OSVGAN: Generative Adversarial Networks for Data Scarce Online Signature Verification

Chandra Sekhar Vorugunti
IIIT SriCity
Chittoor-Dt, 517 646
Andhra Pradesh, India
Chandrasekhar.v@iiits.in

Sai Sasikanth Indukuri
University of Massachusetts
Amherst
MA 01003, US.
sindukuri@umass.edu

Viswanath Pulabaigari
IIIT SriCity
Chittoor-Dt, 517 646
Andhra Pradesh, India.
viswanath.p@iiits.ac.in

Rama Krishna Sai Gorthi
IIT Tirupati
Chittoor-Dt, 517 506
Andhra Pradesh, India.
rkg@iittp.ac.in

Abstract

Impractical to acquire a sufficient number of signatures from the users and learning the inter and intra writer variations effectively with as minimum as one training sample are the two critical challenges need to be addressed by the Online Signature Verification (OSV) frameworks. To address the first challenge, we are generating writer specific synthetic signatures using Auxiliary Classifier GAN, in which a generator is trained with a maximum of 40 signature samples per user. To address the second requirement, we are proposing a Depth wise Separable Convolution based Neural Network, which results in achieving one shot based OSV with reduced parameters. A first of its kind of experimental analysis is done with an increased set of signature samples (five-fold) on two widely used datasets SVC, MOBISIG. The state-of-the-art outcome in almost all categories of experimentation confirms the competence of the proposed OSV framework and qualifies for the real time deployment in limited data applications.

1. Introduction

Signatures encompass an aggregation of individual writing characteristics which are a significant source of information to classify the genuineness of a user trying to login into the system. Based on the data acquisition, OSV systems are classified into offline or online [1,2,7,22,23]. In case of offline signatures, only the static information, i.e. X-axis, Y-axis profiles are available in an image format for verification. In case of online signatures, as shown below, along with X, Y profiles, the dynamic information includes, the pressure, pen angle, tilt of a device, etc. Due to availability of both static (X, Y profile) and dynamic information, OSV frameworks tend to be more robust and accurate.

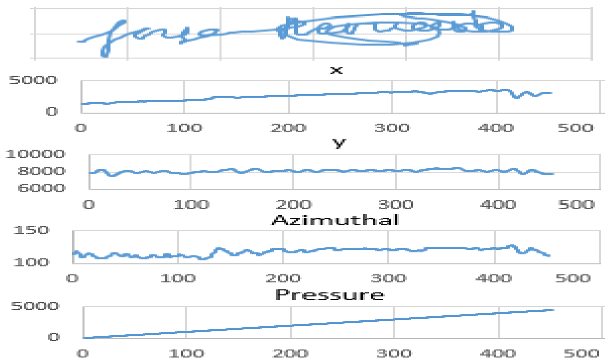

TABLE I. THE PARTICULARS OF WIDELY USED DATASETS IN OSV

DataSet→	SVC	MobiSig	MCYT
Total number of users	40	83	100
Genuine, Forgery samples per user	20,20	40,25	25,25

In literature, several approaches towards online signature verification (OSV) had been detailed which can be primarily grouped into feature-centric techniques [13, 20] that interpret signatures using a collection of local or global features, function-centric techniques which apply distinct methods such as Hidden Markov models [1], Dynamic Time Warping (DTW) [1,10,11], averaging time series [26], interval valued [13,20], sequence matching [10,26], feature fusion based [13], fuzzy based [20], stroke based [18, 24], deep learning based [23,24,25,27] and many more. Recent works [32] confirms that, individual profiles (x, y, pressure, azimuthal) results into lightweight frameworks and higher classification accuracies compared to the compound signature. Hence, in this work, we focus on generating writer specific synthetic profiles to evaluate the genuineness of the test signature.

Even though several OSV frameworks were proposed, still there is a shortfall in OSV systems addressing critical requirements: R1. One/Few shot learning: A light weight OSV framework, which can effectively learn to classify a test signature, when trained with one signature sample per user. R2. An OSV framework, must be tested with more signature samples, to be ratified to deploy in real time environment. Even though very few works are proposed to address R1 [7,14,18,25], based on authors' knowledge, no work is proposed to address R2, as acquiring a greater number of signature samples per user is impractical. As Table 1 suggests, maximum number of signature samples per user 40, which is very less compared to other computer vision problems like Object Detection [31] etc. Hence, to address the above two requirements, our contribution in this work is two-fold:

1. As represented in Fig 1, a novel variant of Auxiliary Classifier-GAN (AC-GAN) based framework, which generates effective and unlimited writer specific synthetic signature samples.

2. As represented in Fig 2, we propose a Depth Wise Separable Convolution (DWS) based OSV framework, through which we achieve one shot learning with reduced parameters compared to standard convolution based neural networks.

2. Proposed OSV framework

2.1 Synthetic Signature profile generation

As depicted in Fig 1, 3 to generate high-quality writer specific synthetic signatures, we have proposed OSVGAN, which is a modified version of AC-GAN [5,28]. AC-GAN is widely used variant of vanilla GAN, in which, addition to the noise vector $'n'$, a corresponding label, $l \sim P_l$ is given as an input to the generator G to generate writer specific synthetic signatures $S_{syn} = G(l, n)$.

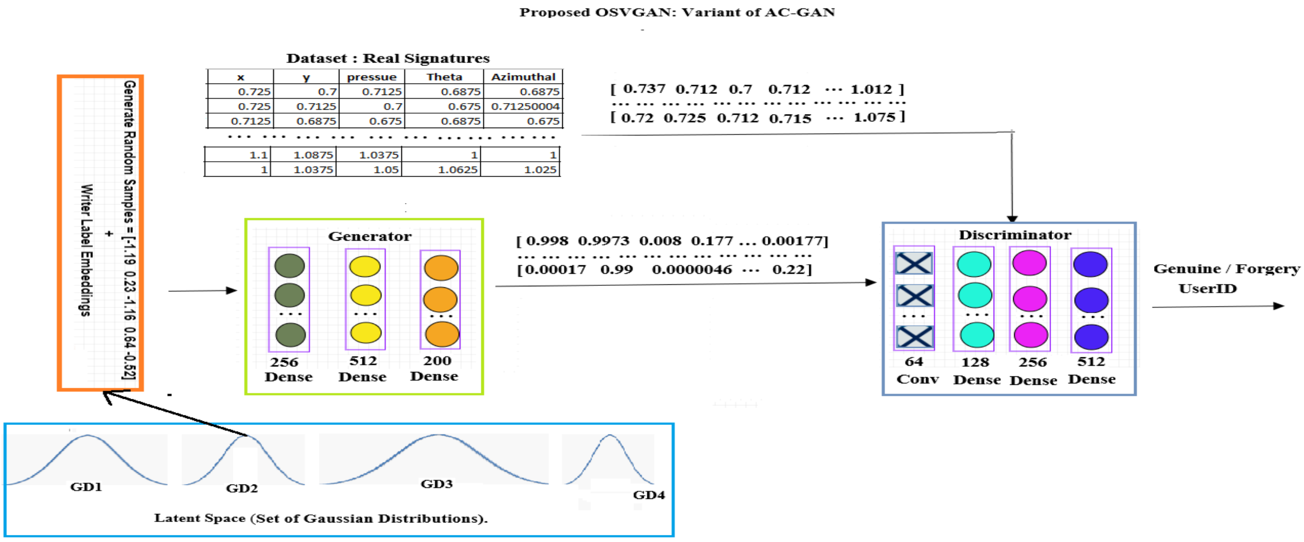

Figure 1. The Proposed OSVGAN architecture, which is a variant of Auxiliary Classifier GAN.

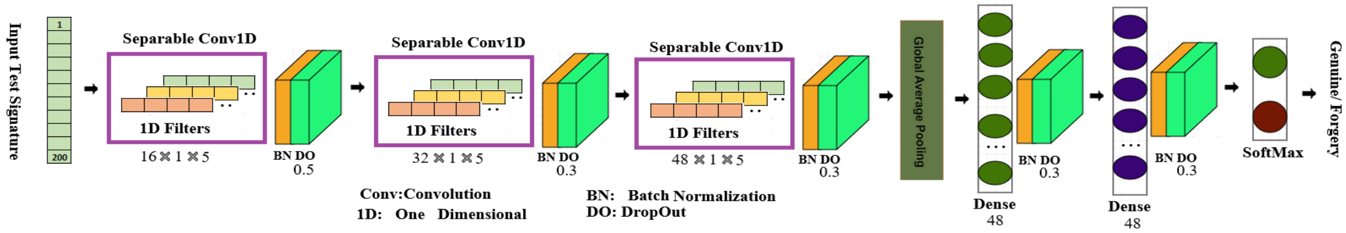

Figure 2. The Proposed Depth Wise Separable Convolution based Neural Network Architecture to classify a test signature.

In vanilla GAN [3,4], Generator G, transforms a random noise 'n' into a 1D vector (signature profiles) or 2D image i.e. $x_G = G(n)$. The noise 'n' typically chosen from an easy-to-sample uniform distribution, typically 'n' $\sim U(-1,1)$. The generator aims to maximize the generated data (synthetic signature profiles) as similar as possible to the target data distribution (signature dataset).

$$P_{data}(S_G) = \int_n P_{data}(S_G, n) = \int_n P_{data}(S_G|n).P_n(n)dn \quad (1)$$

principally GAN attempt to learn a mapping from a basic latent distribution $P_n(n)$ to the complicated data distribution $P_{data}(S_G|n)$. Therefore, the joint optimization problem for the GAN can be represented as given below:

$$Min_G\, Max_D(V(G,D)) = Min_G\, Max_D(E_{x\sim Pdata}[logD(x)]) + E_{z\sim Pz}[\log(1 - D(G(z)))] \quad (2)$$

An Auxiliary Classifier GAN [5], as depicted in Fig1, 'G' takes as input both the class label 'c' (in the signature context, genuine/forgery) and the noise 'n' i.e. $S_{fake} = G(c,n)$. Similarly, the discriminator outputs the probability distributions over signature labels L_S (genuine/forgery) and the class (writer) labels L_w (writer id) i.e. $P(S \mid X), P(L \mid X) = D(X)$. The discriminator's objective function is represented as below:

$$L_S = E[\log P(S = real \mid X_{real})] + E[\log P(S = fake \mid X_{fake})] \quad (3)$$
$$L_w = E[log\, P(L = w \mid X_{real})] + E[log\, P(L = w \mid X_{fake})] \quad (4)$$

The generator and the discriminator compete to maximize L_S - L_W and $L_S + L_W$ respectively. Recently, Swaminathan et al [4] proposed a novel attempt in which, to increase the modelling power of the prior distribution, they have reparametrized [6] the latent generative space of vanilla GAN into a set of Gaussian mixture models and learn the best mixture model specific to each writer. Motivated by Swaminathan et al [4] work, we have reparametrized the latent generative space of Auxiliary Classifier GAN into a set of mixture models and learn the best mixture model specific to each writer.

$$P_z(z) = \sum_{g=1}^{G} \emptyset_i . g(z|\mu_g, \Sigma_g) \quad (5)$$

where $g(z|\mu_g, \Sigma_g)$ represents the probability of the sample z in the normal distribution $N(\mu_g, \Sigma_g)$.

Assuming uniform mixture weights i.e. $\emptyset_i = 1/G$

$$P_z(z) = \sum_{g=1}^{G} \frac{g(z|\mu_g \Sigma_g)}{G} \quad (6)$$

Applying "reparameterization trick" [6] on equation (6), which divides the single Gaussian distribution into 'G' Gaussian distributions. The noise from the i[th] Gaussian distribution is calculated using $z = \mu_i + \sigma_i.\epsilon$, where '$\epsilon$' represents an auxiliary noise variable such that, $\epsilon \sim N(0,1)$. μ_i is a sample from a uniform distribution $U(-1,1)$ and σ_i is set to 0.4. User is advised to read [4] for further analysis.

As depicted in Fig 2, a Gaussian random noise of size 5 is derived from the selected Gaussian distribution and the label embeddings are given as an input the Generator G. The generator generates the corresponding profile of an online signature of size 1*200, which is fed as an input to the discriminator 'D', consists of a one-dimensional convolution layer, followed by three dense layers to classify the synthetic signature profile as real or generated. The generator is trained to generate the synthetic profiles close to the samples from

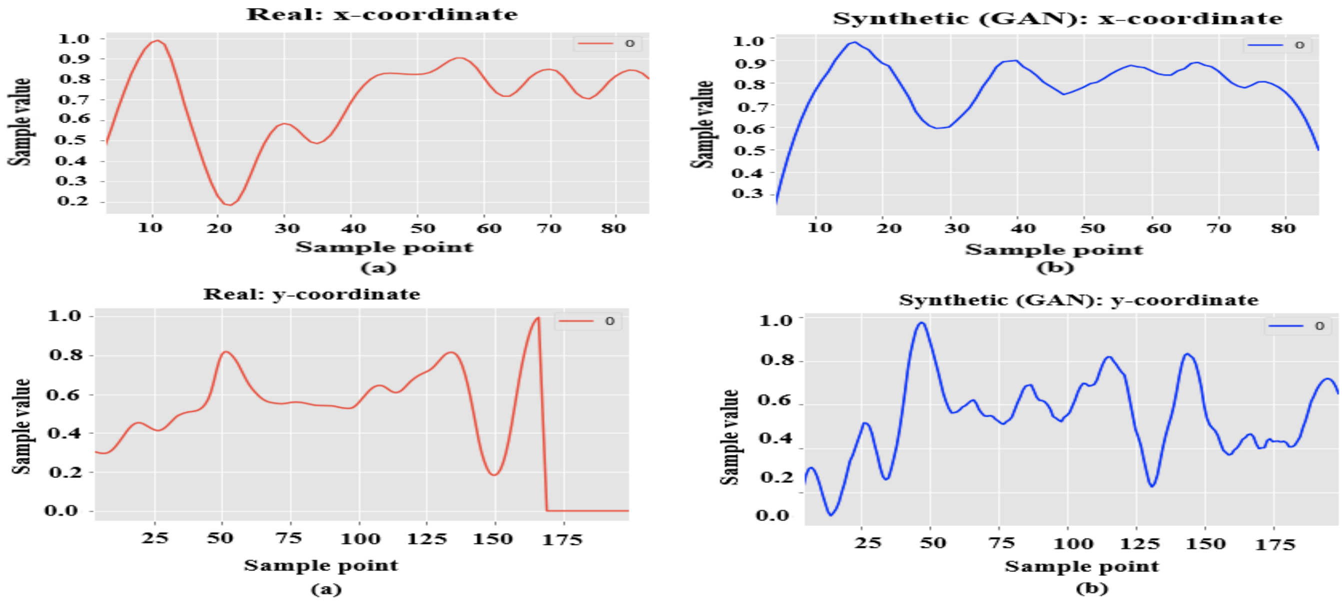

Figure 3. Comparing the real signature profiles (a: red) and the synthetic profiles generated by our proposed model (b:blue).

the target space (signature dataset) through backpropagation of discriminator error in classifying the real and generated signature profiles. Fig 3, depicts the real and synthetic signature profiles generated by our proposed framework.

2.2 Depth-Wise Separable (DWS) Convolution

As depicted in Fig 2, the writer specific synthetic signature profiles generated by our proposed OSVGAN are used during the testing phase of our proposed DWSCNN. Recent works [25, 31], confirms that the DWS convolution outcomes, reduced parameters and operations by a factor of $1/c + 1/(N^2)$ compared to the standard convolution, where the c = number of input channels of an input signature and N = number of kernels. The model proposed in Fig 2 requires 7,350 parameters, whereas the same model with standard convolutions requires 15,361 trainable parameters, which results in a deduction of 47.8% of trainable parameters by using DWS convolutions. DWS convolution is a set of depth-wise convolution and 1×1 point-wise convolution on the outcome of depth-wise convolution (DWC).

$$\text{DepthWiseConv}(\mathrm{I},\mathrm{K})_{(x,y)} = \sum_{a,b}^{A,B} I(x+a, y+b).K_{(a,b)} \quad (7)$$

In DWC, for each input channel ‘c’ , kernel $K_{(a,b)}$ is convolved with an input image $I_{(x,y)}$ to produce an intermediate result. For each input channel, a point wise convolution is carried out on the interim result as given below:

$$\text{PointWiseConv}(\mathrm{I},\mathrm{K})_{(x,y)} = \sum_{c}^{C} W_{(c)} \times f(x,y,c) \quad (8)$$

As depicted in Fig 2, online signature is represented as a (1 * 200 feature vector. If we substitute x = 1, a=1 and b =1, the above equations (7), and (8) represent an online signature which is of one-dimensional feature vectors. A batch normalization is applied on each layer output. A dropout of 50%, 30%, 30% are applied at each DWS layer. The deep representation features captured by the DWS layers passed as an input to the dense layers for classification. A dropout of 30%, 30% are applied at each dense layer. The final SoftMax layer classifies the test signature as genuine/forgery.

3. Experimental Analysis

To appraise our proposed OSVGAN, we have thoroughly evaluated our framework on two extensively used datasets i.e. SVC [5,12] and MOBISIG [20,21]. The experiments are conducted on Ubuntu based GTX1080 GPU machine with 20 GB memory. The proposed framework is experimented with four categories of evaluation, i.e. Skilled_1 (S_01), Skilled_5 (S_05), Skilled_10 (S-10) and Skilled_15 (S_15). Traditionally, if a dataset contains ‘G’ genuine and ‘F’ forgery signature samples per user, in Skilled_N category, for each user, ‘N’ samples of genuine and forgery are used for training and ‘G-N’ and ‘F-N’ samples are used to compute True Acceptance Rate (TAR) and False Acceptance Rate (FAR) per user and an Equal Error Rate (EER) is computed using Receiver Operating Curves (ROC). In this current work, similar to existing works, same number of training samples, i.e. (G-N) are considered to compute TAR. To compute, FAR, apart from (F-N) testing samples, hundred synthetic signature profiles are generated per user using proposed OSVGAN and a total of (F-N) +100 skilled forgery samples are used to compute FAR. We have evaluated our framework using three types of testing samples. 1. Considering both AC-GAN generated synthetic signature samples and handcrafted features. 2. Considering only the AC-GAN generated synthetic samples and 3. Considering only the existing handcrafted features.

Evaluating the model with a greater number of signature samples per user, is a first of its kind of an attempt to address the requirement R1 discussed above. As illustrated in Tables II and III, the proposed OSVGAN realizes one shot learning. The frameworks which results in first highest EER is marked as * and the second highest is marked as **. Even though, the proposed model is evaluated with more testing samples compared to the existing works, in case of SVC, the proposed framework realized state-of-the-art EER in S_01, S_10 and S_15 categories by yielding an EER of 2.86%, 1.42% and 1.07% respectively. As illustrated in table III, in case of MobiSig, the proposed OSVGAN framework yields state of the art EER in all classes of experimentation. In S_01 (one shot learning), the framework achieves an EER of 5.17 with a handcrafted pressure profile.

TABLE II. COMPARISON OF EER (LOWER IS BEST) PERFORMANCE OF VARIOUS RECENT OSV FRAMEWORKS EVALUATED ON SVC DATASET

Technique	S_01	S_05	S_10	S_15
Proposed Model : (GAN+ Handcrafted features): X-Axis	5.17	3.99	4.82	2.75
Proposed Model : (GAN+ Handcrafted features): Y-Axis	4.78	4.2	3.27	2.12
Proposed Model : (GAN+ Handcrafted features): Pressure	6.8	4.27	3.32	1.89
Proposed Model : (Only GAN generated features): X-Axis	2.95	3.14	2.83	2.67
Proposed Model : (Only GAN generated features): Y-Axis	2.86*	5.24	1.42**	1.76
Proposed Model : (Only GAN generated features): Pressure	4.52	2.62	1.48	1.9
Proposed Model : (Handcrafted features): X-Axis	2.87**	3.19	2.71	2.71
Proposed Model : (Handcrafted features): Y-Axis	2.97	5.18	1.46	1.29**
Proposed Model : (Handcrafted features): Pressure	4.6	2.59**	1.23	1.07*
SVM +SPW+ mRMR (10-Samples) [18]	-	-	1.00*	-
LCSS[10]	-	-	5.33	-
Relief-1 [17]	-	-	8.1	-
SPW[18]	-	-	1.00*	-
PDTW(case 2) [22]	-	-	-	-
Relief-2 [17]	-	-	5.31	-
Stroke-Wise [16]	18.25	-	-	-
PCA [16]	-	-	7.05	-
TW [16]	18.63	-	-	-
PDTW [22]	-	-	-	-
Variance selection [17]	-	-	13.75	-
Curvature +Torsion [21]	-	-	6.61	3.10
DTW+ warping path score [11]	-	-		-
RNN+LNPS [14]	-	2.37*	-	-
DTW[9]	-	2.73	-	-
SynSig2Vec -Common Threshold [23]	11.96	4.65	-	-
SynSig2Vec – User Specific Threshold [23]	7.34	-	-	-
Template matching + time-series averaging [26]		2.98	1.80	

TABLE III. COMPARISON OF EER PERFORMANCE OF VARIOUS RECENT OSV FRAMEWORKS EVALUATED ON MOBISIG DATASET.

Technique	S_01	S_05	S_10	S_15
Proposed Model : (GAN+ Handcrafted features): X-Axis	19.74	15.84	14.88	13.3
Proposed Model : (GAN+ Handcrafted features): Y-Axis	17.65	15.01	14.62	13.56
Proposed Model : (GAN+ Handcrafted features): Pressure	13.84	15.05	12.78	12.26
Proposed Model : (Only GAN generated features): X-Axis	14.91	8.52	6.39	6.97
Proposed Model : (Only GAN generated features): Y-Axis	12.78	6.44	5.87	5.68
Proposed Model : (Only GAN generated features): Pressure	7.78**	6.5	2.42**	2.55**
Proposed Model : (Handcrafted features): X-Axis	10.65	6.32	5.87	4.31
Proposed Model : (Handcrafted features): Y-Axis	10.72	6.1*	4.91	4.27
Proposed Model : (Handcrafted features): Pressure	5.17*	6.21**	2.18*	2.14*
Baseline [15]	-	25.45, 19.27	-	-
Stroke-based RNN [24]	16.08, 16.261	-	-	-
Recurrent Adaptation Networks [25]	-	10.87	-	-

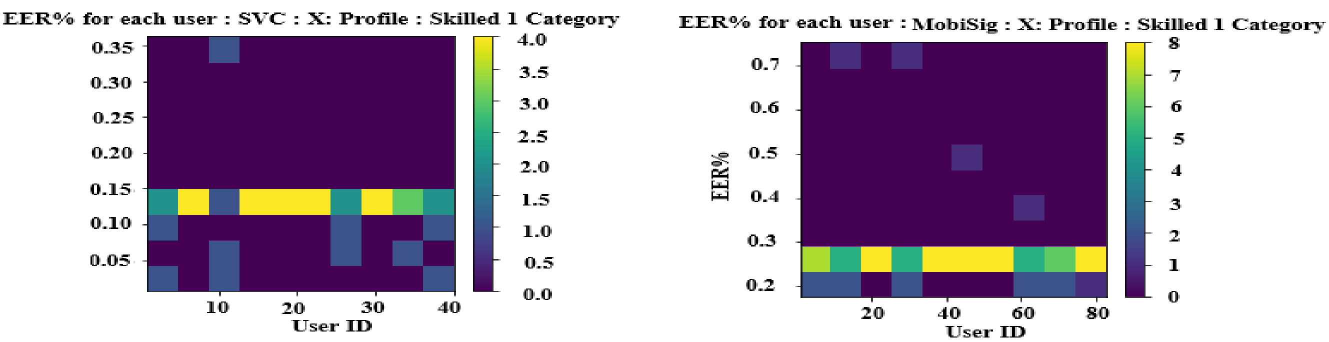

Figure 4. A 2D-histogram representing EER registered for each user in case of SVC and Mobisig datasets under Skilled_1 category.

Figure 4 depicts the EER yielded by the proposed framework for each user in Skilled_1 category of SVC and MobiSig dataset respectively in the form of a 2D-Histogram. Figure 4.a) depicts that the users from 5-10, 15-25, 27-32 contributes to higher EER of the framework compared to others and the average EER varies between 10-15%. Correspondingly, Fig 4.b) delineates that users from 35-60 contributes to higher EER of the framework and the average EER varies between 25 – 30%.

4. Conclusion

In this work, two most challenging requirements of OSV are addressed. First, data scarcity to thoroughly test the framework for real time deployment in critical applications. To address this, we have proposed a first of its kind of an attempt to generate virtually unlimited synthetic signature samples per user from a maximum of 40 signatures per user based on a modified version of AC-GAN. Second, achieving few shot learning, especially one-shot learning to classify the genuineness of test signature with as minimum as one training sample per user. The efficiency of the proposed model is confirmed through state-of-the-art achievement in various categories compared to the frameworks evaluated with reduced number of test samples. In future, to grasp the generative skills of GANs, we will focus on filling the missing and noisy parts of the signatures.

References

[1] B. L. Van, S. Garcia-Salicetti, and B. Dorizzi, “On using the viterbi path along with hmm likelihood information for online signature verification.” IEEE Transactions on Systems, Man, Cybernetics, Part B, vol. 37, no. 5, pp. 1237–1247, 2007.

[2] J. Galbally, J. Fiérrez, M. Diaz, and J. O.Garcia, “Improving the enrollment in dynamic signature verfication with synthetic samples,” in Int Conf on Doc Anal. Recognit. (ICDAR), pp. 1295–1299, Barcelona, Spain, 2009.

[3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In Advances in Neural Information Processing Systems, pages 2672–2680, 2014.

[4] G.Swaminathan,S.R.Kiran and R.V.Babu, DeLiGAN : Generative Adversarial Networks for Diverse and Limited Data, 30TH IEEE conference on Computer viison and pattern recognition (CVPR 2017),pp. 4941-4949,JUL 21-26, 2017.

[5] A.Odena, C.Olah and J.B Shlens, Conditional image synthesis with auxiliary classifier GANs, 34th International Conference on Machine Learning (ICML), vol 70, pp:2642–2651, august 2017.

[6] D.P. Kingma, M.Welling, Auto-Encoding Variational Bayes, International Conference on Learning Representations 2014.

[7] A.Odena, C.Olah, J.Shlens,Conditional Image Synthesis with Auxiliary Classifier GANs,34th International Conference on Machine Learning, ICML,pp:2642-2651,2017.

[8] N. Sae-Bae and N. Memon, “Online signature verification on mobile devices,” IEEE Trans. on Information Forensics and Security, vol. 9, no. 6, pp. 933–947, 2014.

[9] M. Diaz, A. Fischer, R. Plamondon, and M. A. Ferrer, “Towards an automatic on-line signature verifier using only one reference per signer,” in Int. Conf. Document Anal. Recognit. (ICDAR), Tunis, Tunisia, pp. 631–635, 2015.

[10] K.Barkoula, G.Economou, and S.Fotopoulos.: ‘Online signature verification based on signatures turning angle representation using longest common subsequence matching’, Int Journal of Doc Analysis and Recognition, 2013, vol 16, pp. 261–272.

[11] A. Sharma and S. Sundaram, “An enhanced contextual dtw based system for online signature verification using vector quantization,” Pattern Recognition Letters, vol. 84, pp. 22–28, 2016.

[12] A. Fischer and R. Plamondon, “Signature verification based on the kinematic theory of rapid human movements,” IEEE Trans. On Human-Machine Systems, vol. 47, no. 2, pp. 169–180, April 2017.

[13] D. Guru, K. Manjunatha, S. Manjunath, and M. Somashekara, “Interval valued symbolic representation of writer dependent features for online signature verification,” Expert Systems with Applications, vol. 80, pp. 232–243, 2017.

[14] S.Lai, L.Jin, W.Yang: ‘Online Signature Verification Using Recurrent Neural Network and Length-Normalized Path Signature Descriptor’, 14th IAPR Int Conf on Doc Ana and Rec (ICDAR), 2017.

[15] M. Antal, L. Z. Szab´o, and T. Tordai, “Online signature verification on MOBISIG finger-drawn signature corpus,” Mobile Information Systems, vol. 2018, 2018.

[16] M.Diaz, A.Fischer, M. A. Ferrer and R.Plamondon, Dynamic Signature Verification System Based on One Real Signature, IEEE Trasactions On Cybernetics, vol 48, Jan 2018.

[17] Yang, L., Cheng, Y., Wang, X., et al.: ‘Online handwritten signature verification using feature weighting algorithm relief’, Soft Computing, December 2018, Vol 22

[18] B.Kar, A.Mukherjee, P.K.Dutta : ‘Stroke Point Warping-Based Reference Selection and Verification of Online Signature’, IEEE Transactions On Instrumentation and Measurement, Jan 2018, vol. 67

[19] M.Diaz, M.A. Ferrer and J.J.Quintana, Anthropomorphic features for On-line Signatures, IEEE Transactions on Pattern Anaysis and Machine Intelligence, pp:2807 - 2819.,vol 41, Dec 2019.

[20] V.C.Sekhar, P.Mukherjee, D. S. Guru, V.Pulabaigari, Online Signature Verification Based on Writer Specific Feature Selection and Fuzzy Similarity Measure, WORKSHOP ON MEDIA FORENSICS, CVPR, PP:88-96, 2019.

[21] He, L., Tan, H., Huang, Z.: ‘Online handwritten signature verification based on association of curvature and torsion feature with Hausdorff distance’, springer : Multimedia Tools and Applications, Jan 2019.

[22] Al-Hmouz, R., Pedrycz, W., Daqrouq, K., et al.: ‘ Quantifying dynamic time warping distance using probabilistic model in verification of dynamic signatures’, Elsevier-Soft Computing, Jan 2019, vol 23, pp. 407–418

[23] S.Lai, L.Jin, L.Lin, Y.Zhu and H.Mao, SynSig2Vec: Learning Representations from Synthetic Dynamic Signatures for Real-world Verification,arxiv.

[24] C.Li, X.Zhang, F.Lin, Z.Wang, J.Liu and R.Zhang,A Stroke-based RNN for Writer-Independent Online Signature Verification,2019 International Conference on Document Analysis and Recognition (ICDAR), pp:526-532, ICDAR, Australia, 2019.

[25] S.Lai and L.Jin, Recurrent Adaptation Networks for Online Signature Verification, IEEE Transactions on Information Forensics and Security, vol 14 , Issue: 6, pp:1624 - 1637, June 2019.

[26] M.Okawa, Online signature verification using single-template matching with time-series averaging and gradient boosting,Elsevier journal of Pattern Recognition, Volume 102, June 2020.

[27] V.C. Sekhar, G.Rama krishna, P.Viswanath, ‘Online Signature Verification by Few-shot Separable Convolution Based Deep Learning’, 15th International Conference on Document Analysis and Recognition, Sydney, Australia, pp: 1125-1130, 2019.

[28] F.Zhan, H.Zhu, S.Lu; Spatial Fusion GAN for Image Synthesis,The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3653-3662,2019.

[29] E.Heim,Constrained Generative Adversarial Networks for Interactive Image Generation,The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.10753-10761,2019.

[30] J. Li, J.Yang, A.Hertzmann, J.Zhang and T.Xu, LayoutGAN: Synthesizing Graphic Layouts with Vector-Wireframe Adversarial Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence , Jan 2020.

[31] X.Zhou, J.Zhuo, P. Krahenbuhl, Bottom-Up Object Detection by Grouping Extreme and Center Points,The IEEE Conference on Computer Vision and Pattern Recognition (CVPR),pp. 850-859, 2019.

[32] M.Zalasinski, K.Lapa and M.Laskowska,Intelligent Approach to the Prediction of Changes in Biometric Attributes, IEEE TRANSACTIONS ON FUZZY SYSTEMS, Nov 2019.

[33] A.Yang, B.Yang, Z.Ji, Y.Pang and L.Shao, Lightweight group convolutional network for single image super-resolution, Information Sciences, vol 516, pp:220-233, April 2020.