LSTM hidden state dimension of 64, with 2 LSTM layers and 1 fully connected layer.

Figure 2. Evaluation pipeline. CNN (ResNet-50) extracts an emotion feature vector from each frame in the video clip. LSTM then learns sequential dependencies across each emotion feature vector within a video clip. Images used in the figure are from Vecteezy.com. ASD: autism spectrum disorder; CNN: convolutional neural network; conv: convolution layer; h: hidden layer; LSTM: long short-term memory; NT: neurotypical.



Results

We first ran experiments to compare multiple CNN architectures on the facial emotion recognition task across our FER data sets. For each model, we trained on 16 out of the 18 FER data sets, validated on *Affectnet*, and tested on *Cafe* with an 80/10/10 data set split. We found high levels of confusion between anger and disgust, and due to their semantic similarities, we combined the two labels into anger. We trained and evaluated on a ResNet-18, ResNet-50, and VGG-16. The results can be found in Table 3. We used the ResNet-50 as it performed the best out of all 3 and yielded sufficiently salient features for the next stage of training our LSTM. ResNet-18 and ResNet-50 have the same architecture, but with different number of layers specified by the number after the word ResNet.

Subsequently, we compared various sequential models with the FER pretrained ResNet-50 backbone for autism classification. We found that unidirectional LSTM performs the best out of all 3 shown in Table 4.

In our next experiments, we attach a unidirectional LSTM on the end of our ResNet-50 backbone, trained on the FER data sets. Short clips from the *GuessWhat* data set are fed into the CNN, then passed to the LSTM, which is trained to perform ASD classification. We conducted experiments varying the following clip hyperparameters: number of FPS extracted from each video and number of frames per clip used to train our model. Each clip is evaluated separately, and the performance reported is an average of every clip's performance. Our goal was to investigate the impact of the number of different emotions included in each data point. We also sought to quantify whether frequent minute or less frequent substantial changes in emotion is a better indicator of autism.

Through this experiment, our top model obtained an ASD versus NT classification top accuracy of 91.2%, precision of 91.2%, recall of 89.8%, and F_1-score of 90.6% in the experiment with video preprocessed at 4-FPS and 16-frame sequences, as shown in Table 5 and Figure 3. Our results first indicate that facial expressions are significant markers for autism identification. Furthermore, varying clip length and FPS input to our model has a vast influence on performance. This is logically sound, as these two variables together vary the number of consecutive changes in emotion the model sequentially processes. Our results demonstrate that our model performs better on average when trained on greater sequential emotion information (more seconds) per clip and with a greater number of emotions per clip.

Table 3. Comparison of facial emotion recognition performance across different convolutional neural network architectures.

	Precision	Recall	*F* score	Accuracy
ResNet-18	76	68	68	68
ResNet-50[a]	80	73	74	73
VGG-16	77	72	72	72

[a]Best-performing model.

Table 4. Comparison of autism spectrum disorder classification performance across different sequential processing architectures all preprocessed at 4 frames per second (FPS) and 16 frames per data point.

4 FPS, 16 frames	Precision	Recall	*F* score	Accuracy
Unidirectional LSTM[a,b]	91.2	89.8	90.6	91.2
Bidirectional LSTM	81.8	87	83.4	83.2
GRU[c]	84.4	82.8	82.2	82.8

[a]LSTM: long short-term memory.

[b]Best-performing model.

[c]GRU: gated recurrent unit.

Figure 3. (a) F1-scores of data preprocessing variations; (b) confusion matrix for autism spectrum disorder (ASD) versus neurotypical (NT) classification on 4 frames per second, 16 frames data set. FPS: frames per second.

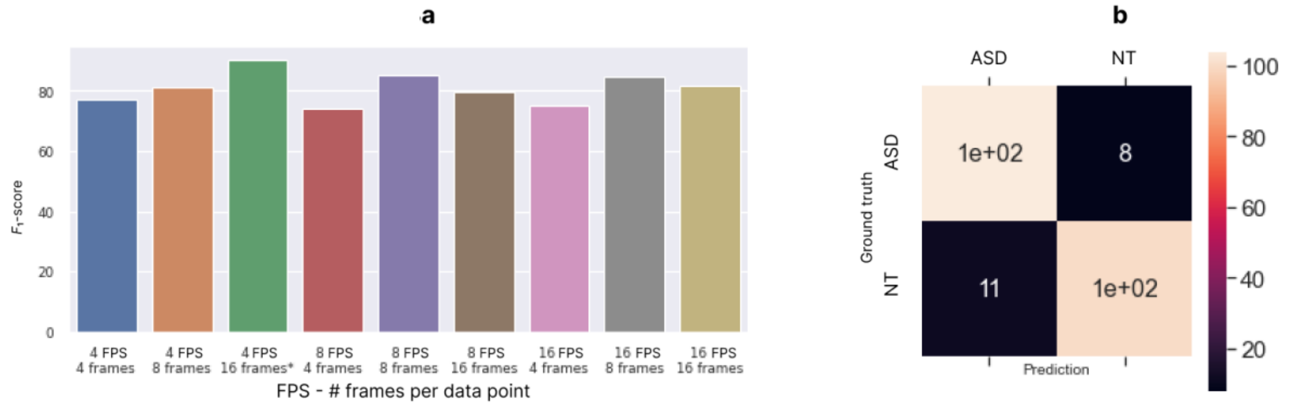


Table 5. ASD Classification using CNN[a] pretrained on the FER[b] data sets. Comparing experiment results of videos in GuessWhat data set extracted at different FPS[c] and different frame lengths per clip.

FPS and frames (seconds)	Precision	Recall	*F* score	Accuracy	Average number of emotions
04 FPS 04 frames (1)	77.6	78.6	77.2	78.6	2.18
04 FPS 08 frames (2)	82.8	80.6	81.2	82.8	3.79
04 FPS 16 frames (4)[d]	91.2	89.8	90.6	91.2	7.27
08 FPS 04 frames (0.5)	76.2	74.6	74.4	78.8	2.22
08 FPS 08 frames (1)	85.8	83.6	85.2	87.2	3.87
08 FPS 16 frames (2)	88.8	72.6	79.8	82.8	5.80
16 FPS 04 frames (0.25)	76.2	75.4	75.2	77.0	1.87
16 FPS 08 frames (0.5)	86.4	84.4	85.0	87.0	3.29
16 FPS 16 frames (1)	81.6	81.6	81.6	82.6	5.12

[a]CNN: convolutional neural network.

[b]FER: facial expression recognition.

[c]FPS: frames per second.

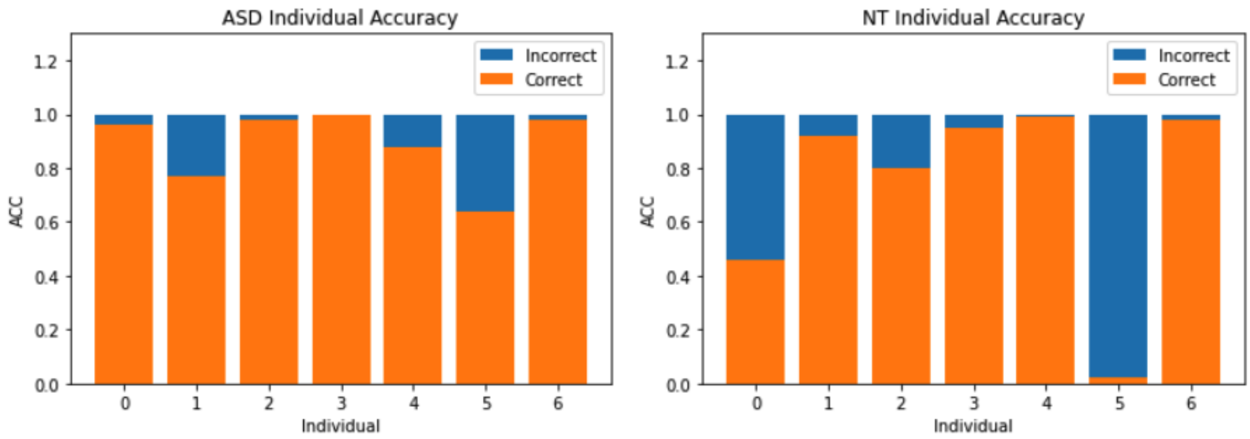
[d]Best-performing preprocessing method.

We also investigated the accuracies on the child level rather than on the clip level for 1 data preprocessing variation (16-FPS and 8-frame clip). Our results illustrate that our model is confident when classifying on the individual level; for the most part, the model is either correct or incorrect with high confidence (Figure 4).

Our results show facial expression extraction is vital for ASD classification; we had a 28.8% increase in F_1-score when comparing results from CNN pretrained on ImageNet and CNN pretrained on the FER data sets.

Lastly, to quantify the influence of emotion features on ASD detection, we ran an experiment with a backbone (ResNet-18), which is only pretrained on ImageNet (not FER data sets). We attached a unidirectional LSTM, as outlined in the previous section, to the model and trained and evaluated on ASD detection. This model received an accuracy of 45.8%, precision of 45.8, recall of 100, and F_1-score of 62.8 on 4 FPS per video and 16 frame-long data points. The poor results were likely caused by overfitting to unrelated features (ImageNet) and highlight the importance of pretraining on FER data sets for emotion extraction.

Figure 4. Model accuracy on the individual level for children with autism spectrum disorder (ASD) and neurotypical (NT) children.



Discussion

Principal Results

Our top-performing model obtained an F_1-score of 91.6%, with 4-FPS and 16-frame sequences when pretraining the CNN on the FER data sets and sequentially processing them through an LSTM. Our extensive experiments indicate that ResNet-50 performed best as compared to other image encoders, and the unidirectional LSTM performed best among the sequential models.

Limitations

The first limitation we faced in our experiments was our modest data set size. With 126 ninety-second long videos, our data set is intrinsically less diverse and as a result leads to some signs of overfitting in our experimentation. We also faced a slight class imbalance of patients with ASD and NT patients (74 vs 52), an imbalance in male versus female participants (14 vs 58) for ASD, and 18 versus 34 for NT.

There are inherent inconsistencies and shortcomings with naturally recorded videos in the *GuessWhat* data set such as jittery camera motion, forced emotional expressions, and so on. Furthermore, *GuessWhat* videos are self-labeled by the child's parent, and although they are likely to be true, they are unverified.

Future Work

Future directions for this work could be improving the facial feature extractor CNN both on the model side and by training on a larger, more diverse data set. This could be done with self-supervised learning methods on larger and cheaper unlabeled emotion data sets [76]. Another direction to improve patient privacy [77-79] could be on-device model personalization, where training takes place on a patient's device and not on our servers.

Our LSTM can be replaced and compared with more recent methods such as vision transformers [80]. We could also benefit from a more extensive, balanced *GuessWhat* data set to improve model generalizability. This would include increasing patient ethnic diversity, age diversity, and gender balance.

Multimodal learning, where a model is trained on the integration of multiple modalities such as video and audio, is another important next direction. Past works used audio for ASD classification [81] and took advantage of eye gaze to learn statistically different features between patients with ASD and NT patients [82]. Combining these and other methods with multimodal learning could even further boost performance and robustness.

Conclusions

In this study, we showed computer vision models trained on heterogeneous, at-home video recordings to be a promising method for autism diagnosis. Our results also indicate children's facial expressions and the rate at which their facial expressions change are salient features for machine learning–based diagnosis. More specifically, on average, the greater number of seconds of video per clip and more emotions per clip both led to improved performance. Our model detects autism in children with high success on inconsistent video recording quality, suggesting its diagnostic viability in a real-world setting. Our top-performing model received an F_1-score of 91.6% and consists of a ResNet-50 model as the facial feature extractor, a unidirectional LSTM to process the video sequences, and a video data set preprocessed at 4 FPS and 16 frames per data point. To our knowledge, we are the first to propose a deep learning–based architecture to detect autism by extracting facial expression features from videos collected in home settings and processing these features sequentially. We also conducted a robust comparison of various deep learning architectures on the *GuessWhat* data set to detect autism.

Acknowledgments

This paper was supported in part by funds to DPW from the National Institutes of Health (1R01EB025025-01, 1R21HD091500-01, 1R01LM013083, and 1R01LM013364), the National Science Foundation (Award 2014232), The Hartwell Foundation, Bill and Melinda Gates Foundation, Coulter Foundation, Lucile Packard Foundation, the Weston Havens Foundation, and program grants from Stanford's Human Centered Artificial Intelligence Program, Stanford's Precision Health and Integrated Diagnostics Center (PHIND), Stanford's Beckman Center, Stanford's Bio-X Center, Predictives and Diagnostics Accelerator (SPADA) Spectrum, Stanford's Spark Program in Translational Research, Stanford mediaX, and Stanford's Wu Tsai Neurosciences Institute's Neuroscience: Translate Program. We also acknowledge generous support from David Orr, Imma Calvo, Bobby Dekesyer, and Peter Sullivan. PW would like to acknowledge support from Mr Schroeder and the Stanford Interdisciplinary Graduate Fellowship (SIGF) as the Schroeder Family Goldman Sachs Graduate Fellow.

Conflicts of Interest

DPW is the founder of Cognoa.com. This company is developing digital health solutions for pediatric health care. ES, OCM, SS, AH, AK, and PW declare no conflicts of interest.

References

1. Maenner MJ, Shaw KA, Bakian AV, Bilder DA, Durkin MS, Esler A, et al. Prevalence and characteristics of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2018. MMWR Surveill Summ 2021 Dec 03;70(11):1-16 [FREE Full text] [doi: 10.15585/mmwr.ss7011a1] [Medline: 34855725]
2. Elder J, Kreider C, Brasher S, Ansell M. Clinical impact of early diagnosis of autism on the prognosis and parent-child relationships. PRBM 2017 Aug;Volume 10:283-292. [doi: 10.2147/prbm.s117499]
3. Hyman S, Levy SE, Myers SM, Council on Children with Disabilities, Section on Developmental and Behavioral Pediatrics. Identification, Evaluation, and Management of Children With Autism Spectrum Disorder. Pediatrics 2020 Jan;145(1):3447. [doi: 10.1542/peds.2019-3447] [Medline: 31843864]
4. van 't Hof M, Tisseur C, van Berckelear-Onnes I, van Nieuwenhuyzen A, Daniels AM, Deen M, et al. Age at autism spectrum disorder diagnosis: A systematic review and meta-analysis from 2012 to 2019. Autism 2021 May;25(4):862-873 [FREE Full text] [doi: 10.1177/1362361320971107] [Medline: 33213190]
5. Clark MLE, Vinen Z, Barbaro J, Dissanayake C. School Age Outcomes of Children Diagnosed Early and Later with Autism Spectrum Disorder. J Autism Dev Disord 2018 Jan 14;48(1):92-102. [doi: 10.1007/s10803-017-3279-x] [Medline: 28905160]
6. Zablotsky B, Colpe LJ, Pringle BA, Kogan MD, Rice C, Blumberg SJ. Age of parental concern, diagnosis, and service initiation among children with autism spectrum disorder. Am J Intellect Dev Disabil 2017 Jan;122(1):49-61 [FREE Full text] [doi: 10.1352/1944-7558-122.1.49] [Medline: 28095057]
7. Shepherd CA, Waddell C. A Qualitative Study of Autism Policy in Canada: Seeking Consensus on Children's Services. J Autism Dev Disord 2015 Nov 24;45(11):3550-3564 [FREE Full text] [doi: 10.1007/s10803-015-2502-x] [Medline: 26105592]
8. Sturm A, Williams J, Kasari C. Who gains and who loses? Sociodemographic disparities in access to special education services among autistic students. Autism Res 2021 Aug;14(8):1621-1632. [doi: 10.1002/aur.2517] [Medline: 33904253]
9. Bilaver LA, Sobotka SA, Mandell DS. Understanding racial and ethnic disparities in autism-related service use among Medicaid-enrolled children. J Autism Dev Disord 2021 Sep;51(9):3341-3355 [FREE Full text] [doi: 10.1007/s10803-020-04797-6] [Medline: 33219917]
10. Drahota A, Sadler R, Hippensteel C, Ingersoll B, Bishop L. Service deserts and service oases: Utilizing geographic information systems to evaluate service availability for individuals with autism spectrum disorder. Autism 2020 Nov;24(8):2008-2020 [FREE Full text] [doi: 10.1177/1362361320931265] [Medline: 32564619]
11. Voss C, Schwartz J, Daniels J, Kline A, Haber N, Washington P, et al. Effect of Wearable Digital Intervention for Improving Socialization in Children With Autism Spectrum Disorder: A Randomized Clinical Trial. JAMA Pediatr 2019 May 01;173(5):446-454 [FREE Full text] [doi: 10.1001/jamapediatrics.2019.0285] [Medline: 30907929]
12. Chi NA, Washington P, Kline A, Husic A, Hou C, He C, et al. Classifying Autism From Crowdsourced Semistructured Speech Recordings: Machine Learning Model Comparison Study. JMIR Pediatr Parent 2022 Apr 14;5(2):e35406 [FREE Full text] [doi: 10.2196/35406] [Medline: 35436234]
13. Vabalas A, Gowen E, Poliakoff E, Casson AJ. Applying Machine Learning to Kinematic and Eye Movement Features of a Movement Imitation Task to Predict Autism Diagnosis. Sci Rep 2020 May 20;10(1):8346 [FREE Full text] [doi: 10.1038/s41598-020-65384-4] [Medline: 32433501]
14. Crippa A, Salvatore C, Perego P, Forti S, Nobile M, Molteni M, et al. Use of Machine Learning to Identify Children with Autism and Their Motor Abnormalities. J Autism Dev Disord 2015 Jul 5;45(7):2146-2156. [doi: 10.1007/s10803-015-2379-8] [Medline: 25652603]
15. Hou C, Kalantarian H, Washington P, Dunlap K, Wall D. Leveraging video data from a digital smartphone autism therapy to train an emotion detection classifier. medRxiv. Preprint posted online August 1, 2021 [FREE Full text]

16. Banerjee A, Washington P, Mutlu C, Kline A, Wall D. Training and profiling a pediatric emotion recognition classifier on mobile devices. Arxiv. URL: https://arxiv.org/abs/2108.11754 [accessed 2022-09-05]
17. Tariq Q, Fleming SL, Schwartz JN, Dunlap K, Corbin C, Washington P, et al. Detecting Developmental Delay and Autism Through Machine Learning Models Using Home Videos of Bangladeshi Children: Development and Validation Study. J Med Internet Res 2019 Apr 24;21(4):e13822 [FREE Full text] [doi: 10.2196/13822] [Medline: 31017583]
18. Washington P, Leblanc E, Dunlap K, Penev Y, Varma M, Jung JY, et al. Selection of trustworthy crowd workers for telemedical diagnosis of pediatric autism spectrum disorder. 2021 Presented at: Pacific Symposium on Biocomputing; 2020; Hawaii p. 2020-2025. [doi: 10.1142/9789811232701_0002]
19. Washington P, Chrisman B, Leblanc E, Dunlap K, Kline A, Mutlu C, et al. Crowd annotations can approximate clinical autism impressions from short home videos with privacy protections. Intell Based Med 2022;6 [FREE Full text] [doi: 10.1016/j.ibmed.2022.100056] [Medline: 35634270]
20. Leblanc E, Washington P, Varma M, Dunlap K, Penev Y, Kline A, et al. Feature replacement methods enable reliable home video analysis for machine learning detection of autism. Sci Rep 2020 Dec 04;10(1):21245 [FREE Full text] [doi: 10.1038/s41598-020-76874-w] [Medline: 33277527]
21. Washington P, Kalantarian H, Kent J, Husic A, Kline A, Leblanc E, et al. Training Affective Computer Vision Models by Crowdsourcing Soft-Target Labels. Cognit Comput 2021 Sep 27;13(5):1363-1373 [FREE Full text] [doi: 10.1007/s12559-021-09936-4] [Medline: 35669554]
22. Washington P, Kline A, Mutlu O, Leblanc E, Hou C, Stockham N, et al. Activity recognition with moving cameras and few training examples: Applications for detection of autism-related headbanging. 2021 Presented at: CHI Conference on Human Factors in Computing Systems; 2021; Yokohama. [doi: 10.1145/3411763.3451701]
23. Lakkapragada A, Kline A, Mutlu O, Paskov K, Chrisman B, Stockham N, et al. The Classification of Abnormal Hand Movement to Aid in Autism Detection: Machine Learning Study. JMIR Biomed Eng 2022 Jun 6;7(1):e33771 [FREE Full text] [doi: 10.2196/33771]
24. Washington P, Kalantarian H, Kent J, Husic A, Kline A, Leblanc E, et al. Improved Digital Therapy for Developmental Pediatrics Using Domain-Specific Artificial Intelligence: Machine Learning Study. JMIR Pediatr Parent 2022 Apr 08;5(2):e26760 [FREE Full text] [doi: 10.2196/26760] [Medline: 35394438]
25. Hyman SL, Levy SE, Myers SM, Council on Children With Disabilities, Section on Developmental and Behavioral Pediatrics. Identification, evaluation, and management of children with autism spectrum disorder. Pediatrics 2020 Jan;145(1):2019-3447. [doi: 10.1542/peds.2019-3447] [Medline: 31843864]
26. Signs and symptoms of autism spectrum disorder. Centers for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/ncbddd/autism/signs.html [accessed 2022-09-05]
27. Zunino A, Morerio P, Cavallo A, Ansuini C, Podda J, Battaglia F, et al. Video gesture analysis for autism spectrum disorder detection. 2018 Presented at: 24th International Conference on Pattern Recognition (ICPR); November 29, 2018; Beijing, China. [doi: 10.1109/ICPR.2018.8545095]
28. Li B, Mehta S, Aneja D, Foster C, Ventola P, Shic F, et al. A facial affect analysis system for autism spectrum disorder. 2019 Presented at: International Conference on Image Processing (ICIP); September 22-25, 2019; Taipei, Taiwan p. 4549-4553. [doi: 10.1109/ICIP.2019.8803604]
29. Voss C, Washington P, Haber N, Kline A, Daniels J, Fazel A, et al. Superpower glass: delivering unobtrusive real-time social cues in wearable systems. 2016 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing; 2016; Heidelberg. [doi: 10.1145/2968219.2968310]
30. Kline A, Voss C, Washington P, Haber N, Schwartz H, Tariq Q, et al. Superpower Glass. GetMobile: Mobile Comp. and Comm 2019 Nov 14;23(2):35-38. [doi: 10.1145/3372300.3372308]
31. Washington P, Voss C, Haber N, Tanaka S, Daniels J, Feinstein C, et al. A wearable social interaction aid for children with autism. 2016 Presented at: 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems; 2016; San Jose. [doi: 10.1145/2851581.2892282]
32. Washington P, Voss C, Kline A, Haber N, Daniels J, Fazel A, et al. SuperpowerGlass. Proc. ACM Interact. Mob. Wearable Ubiquitous Technol 2017 Sep 11;1(3):1-22. [doi: 10.1145/3130977]
33. Szalay KZ, Csermely P. Perturbation centrality and turbine: a novel centrality measure obtained using a versatile network dynamics tool. PLoS One 2013;8(10):e78059 [FREE Full text] [doi: 10.1371/journal.pone.0078059] [Medline: 24205090]
34. Washington P, Park N, Srivastava P, Voss C, Kline A, Varma M, et al. Data-Driven Diagnostics and the Potential of Mobile Artificial Intelligence for Digital Therapeutic Phenotyping in Computational Psychiatry. Biol Psychiatry Cogn Neurosci Neuroimaging 2020 Aug;5(8):759-769 [FREE Full text] [doi: 10.1016/j.bpsc.2019.11.015] [Medline: 32085921]
35. Washington P, Leblanc E, Dunlap K, Penev Y, Kline A, Paskov K, et al. Precision Telemedicine through Crowdsourced Machine Learning: Testing Variability of Crowd Workers for Video-Based Autism Feature Recognition. J Pers Med 2020 Aug 13;10(3):86 [FREE Full text] [doi: 10.3390/jpm10030086] [Medline: 32823538]
36. Washington P, Paskov K, Kalantarian H, Stockham N, Voss C, Kline A, et al. Feature Selection and Dimension Reduction of Social Autism Data. 2020 Presented at: Pac Symp Biocomput; 2020; The Big Island of Hawaii p. 707-718. [doi: 10.1142/9789811215636_0062]

37. Washington P, Kalantarian H, Tariq Q, Schwartz J, Dunlap K, Chrisman B, et al. Validity of Online Screening for Autism: Crowdsourcing Study Comparing Paid and Unpaid Diagnostic Tasks. J Med Internet Res 2019 May 23;21(5):e13668. [doi: 10.2196/13668]
38. Cook A, Mandal B, Berry D, Johnson M. Towards Automatic Screening of Typical and Atypical Behaviors in Children With Autism. 2019 Presented at: IEEE International Conference on Data Science and Advanced Analytics (DSAA); 2019; Washington. [doi: 10.1109/DSAA.2019.00065]
39. Dawson G, Campbell K, Hashemi J, Lippmann SJ, Smith V, Carpenter K, et al. Atypical postural control can be detected via computer vision analysis in toddlers with autism spectrum disorder. Sci Rep 2018 Nov 19;8(1):17008 [FREE Full text] [doi: 10.1038/s41598-018-35215-8] [Medline: 30451886]
40. Deveau N, Washington P, Leblanc E, Husic A, Dunlap K, Penev Y, et al. Machine learning models using mobile game play accurately classify children with autism. Intell Based Med 2022;6:100057 [FREE Full text] [doi: 10.1016/j.ibmed.2022.100057] [Medline: 36035501]
41. Hauser M, Sariyanidi E, Tunc B, Zampella C, Brodkin E, Schultz R. Using natural conversations to classify autism with limited data: Age matters. 2019 Presented at: Proceedings of the Sixth Workshop on Computational Linguistics and Clinical Psychology; June 6, 2019; Minneapolis, Minnesota, USA. [doi: 10.18653/v1/w19-3006]
42. Rudovic O, Utsumi Y, Lee J, Hernandez J, Ferrer E, Schuller B, et al. CultureNet: A Deep Learning Approach for Engagement Intensity Estimation from Face Images of Children with Autism. 2018 Presented at: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); January 06, 2019; Madrid, Spain. [doi: 10.1109/IROS.2018.8594177]
43. Owada K, Kojima M, Yassin W, Kuroda M, Kawakubo Y, Kuwabara H, et al. Computer-analyzed facial expression as a surrogate marker for autism spectrum social core symptoms. PLoS One 2018 Jan 2;13(1):e0190442 [FREE Full text] [doi: 10.1371/journal.pone.0190442] [Medline: 29293598]
44. Guha T, Yang Z, Ramakrishna A, Grossman R, Darren H, Lee S, et al. On Quantifying Facial Expression-Related Atypicality of Children with Autism Spectrum Disorder. In: Proc IEEE Int Conf Acoust Speech Signal Process. 2015 Apr Presented at: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); April 19-24, 2015; South Brisbane, QLD, Australia p. 803-807 URL: https://europepmc.org/abstract/MED/26705397 [doi: 10.1109/ICASSP.2015.7178080]
45. Leo M, Carcagnì P, Distante C, Spagnolo P, Mazzeo P, Rosato A, et al. Computational Assessment of Facial Expression Production in ASD Children. Sensors (Basel) 2018 Nov 16;18(11):3993 [FREE Full text] [doi: 10.3390/s18113993] [Medline: 30453518]
46. Manfredonia J, Bangerter A, Manyakov NV, Ness S, Lewin D, Skalkin A, et al. Automatic Recognition of Posed Facial Expression of Emotion in Individuals with Autism Spectrum Disorder. J Autism Dev Disord 2019 Jan 8;49(1):279-293. [doi: 10.1007/s10803-018-3757-9] [Medline: 30298462]
47. Jiang M, Francis S, Srishyla D, Conelea C, Zhao Q, Jacob S. Classifying Individuals with ASD Through Facial Emotion Recognition and Eye-Tracking. In: Annu Int Conf IEEE Eng Med Biol Soc. 2019 Jul Presented at: 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2019; Berlin p. 6063-6068. [doi: 10.1109/EMBC.2019.8857005]
48. Harun N, Hamzah N, Zaini N, Sani M, Norhazman H, Yassin I. EEG Classification Analysis for Diagnosing Autism Spectrum Disorder based on Emotions. Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 2018:87-93.
49. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall DP. Guess What?: Towards Understanding Autism from Structured Video Using Facial Affect. J Healthc Inform Res 2019 Oct 2;3(1):43-66 [FREE Full text] [doi: 10.1007/s41666-018-0034-9] [Medline: 33313475]
50. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall DP. Guess What?: Towards Understanding Autism from Structured Video Using Facial Affect. J Healthc Inform Res 2019 Oct 2;3(1):43-66 [FREE Full text] [doi: 10.1007/s41666-018-0034-9] [Medline: 33313475]
51. Jiankang D, Jia G, Evangelos V, Irene K, Stefanos Z. RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild. 2020 Presented at: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); June 13-19, 2020; Seattle, WA, USA. [doi: 10.1109/CVPR42600.2020.00525]
52. van der Schalk J, Hawk ST, Fischer AH, Doosje B. Moving faces, looking places: validation of the Amsterdam Dynamic Facial Expression Set (ADFES). Emotion 2011 Aug;11(4):907-920. [doi: 10.1037/a0023853] [Medline: 21859206]
53. Mollahosseini A, Hasani B, Mahoor MH. AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild. IEEE Trans. Affective Comput 2019 Jan 1;10(1):18-31. [doi: 10.1109/taffc.2017.2740923]
54. LoBue V, Thrasher C. The Child Affective Facial Expression (CAFE) set: validity and reliability from untrained adults. Front Psychol 2014 Jan 06;5:1532 [FREE Full text] [doi: 10.3389/fpsyg.2014.01532] [Medline: 25610415]
55. Lucey P, Cohn J, Kanade T, Saragih J, Ambadar Z, Matthews I. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. 2018 Presented at: IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops; 2010; San Francisco. [doi: 10.1109/CVPRW.2010.5543262]

56. Dalrymple KA, Gomez J, Duchaine B. The Dartmouth Database of Children's Faces: acquisition and validation of a new face stimulus set. PLoS One 2013 Nov 14;8(11):e79131 [FREE Full text] [doi: 10.1371/journal.pone.0079131] [Medline: 24244434]
57. Benitez-Quiroz C, Srinivasan R, Martinez A. EmotioNet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild. 2016 Presented at: IEEE Conference on Computer Vision and Pattern Recognition; 2016; Las Vegas. [doi: 10.1109/CVPR.2016.600]
58. O'Reilly H, Pigat D, Fridenson S, Berggren S, Tal S, Golan O, et al. The EU-Emotion Stimulus Set: A validation study. Behav Res Methods 2016 Jun 30;48(2):567-576. [doi: 10.3758/s13428-015-0601-4] [Medline: 26424443]
59. Ebner NC, Riediger M, Lindenberger U. FACES--a database of facial expressions in young, middle-aged, and older women and men: development and validation. Behav Res Methods 2010 Feb;42(1):351-362. [doi: 10.3758/BRM.42.1.351] [Medline: 20160315]
60. Michael L, Miyuki K, Jiro G. The Japanese Female Facial Expression (JAFFE) Dataset Data set. Zenodo. 1998. URL: https://zenodo.org/record/3451524#.YyImFuzMJoN [accessed 2022-09-14]
61. Goeleven E, De Raedt R, Leyman L, Verschuere B. The Karolinska Directed Emotional Faces: A validation study. Cognition & Emotion 2008 Sep;22(6):1094-1118. [doi: 10.1080/02699930701626582]
62. Kamińska D, Aktas K, Rizhinashvili D, Kuklyanov D, Sham AH, Escalera S, et al. Two-Stage Recognition and beyond for Compound Facial Emotion Recognition. Electronics 2021 Nov 19;10(22):2847. [doi: 10.3390/electronics10222847]
63. Aifanti N, Papachristou C, Delopoulos A. The MUG facial expression database. 2010 Presented at: 11th International Workshop on Image Analysis for Multimedia Interactive Services WIAMIS 10; 2010; Desenzano del Garda.
64. Egger HL, Pine DS, Nelson E, Leibenluft E, Ernst M, Towbin KE, et al. The NIMH Child Emotional Faces Picture Set (NIMH-ChEFS): a new set of children's facial emotion stimuli. Int J Methods Psychiatr Res 2011 Sep 24;20(3):145-156 [FREE Full text] [doi: 10.1002/mpr.343] [Medline: 22547297]
65. Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare TA, et al. The NimStim set of facial expressions: judgments from untrained research participants. Psychiatry Res 2009 Aug 15;168(3):242-249 [FREE Full text] [doi: 10.1016/j.psychres.2008.05.006] [Medline: 19564050]
66. Zhao G, Huang X, Taini M, Li SZ, Pietikäinen M. Facial expression recognition from near-infrared videos. Image and Vision Computing 2011 Aug;29(9):607-619. [doi: 10.1016/j.imavis.2011.07.002]
67. Langner O, Dotsch R, Bijlstra G, Wigboldus DHJ, Hawk ST, van Knippenberg A. Presentation and validation of the Radboud Faces Database. Cognition & Emotion 2010 Dec;24(8):1377-1388. [doi: 10.1080/02699930903485076]
68. Maack JK, Bohne A, Nordahl D, Livsdatter L, Lindahl ÅAW, Øvervoll M, et al. The Tromso Infant Faces Database (TIF): Development, Validation and Application to Assess Parenting Experience on Clarity and Intensity Ratings. Front Psychol 2017 Mar 24;8:409 [FREE Full text] [doi: 10.3389/fpsyg.2017.00409] [Medline: 28392772]
69. Yang T, Yang Z, Xu G, Gao D, Zhang Z, Wang H, et al. Tsinghua facial expression database - A database of facial expressions in Chinese young and older women and men: Development and validation. PLoS One 2020 Apr 15;15(4):e0231304 [FREE Full text] [doi: 10.1371/journal.pone.0231304] [Medline: 32294105]
70. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall DP. Guess What?: Towards Understanding Autism from Structured Video Using Facial Affect. J Healthc Inform Res 2019 Jun;3(2):43-66 [FREE Full text] [doi: 10.1007/s41666-018-0034-9] [Medline: 33313475]
71. Kalantarian H, Washington P, Schwartz J, Daniels J, Haber N, Wall D. A gamified mobile system for crowdsourcing video for autism research. 2018 Presented at: IEEE International Conference on Healthcare Informatics (ICHI); 2018; New York City. [doi: 10.1109/ICHI.2018.00052]
72. Kalantarian H, Jedoui K, Washington P, Tariq Q, Dunlap K, Schwartz J, et al. Labeling images with facial emotion and the potential for pediatric healthcare. Artif Intell Med 2019 Jul;98:77-86 [FREE Full text] [doi: 10.1016/j.artmed.2019.06.004] [Medline: 31521254]
73. Kalantarian H, Jedoui K, Dunlap K, Schwartz J, Washington P, Husic A, et al. The Performance of Emotion Classifiers for Children With Parent-Reported Autism: Quantitative Feasibility Study. JMIR Ment Health 2020 Apr 01;7(4):e13174 [FREE Full text] [doi: 10.2196/13174] [Medline: 32234701]
74. Deng J, Dong W, Socher R, Li L, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. 2009 Presented at: IEEE Conference on Computer Vision and Pattern Recognition; 2009; Miami. [doi: 10.1109/CVPR.2009.5206848]
75. Kingma D, Ba J. Adam: A Method for Stochastic Optimization. 2015 Presented at: International Conference on Learning Representations (ICLR); 2015; San Diego.
76. Surabhi S, Shah B, Washington P, Mutlu O, Leblanc E, Mohite P, et al. TikTok for Good: Creating a Diverse Emotion Expression Database. 2022 Presented at: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops; 2022; New Orleans, US. [doi: 10.1109/CVPRW56347.2022.00279]
77. Kotz D, Gunter CA, Kumar S, Weiner JP. Privacy and Security in Mobile Health: A Research Agenda. Computer 2016 Jun;49(6):22-30. [doi: 10.1109/mc.2016.185]
78. Washington P, Yeung S, Percha B, Tatonetti N, Liphardt J, Wall D. Achieving trustworthy biomedical data solutions. Biocomputing 2021 2021:1-13. [doi: 10.1142/9789811232701_0001]

79. Washington P, Tariq Q, Leblanc E, Chrisman B, Dunlap K, Kline A, et al. Crowdsourced privacy-preserved feature tagging of short home videos for machine learning ASD detection. Sci Rep 2021 Apr 07;11(1):7620 [FREE Full text] [doi: 10.1038/s41598-021-87059-4] [Medline: 33828118]
80. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. 2021 Presented at: International Conference on Learning Representations, 2021; 2021; Vienna.
81. Chi NA, Washington P, Kline A, Husic A, Hou C, He C, et al. Classifying Autism From Crowdsourced Semistructured Speech Recordings: Machine Learning Model Comparison Study. JMIR Pediatr Parent 2022 Apr 14;5(2):e35406 [FREE Full text] [doi: 10.2196/35406] [Medline: 35436234]
82. Varma M, Washington P, Chrisman B, Kline A, Leblanc E, Paskov K, et al. Identification of Social Engagement Indicators Associated With Autism Spectrum Disorder Using a Game-Based Mobile App: Comparative Study of Gaze Fixation and Visual Scanning Methods. J Med Internet Res 2022 Feb 15;24(2):e31830 [FREE Full text] [doi: 10.2196/31830] [Medline: 35166683]

Abbreviations

ASD: autism spectrum disorder
CNN: convolutional neural network
FER: facial expression recognition
FPS: frames per second
HIPAA: Health Insurance Portability and Accountability Act
LSTM: long short-term memory
NT: neurotypical

Edited by A Mavragani; submitted 30.05.22; peer-reviewed by G Nneji, W Shadid, T Zhang; comments to author 22.06.22; revised version received 25.07.22; accepted 09.08.22; published 14.11.22

Please cite as:
Sleiman E, Mutlu OC, Surabhi S, Husic A, Kline A, Washington P, Wall DP
Deep Learning–Based Autism Spectrum Disorder Detection Using Emotion Features From Video Recordings: Model Development and Validation
JMIR Biomed Eng 2022;7(2):e39982
URL: https://biomedeng.jmir.org/2022/2/e39982
doi: 10.2196/39982
PMID: