Workshops at 18th International Conference on Intelligent Environments (IE2022)
H.H. Alvarez Valera and M. Luštrek (Eds.)

doi:10.3233/AISE220022

Unsupervised Domain Adaptation for Facial Emotion Recognition in Autistic Children

Asha Kurian[a,1] and Shikha Tripathi[b,2]
[a] *Dept. of MCA, PES University, Bangalore, India*
[b] *Dept. of ECE, PES University, Bangalore, India*

Abstract. Autism is a neurodevelopmental disorder characterized by deficits in social, interpersonal interaction and communication skills. A generalized facial emotion recognition model does not scale well when confronted with the emotions of autistic children due to the domain shift inherent in the distributions of the source (neurotypical) and the target (autistic) population. The dearth of labeled datasets in the field of autism exacerbates the problem. Domain adaptation using a generative adversarial model (GAN) counters this disparity by creating an adversarial model that aligns features of the source and target domains using adversarial training. This paper looks at building a facial emotion classifier model that can identify the idiosyncrasies associated with an autistic child's facial expression by generating feature-invariant representations of the source and target distribution. The objective of the paper is two-fold – a) build a discriminative classifier to identify the emotions of autistic children accurately b) to train a feature generator to produce an invariant feature representation of the source and target domains taking into account their similar yet different data distributions, in the presence of unlabeled target data. Investigation into automatic recognition and classification of the facial expressions of the autistic population has not been pursued extensively vis-a-vis a neurotypical population due to the complexities associated with eliciting and interpreting data obtained from autistic children.

Keywords. Autism, Emotion Recognition, Generative Adversarial Modeling, Domain Adaptation

1. Introduction

Affect recognition is the ability to automatically deduce the internal affective states of an individual. Inner emotions of a person have been quantified by using sensors that measure the tangible manifestations of affect, namely facial expression, gestures, voice, and physiological signals [1,2]. Human emotions like happiness, sadness, fear, etc. are typically manifested non-verbally through facial expressions [1]. But these expressions of affect cannot be generalized across people. Various factors like personality traits (introvert vs extrovert), cultural background, ethnicity, and facial physiology determine the extent and nature of emotions displayed through facial expressions [3].

Children affected by autism spectrum disorder (ASD) have impaired emotion recognition and emotion expression skills [4]. They find it difficult to express,

[1] Research Scholar, PES University, Bangalore, India; E-mail: ashaj@pes.edu
[2] Head of Research, PES University, Bangalore, India; E-mail: shikha@pes.edu

acknowledge their own emotions (facial emotion expression) or understand those of their peers (facial emotion recognition) [5,6,7]. Due to the atypicality of emotions exhibited by autistic children, facial expression markers cannot provide an efficient indicator of the mental disposition of the individual. As extrapolated by various studies [8, 9, 10], episodes of anxiety attacks, depression, and schizophrenia are predominantly high in children diagnosed with ASD.

A generalized emotion recognition model cannot be standardized in such scenarios because there is no one standard that covers all categories of people and the range of emotions they display. The objective of this paper is to develop a discriminative classifier that can identify the facial emotions of autistic children, in the absence of annotated datasets. Representative features of the target ASD population are oriented to that of a typically developing (TD) source population by exploiting the power of massive annotated datasets available in plenty in the source domain. Sensing the true inner emotive states of individuals and regulating them can be life-saving in many situations in case of neurological and cognitive disorders like schizophrenia, dementia, autism, depression, PTSD, etc. It also equips the caregiver or clinician to administer timely intervention and effective applied behavior therapy [11, 12].

2. Generative Adversarial Networks (GAN)

GANs follow a paradigm of deep (unsupervised or semi-supervised) generative modeling that enable neural networks to learn feature representations from data distributions in the absence of extensively annotated data [13]. GAN training can be characterized as an adversarial duel between two competing neural networks, a generator G and a discriminator D trained in unison, primarily to generate photo-realistic images from a random noise distribution [14]. GANs have now been adapted to a wide range of computing disciplines like computer vision, image synthesis, image-to-image translation, image super-resolution, video synthesis, and natural language processing. This has led to the evolution of GANs in a variety of extensions - DCGAN, AuxGAN, cGAN, InfoGAN, StyleGAN, CycleGAN, Pix2Pix, and BigGAN [15,16,17,18].

The competing neural network models are typically constituted from a stack of convolutional and fully connected dense layers. The generator learns to produce a distribution of images from a latent space, which bears a close semblance to the data distribution of the real images. The discriminator presented with both the synthetic and real images becomes an expert at tagging them as fake/real respectively [19,20,21]. Updating the model weights is based on the optimization strategy to maximize the discriminator objective function $E_x[\log(D(x)] + E_z[\log(1-D(G(z)))]$ and minimize the generator cost function $E_z[\log(1-D(G(z)))]$.

2.1. Unsupervised Domain Adaptation

Deep neural architectures learn meaningful feature representations in the presence of massive amounts of annotated data. Often, obtaining large, annotated datasets can be cumbersome, expensive, time-consuming, and sometimes infeasible. Domain adaptation adapts the concept of adversarial training in GANs to build classifiers that predict labels from an unlabeled target domain by harnessing the power of extensive, labeled datasets available in a related source domain, albeit with a shift in data distribution [22, 23, 24, 25, 26]. This approach has been applied in situations where the target data is partially labeled (semi-supervised domain adaptation) or totally unlabeled (un-supervised domain adaptation) [27]. In the course of training, a deep feed-forward network with the domain adaptation component embedded enables a decision model to generate feature vectors that are discriminative for a classifier to label [28]. Concurrently, the learned feature representations from both the domains are mapped together to have a near-identical data distribution so that a shift in the domain distribution of the source and target does not impede the classifier's judgment [27].

3. Methodology

The source domain constitutes images from the neurotypical population and the target domain distribution comprises images of facial expressions of autistic children. Figure 1 depicts the proposed architecture for unsupervised domain adaptation.

The architecture comprises of three components:

i) Feature Generator: Constituted from layers of convolution neural network, it doubles as an extractor of feature vectors for the source domain as well as a domain invariant feature generator for both the distributions put together. The embeddings of the source images coupled with the class labels (source dataset is labeled) is provided as input to the emotion classifier model.
ii) Label Classifier: Predicts an emotion label by interpreting the feature embeddings provided by the feature generator.
iii) Domain Discriminator: Takes in features of both source and target images given by the generator and generates a prediction as to its true domain.

In a typical training iteration, each of the network components is involved in training as follows - The feature generator extracts embeddings from source images, coupled with their labels are fed into an emotion classifier to train on. Simultaneously, the discriminator presented with images from both the source and the target (labeled as real/fake), is a binary classifier that gets better at pronouncing a verdict as to the domains they originate from.

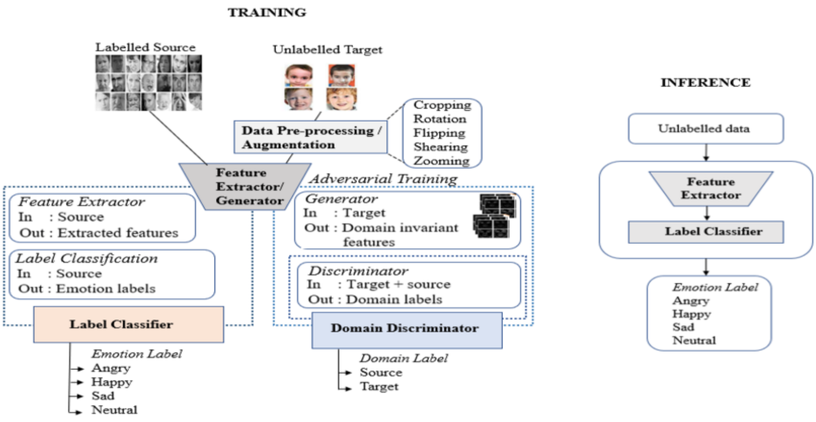

Figure 1. Proposed architecture for unsupervised domain adaptation.

Both the classifier and the discriminator networks are tuned independently and in isolation on mini-batches of data for the dual tasks of emotion classification and domain discrimination respectively. Unsupervised domain adaptation pits the generator and discriminator against each other in adversarial training. The feature generator weights are updated only as part of the GAN adversarial training. During GAN training, images from both the source and target with their labels interchanged are passed to the generator. The extracted feature vectors are further consumed by the discriminator. This rings in domain confusion with a corresponding increase in the GAN loss, a standard cross-entropy loss. As a ripple effect, the generator updates its weights as part of the backpropagation process with a goal to minimize this loss. The GAN is stabilized only when the discriminator and the GAN losses are in equilibrium. Network weights are updated, after each batch is encountered, by taking samples from the source for label prediction, and both the source and target for domain adaptation. Figure 2 details the architecture of the neural network model.

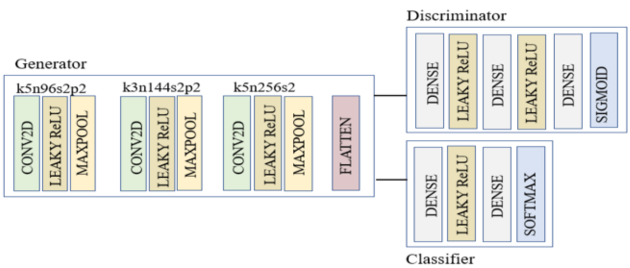

Figure 2. Architecture of the network with kernel size (k), number of channels (n), stride (s), and pool size(p)

3.1. Datasets Used

The GAN was evaluated using two renowned datasets widely used in facial emotion recognition studies, FER-2013 and CK+ datasets. FER-2013 contains around 28000 images, 48X48 in dimension, of individuals expressing varied emotion like anger, sad, happy and neutral. CK+ dataset consists of around 920 images of individuals showing varied expressions of facial affect. Images from these datasets are considered as the source domain. Preprocessing of the datasets was performed to normalize the images before feature extraction.

To obtain a target dataset of autistic children was difficult since there is no corpus of emotion expressions of autistic children. The target dataset consists of the freely available Autism Dataset available on Kaggle that comprises of 1200 images. A holdout manually annotated test data was set aside from the Kaggle dataset for evaluation of the GAN.

3.2. Data Pre-processing and Data Augmentation

Faces were detected and cropped from the original images in the target dataset using the MTCNN architecture for face detection. These images were further normalized, rescaled to 48X48 pixels, and the grayscale version was taken as the target domain. To alleviate the problem of an unbalanced and small domain set, the target data was further subjected to a slew of data augmentation approaches. Images were rotated at an angle of 40^{o}, horizontally flipped to provide a mirror of the image in the horizontal direction, distorted versions generated via shearing, and randomly zoomed. At each timestep during training, a batch of 50 samples from the source dataset was combined with a batch of 50 augmented target samples to form a mini-batch and presented to the domain discriminator.

4. Results and Discussion

GAN models are notoriously hard to train as it involves the concurrent training of two competing neural network models, a feature generator and a domain discriminator. The neural network models are in competition with each other. An improvement in the loss/accuracy of one model comes at the cost of degradation in the other.

4.1. Experimental Set-up

Various network architectures and hyperparameter values have been experimented for the generator, classifier, and discriminator models in accordance with the training heuristics essayed by experts [29, 30, 31, 32, 36]. Unlike normal deep neural networks where high accuracy of the model indicates convergence, GANs are said to attain equilibrium when both the discriminator and generator losses are in a balance. Contrary to the usual practice of tagging images from the source and target domains with hard labels, some best practices like providing a random sample of soft labels for training the discriminator was followed in this experiment.

4.2. Metrics for GAN Evaluation

The performance of the model has been graded taking into consideration three evaluation measures.

- Convergence of the discriminator and GAN loss
- Accuracy score of the GAN model
- Comparison against VGG-16 pre-trained model and transfer learning
- Comparison against previous work done on emotion recognition in the field of autism

The generator and discriminator losses go through a lot of fluctuation as each of the network models tries to outdo the other. As each of the network components gets better at their task, their corresponding losses start to stabilize. Accuracy score is a metric used in the evaluation of GANs that indicates the disparity between the predicted values by GAN and the actual test set values [33, 34, 35]. It is a reasonable indicator of the task performance of a network that has been trained on the facial emotions of neurotypical adults and tested on the facial emotion expressions of Autistic children. An approximate accuracy of 71.53% is obtained against the holdout Autism test dataset. No data augmentation was employed during testing.

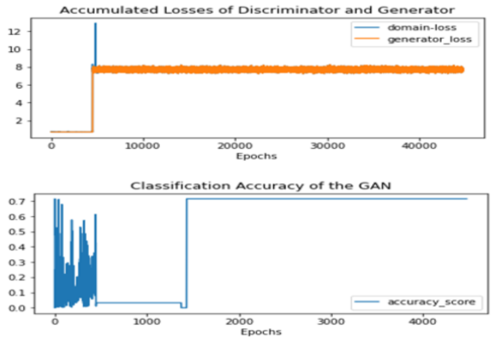

Figure 3. Line plot of the accumulated losses and accuracy score of the model trained with FER-2013.

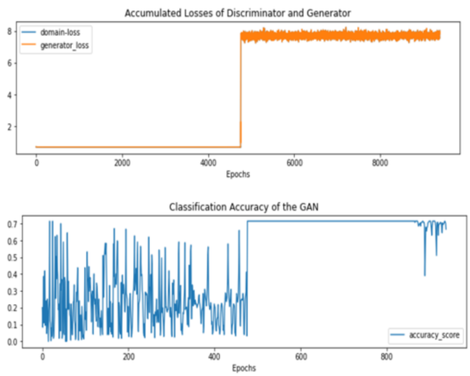

Figure 4. Line plot of the accumulated losses and accuracy score of the model trained with CK+

As indicated by the line plots in Figure 3 and 4, the accumulated losses of the generator and the discriminator converge around 4800 epochs with an accuracy of 71.533% for the holdout test data. The same results were obtained when the GAN was trained on both the FER-2013 and CK+ datasets. The experiments were repeated ten times with each dataset to discount any variations that may have occurred as part of training. Table 1 displays the accuracy score of the GAN model trained on FER-2013 and CK+ datasets and tested on the Autism dataset.

Table 1. Comparison of the accuracy score on FER-2013 and CK+ datasets

Source Dataset	Target Dataset	Epoch at Convergence	Accuracy Score
FER-2013	Kaggle Autism dataset	~5000	71.533%
CK+	Kaggle Autism dataset	~5000	71.533%

A comparison of the GAN model was also made against a pre-trained VGG16 model that had been fine-tuned with the FER-2013 dataset for 5000 epochs and evaluated on the Autism dataset. As indicated by Figure 5, the training accuracy touches a remarkable rate of 99.1%, but the test time accuracy remains static at around 60 % when transfer learning is employed, a typical indicator of data overfitting.

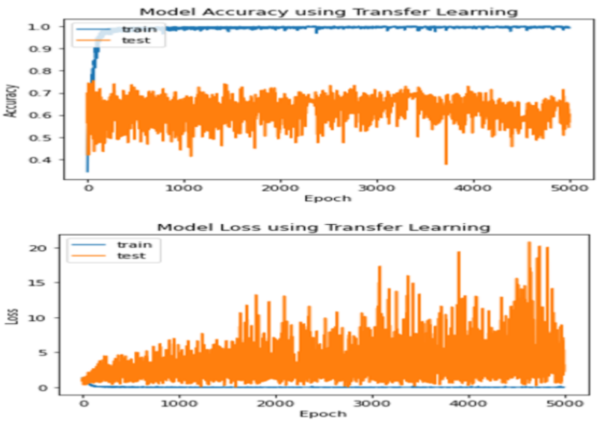

Figure 5. Training and test accuracy/loss of VGG-16 using transfer learning

Numerous studies have attempted to recognize the emotions of Autistic children using data elicited from facial images and/or physiological signals summarized in Table 2.

Table 2. Previous work done on emotion recognition in Autism

Author	Data Collection	Sample Size	Classifier	Ground Truth	Accuracy	Evaluation Metric/ Comparison
Krupa et al, (2016) [37]	GSR, HRV on stimuli	N = 10 ASD N = 10 TD	SVM recognizes emotions: neutral, happy, involvement	Analysis of variation in signals.	Neutral:93.3 % Happy, involvement: 90%	Sensitivity, specificity, accuracy/ Comparison with TD samples
Chu, Hui-Chuan, et al., (2018) [38]	Facial images on applying stimuli	N = 15 ASD	SVM recognizes basic and transition emotions	Manual tagging by parent / experts	98%	F1-score, precision, recall, AUC/ Compared to a neural network
Sarabadani, Sarah, et al., (2018) [39]	Facial images on applying stimuli	N = 15 ASD	KNN, LDA, SVM and ensemble classifier.	Manual tagging by parent / child	78%	Comparison of the accuracy of the mentioned classifiers
Rudovic, Ognjen, et al., (2018) [40]	Multimo-dal	N = 35 ASD	Autoencoder for multimodal data fusion	Manual annotatio n by experts	ICC ~60%	Intra-class co-relation
Puli, Akshay, and Azadeh Kushki, (2019) [41]	ECG and acceleromet ry signals	N = 15 ASD	Modified Kalman filter for anxiety detection	Analysis of variation in signals	93%	Sensitivity, specificity/ Comparison of accuracy of KNN, LR, SVM, Decision Trees

Kalantarian, Haik, et al., (2019) [42]	Facial images on game playing	N = 8 ASD	Ensemble based classifiers on the cloud.	Manual annotation by parent	Disgust: 94%, Neutral: 81%, Surprise: 92%, Scared: 56%	Comparison of the accuracy of ensemble classifiers
Han, Jing, et al, (2020) [43]	Facial images from HRI	FERET, CK+, N = 15 ASD	RVFLN and transfer learning	Annotation by the system	87%	Comparison of accuracy with stacked autoencoder and CNN
Jarraya, Salma Kammoun, et.al, (2020) [44]	Video samples of ASD in normal and meltdown state	N = 13 ASD	Feed Forward, Cascade Feed Forward, RNN and LSTM	Analysis of video by expert	RNN: 85.8%	Comparison with the other classifiers

Investigation into automatic recognition and classification of the facial expressions of the autistic population has not been pursued extensively vis-a-vis a neurotypical population. Undoubtedly, this has been due to the complexities associated with eliciting and interpreting data obtained from autistic children. Reference [37] considers only the presence or absence of emotion, while [38] traces the transition of emotion. [39], [41] look at detecting only anxiety, [40] gives a low accuracy, and [44] at emotions during a meltdown state. The proposed work, in comparison, classifies basic emotions with an accuracy of 71.53%. Additionally, the existing corpus has worked with a small cross-section of the autistic population that is not indicative of the entire ASD population [42], [43]. In this paper, a substantially larger dataset of autistic children has been considered. Coupled with data augmentation, the facial images seen by a deep neural network multiply manifold during training. The technique of classification proposed in this paper is also significantly different. The impact of the classification technique and the accuracy obtained can be gauged from a standpoint, taking into account the challenging task of acquiring labeled data of ASD children that has been validated by manual annotators and/or experts. The proposed method is designed to perform in the presence of unlabeled datasets by leveraging the power of massive amounts of data available in a similar, related domain.

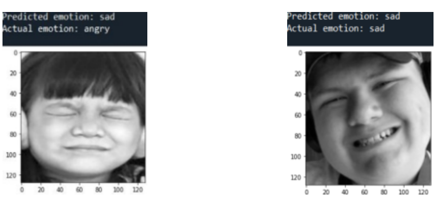

Figure 6. False and true predictions on images in the dataset

Conclusion

This paper explores an approach of unsupervised domain adaptation to classify the facial emotions of children diagnosed with autism. Autistic children experience deficits in recognizing and expressing their own emotions. Hence, generalized classifiers fail at accurately detecting the emotions of ASD children, taking into account the variability in their facial expressions. Also, labeled datasets in the field of autism are practically non-existent. These challenges can be circumvented by developing deep neural models with the proficiency to learn data distributions from unannotated data that is significantly identical to that of the labeled data. This work explores the feasibility of building classifiers to recognize the emotions of autistic children using unsupervised domain adaptation. The advantages of using unsupervised domain adaptation are two-fold. It eliminates the requirement of amassing large annotated datasets that are essential for any classifier to train on. At the same time, also handles the shift in domain distribution of the source and target domains. At the offset, though many of the cited literature exhibit higher accuracy than the proposed method, the accuracy of this work can be improved by training on a larger dataset of real-time ASD children's data. Various other cross-validation metric measures can also be applied for a more subjective evaluation.

References

[1] Tarnowski P, Kołodziej M, Majkowski A, Rak RJ. Emotion recognition using facial expressions. Procedia Computer Science. 2017 Jan 1;108:1175-84.

[2] Tian YI, Kanade T, Cohn JF. Recognizing action units for facial expression analysis. IEEE Transactions on pattern analysis and machine intelligence. 2001 Feb;23(2):97-115.

[3] Kanade T, Cohn JF, Tian Y. Comprehensive database for facial expression analysis. InProceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580) 2000 Mar 28 (pp. 46-53). IEEE.

[4] Smith, Miriam J. Law, et al. "Detecting subtle facial emotion recognition deficits in high-functioning autism using dynamic stimuli of varying intensities." Neuropsychologia 48.9 (2010): 2777-2781.

[5] Harms, Madeline B., Alex Martin, and Gregory L. Wallace. "Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies." Neuropsychology review 20.3 (2010): 290-322.

[6] Uljarevic, Mirko, and Antonia Hamilton. "Recognition of emotions in autism: a formal meta-analysis." Journal of autism and developmental disorders 43.7 (2013): 1517-1526.

[7] Grossard, Charline, et al. "Children with autism spectrum disorder produce more ambiguous and less socially meaningful facial expressions: an experimental study using random forest classifiers." Molecular Autism 11.1 (2020): 1-14.

[8] Zahn TP, Rumsey JM, Van Kammen DP. Autonomic nervous system activity in autistic, schizophrenic, and normal men: effects of stimulus significance. Journal of Abnormal Psychology. 1987 May;96(2):135.

[9] Hirstein, William, Portia Iversen, and Valayanur S. Ramachandran. "Autonomic responses of autistic children to people and objects." Proceedings of the Royal Society of London. Series B: Biological Sciences 268.1479 (2001): 1883-1888.

[10] Sato, Wataru, et al. "Impaired social brain network for processing dynamic facial expressions in autism spectrum disorders." BMC neuroscience 13.1 (2012): 99.

[11] KALIOUBY RE, Picard R, Baron‐Cohen S. Affective computing and autism. Annals of the New York Academy of Sciences. 2006 Dec;1093(1):228-48.

[12] Amaral DG, Schumann CM, Nordahl CW. Neuroanatomy of autism. Trends in neurosciences. 2008 Mar 1;31(3):137-45.

[13] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. Advances in neural information processing systems. 2014;27.

[14] Wang Z, She Q, Ward TE. Generative adversarial networks in computer vision: A survey and taxonomy. ACM Computing Surveys (CSUR). 2021 Feb 9;54(2):1-38.

[15] Zhu JY, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. InProceedings of the IEEE international conference on computer vision 2017 (pp. 2223-2232).

[16] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434. 2015 Nov 19.

[17] Mirza M, Osindero S. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784. 2014 Nov 6.

[18] Chen X, Duan Y, Houthooft R, Schulman J, Sutskever I, Abbeel P. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. Advances in neural information processing systems. 2016;29.
[19] Dumoulin V, Belghazi I, Poole B, Mastropietro O, Lamb A, Arjovsky M, Courville A. Adversarially learned inference. arXiv preprint arXiv:1606.00704. 2016 Jun 2.
[20] Donahue J, Krähenbühl P, Darrell T. Adversarial feature learning. arXiv preprint arXiv:1605.09782. 2016 May 31.
[21] Denton EL, Chintala S, Fergus R. Deep generative image models using a laplacian pyramid of adversarial networks. Advances in neural information processing systems. 2015;28.
[22] Baktashmotlagh M, Harandi MT, Lovell BC, Salzmann M. Unsupervised domain adaptation by domain invariant projection. InProceedings of the IEEE International Conference on Computer Vision 2013 (pp. 769-776).
[23] Ben-David S, Blitzer J, Crammer K, Kulesza A, Pereira F, Vaughan JW. A theory of learning from different domains. Machine learning. 2010 May;79(1):151-75.
[24] Glorot X, Bordes A, Bengio Y. Domain adaptation for large-scale sentiment classification: A deep learning approach. InICML 2011 Jan 1.
[25] Gong B, Grauman K, Sha F. Connecting the dots with landmarks: Discriminatively learning domain-invariant features for unsupervised domain adaptation. InInternational conference on machine learning 2013 Feb 13 (pp. 222-230). PMLR.
[26] Csurka G, editor. Domain adaptation in computer vision applications. Springer International Publishing; 2017 Sep 10.
[27] Saenko K, Kulis B, Fritz M, Darrell T. Adapting visual category models to new domains. InEuropean conference on computer vision 2010 Sep 5 (pp. 213-226). Springer, Berlin, Heidelberg.
[28] Gopalan R, Li R, Chellappa R. Domain adaptation for object recognition: An unsupervised approach. In2011 international conference on computer vision 2011 Nov 6 (pp. 999-1006). IEEE.
[29] Chintala S, Denton E, Arjovsky M, Mathieu M. How to train a GAN? Tips and tricks to make GANs work. Github. com. 2016 Dec.
[30] Salimans T, Goodfellow I, Zaremba W, Cheung V, Radford A, Chen X. Improved techniques for training gans. Advances in neural information processing systems. 2016;29.
[31] Arjovsky M, Bottou L. Towards principled methods for training generative adversarial networks. arXiv preprint arXiv:1701.04862. 2017 Jan 17.
[32] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. InInternational conference on machine learning 2015 Jun 1 (pp. 448-456). PMLR.
[33] Shmelkov K, Schmid C, Alahari K. How good is my GAN?. InProceedings of the European Conference on Computer Vision (ECCV) 2018 (pp. 213-229).
[34] Ravuri S, Vinyals O. Classification accuracy score for conditional generative models. Advances in neural information processing systems. 2019;32.
[35] Lesort T, Stoian A, Goudou JF, Filliat D. Training discriminative models to evaluate generative ones. InInternational Conference on Artificial Neural Networks 2019 Sep 17 (pp. 604-619). Springer, Cham.
[36] Arora S, Ge R, Liang Y, Ma T, Zhang Y. Generalization and equilibrium in generative adversarial nets (gans). InInternational Conference on Machine Learning 2017 Jul 17 (pp. 224-232). PMLR.
[37] Krupa N, Anantharam K, Sanker M, Datta S, Sagar JV. Recognition of emotions in autistic children using physiological signals. Health and technology. 2016 Jul;6(2):137-47.
[38] Chu HC, Tsai WW, Liao MJ, Chen YM. Facial emotion recognition with transition detection for students with high-functioning autism in adaptive e-learning. Soft Computing. 2018 May;22(9):2973-99.
[39] Sarabadani S, Schudlo LC, Samadani AA, Kushski A. Physiological detection of affective states in children with autism spectrum disorder. IEEE Transactions on Affective Computing. 2018 Mar 27;11(4):588-600.
[40] Rudovic O, Lee J, Dai M, Schuller B, Picard RW. Personalized machine learning for robot perception of affect and engagement in autism therapy. Science Robotics. 2018 Jun 27;3(19):eaao6760.
[41] Puli A, Kushki A. Toward automatic anxiety detection in autism: a real-time algorithm for detecting physiological arousal in the presence of motion. IEEE Transactions on Biomedical Engineering. 2019 May 27;67(3):646-57.
[42] Kalantarian H, Jedoui K, Washington P, Tariq Q, Dunlap K, Schwartz J, Wall DP. Labeling images with facial emotion and the potential for pediatric healthcare. Artificial intelligence in medicine. 2019 Jul 1;98:77-86.
[43] Han J, Xie L, Liu J, Li X. Personalized broad learning system for facial expression. Multimedia Tools and Applications. 2020 Jun;79(23):16627-44.
[44] Jarraya SK, Masmoudi M, Hammami M. Compound emotion recognition of autistic children during meltdown crisis based on deep spatio-temporal analysis of facial geometric features. IEEE Access. 2020 Apr 8;8:69311-26.