DeepSign - Deep Online Handwritten Signature Verification

Sai Abhinav Reddy Badinehal
Email: saiabhin@buffalo.edu
University At Buffalo

Sai Sandeep Lankisetty
Email: slankise@buffalo.edu
Univeristy At Buffalo

***Abstract*—Accurately verifying online handwritten signatures is still a challenging task.There has been many researches around in recent years applying deep learning methods for the signature verification tasks, but they couldn't outperform traditional machine learning methods. In our work we are going to use Deep learning method called TA-RNN (Time Aligned-Recurrent Neural Network) to tackle this problem. TA-RNN which is combination of Dynamic Time Warping (DTW) and Recurrent Neural Networks (RNN's) will extract more meaningful time features by capturing the temporal dynamics of online handwritten signatures and thereby accurately help in training a robust system against the forgeries. However, to test deep learning to its limits the data must be large, for that we have selected MOBISIG database. Our Approach is able to outperform previous state of art methods by giving 9.03% EER in case of finger skilled forgery scenarios with 4 training signatures per user.**

1. Introduction

Behavioural Biometrics such as handwritten signatures can be used for verifying authenticity of an individual. Handwritten Signatures are of 2 types i) Static Signatures – Also called as Offline signatures which are acquired by signing with an ink pen on the paper and digitizing the image. Verifying these signatures can be done from the captured images. ii) Dynamic Signatures – Also called as Online signatures are acquired by signing over an electronic device which results not only in image but also in capturing temporal dynamics. Here, comes the advantage with Online signatures, temporal dynamics such as pressure, angle, speed, acceleration can be used to verify information precisely. Although the handwritten signatures verification has been evolved a lot recently, still it is considered as a rigorous task because of the limited practical availability of genuine references. In this work we are taking up the task to verify the online handwritten signatures.

Online Handwritten Signatures has 2 ways of writing input i) Stylus – An Electronic instrument with which signatures can be recorded on electronic device. With this the pen up trajectories can also be detected when pen is not in contact with tablet on which we sign. This information can be helpful in verifying the authenticity of user. ii) Finger – Human finger is used to draw signatures. This is difficult to verify because some of additional information such as pen up trajectories will not be available. Subjects may employ different variations between signatures due to finger moments and writing styles. Our work here is mainly focussed in verifying online handwritten signatures which were created using finger as their input.

However, there has been many researches around verifying Online handwritten signatures using traditional machine learning methods such as Support Vector Machine (SVM), Dynamic Time Warping (DTW) etc., deep learning methods are still in race to verify them accurately, its obvious because we know Deep learning models are data hungry models they need to be trained with more data in order to provide accurate results. In field of Online handwritten signatures the publicly available data is scarce which is similar with other biometric traits such as face.

The Authors of this paper [1] were motivated by the above problem and introduced DeepSignDB and TA-RNN deep learning Approach. DeepSignDB is collection of multiple databases with 1526 unique users which has its data created using both stylus and finger. Time Aligned Recurrent Neural Network Approach combines the potentiality of Dynamic Time warping and Recurrent Neural Network to capture temporal time sequences and align them to extract more meaning features and train the network to distinguish between genuine and forgery signatures. A verification System typically takes two signatures (Enrolled and Test) and checks the similarity between them to verify the authenticity. More on this verification system will be discussed in later sections.

Our work here contributes to i) Perform the experiment with authors proposed TA-RNN deep learning approach on MOBISIG database. ii) Check how efficient this TA-RNN is and verify will it be able to outperform previous state of art approaches on this MOBISIG database. iii) Evaulating and interpreting the results obtained.

Remaining part of the paper is as, Section II describes Related research work on with deep learning approaches, Section III details about the Approach that's being used, Section IV gives description about Data set, Section V explains the experimental protocol and interpret the results obtained. Finally we conclude and give some insights about Future Scope in Section VI.

2. Related-Research Work

Many researches were made on verifying online handwritten signatures, First research study was on [2] "Exploring Recurrent Neural Network". This study's objective is to look at the possibilities of Recurrent Neural Networks (RNNs) for online handwritten signature biometrics, with a focus on improving the efficiency and precision of signature verification systems. The authors goal is to find solutions to issues including intra-class variances, attempted forgeries, and incomplete enrollment data. The recommended method effectively represents the temporal dynamics of online signature data using RNNs, more specifically Long Short-Term Memory (LSTM) networks. These networks are especially well suited for activities requiring signature verification because they have the capacity to capture the fundamental sequential information present in the signatures. The use of data augmentation techniques also enhances the system's generalization capabilities.

Lai and Jin experiment [3] on mobisig database using RAN's (Recurrent Adaptation Network) was to improve the robustness against forgeries, intra-class variation. Here they combined the Gated Auto Regressive Units(GARU) with Length-Normalized Path signature (LNPS) to extract more robust features. They also included MCYT-100 and eBiosign to test their approach. Finally, they were able to achieve 10.9% EER rate for skilled forgery which turns around to be remarkable among other finger online handwritten signature verification studies.

Hefny and Moustafa [4] considered legendre polynomial co-efficients as features and carried out research using basic approaches such as MLP (Multi layer Perceptron). By skillfully addressing forgery attempts and intra-class changes, the main goal is to improve the performance of signature verification systems. Their Approach achieved 0.5% EER on Sigcomp2011 Dutch dataset.

A stroke based LSTM approach proposed by Li [5] makes use of stroke-level data to accurately and reliably capture the temporal and spatial properties of signatures. Their Approach seemed to work but due to poor generalisation capacity of network the results achieved were more than 10% EER which failed to outperform Lai and Jin's work.

3. Proposed Approach – (TA – RNN)

First step in this approach would be taking the signatures we want to verify, as input and extracting all their temporal dynamics from the file. Next the process is divided into two stages as shown in figure 1.

3.1. Time Functions Alignment

Aligning time functions is necessary step here because, when you want to verify two signatures you are actually comparing the similarity between them. To compare the similarity between the time sequences they must be aligned properly to calculate the similarity score. For this, we are using DTW which is also known as Dynamic Time Warping [7]. The main aim of DTW is to optimally wrap two time series between their time axes and compute the similarity [8].

In this context DTW aligns two signatures time sequences to allow corresponding stokes, points to be compared and evaluated. DTW recognises changes in execution speed, timing, and duration that may be brought about by differences in writing styles, personality attributes, or other factors by matching the time functions.By establishing correspondences between similar components of real and fake signatures, the alignment makes it easier for one to contrast two things precisely and determine how similar they are.

In this Approach DTW is used to convert the real time functions to aligned time functions and send them to RNN as input. Doing so, RNN could extract more meaningful features all time sequences are optimally warped.

3.2. Recurrent Neural Network

This architecture is similar to that proposed in [2]. The aligned time functions are now sent to this Recurrent Neural Networks where the primary goal is learn a dissimilarity metric from data by minimising the cost function that makes the metric small for same pair of genuine signatures which will be labelled as 0 and large for different pair of subjects which will be labelled as 1.The proposed RNN architecture consists of three layers.

First layer, consists of two parallel BGRU Hidden layers of 46 memory blocks each where the weights are shared between them. The reason for 2 parallel BGRU hidden layers is to consider the time sequences in for present and future context which makes the model to understand the upcoming time sequences and predict accordingly. Also the reason for weights being shared is that when we are working with long time sequences there is chance for overfitting. To avoid that weights are being shared. The output of this layer concatenated and results in a hidden sequence.

Second layer, consists of a BGRU hidden layer of 23 memory blocks takes in that hidden sequence and results in fixed dimensional representation. This is sent to Feed forward Neural Network layer.

The final layer, Feed-forward NN layer with a sigmoid activation function classifies with output of 0 or 1 telling either the test signature is genuine or fake.

4. MOBISIG Database Description

MOBISIG dataset has signatures of 83 users. In that 43 are men and 39 are women. Each user has a dedicated folder in the dataset. Each user has 65 files in which 45 are genuine and 20 are forgery signature files. The signature are recorded in the 3 sessions. In each session 15 genuine signatures were recorded and in the last 2 sessions 10 forgery signatures were recorded. So on a whole there are 45 x 83 = 3735 genuine signatures and 20 x 83 = 1660 forgery

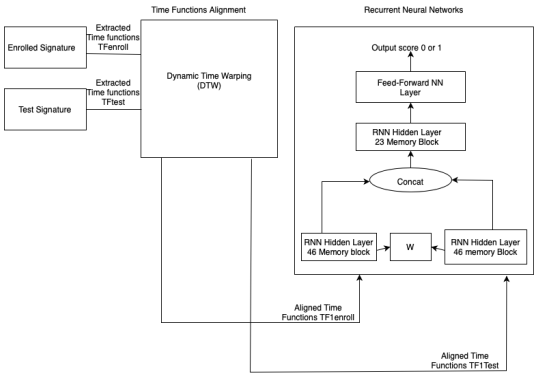

Figure 1. TA-RNN Architecture

signatures which sums up to a total of 5395 signatures in this dataset. Here data was collected using Nexus 9 Tablet. Each Signature file consists of these discrete values shown in figure 2.

Number	Features
1-2	X,Y-coordinate
3	Pen- pressure
4	Time stamp
5	Finger area
6-7	Velocities-X,Y
8-10	Acceleration-X,Y,Z
11-13	Gyroscope-X,Y,Z

Figure 2. Mobisig dataset features

5. Experimental Results

In this section we will discuss about the protocol we used to perform the experiment. In Later sections, we will evaluate our results by comparing them with previous state of art approaches.

5.1. Experimental Protocol

Firstly, we have divided the mobisig database into two seperate datasets. One for development and other for evaluation. Mobisig contains 83 users out of which 70 users are used for development and remaining 13 users for evaluation.

Secondly, to select data from user we have followed the aspects such as i)Number of training Signatures – First, we had chosen 4 signatures per user i.e., 4 genuine and 4 forgery per user. ii) Impostor Scenario – Here we considered only skilled forgeries because we want to show that our TA-RNN can improve previous state of art of 10.9% skilled on mobisig.

After, following the aspects mentioned above, while performing the experiments they are divided into training and validation on 0.3 test train split. Now, in each dataset there is genuine and forgery data so, what we are doing is we are training model by each genuine-genuine comparisons and each genuine-forgery comparisons.

Finally, with the unseen data of remaining users we test the model performance and do complete analysis of the above scenarios.

5.2. Results

Study	Classifiers	Database		Experimental Protocol			Performance in EER
		Name	#Users	#Train Users	Input	#Train Sig.	
Lai and Jin (2018)	GARU + DTW	Mobisig	83	70	Finger	5	Skilled = 10.9%
Li *et al.* (2019)	LSTM	Mobisig	83	70	Finger	1	Skilled = 16.1%
R. Tolsana	TA-RNN's	DeepSignDB	1526	1084	Finger	1	Skilled = 13.8%
						4	Skilled = 11.3%
Our Work	TA-RNN's	Mobisig	83	70	Finger	4	Skilled= 9.03%

Figure 3. Comparison of state of Art with our work

In this section we will be discussing the results we have obtained and intrepret them.

To begin with the experiment as discussed before, we have started by randomly taking 4 training samples per user with 70 X 4 = 280 genuine and forgery signatures. Next we started comparing each user genuine-genuine and genuine-forgery signatures. So per user we will be doing 16 genuine-genuine and 16 forgery-genuine comparisons and for 70 users summing upto 1120 comparisons in each cases while training the model

We have used Google colab to perform this experiment by using adam optimizer and Binary Cross Entropy loss (because its binary classification task) with 0.001 learning rate.

Analysing the results, as we can see in figure 3, First our intention was to start with lower training samples and then get an estimate on how model is performing on them. To that work as we can see in the table we got 9.03% EER which is less than state of art approach EER when they did it with GARU and DTW which had remarkable result of 10.9% in the case of finger scenarios.

So, Instead increasing the training samples we kept on repeating this with another set of random signatures per user to see if there is any drop in EER but then EER ranged in between 8-10 % everytime we did it.

Later, On analysing the graph plot of EER. This graph shows the relationship between FAR (false Acceptance Rate)

and FRR (false Rejection Rate). The point where both FAR and FRR meet define the EER. There is threshold on X-axis and Error Rate on Y-axis. This threshold is used to determine the class prediction based on the model's output or a similarity/distance measure. If we see in the graph as threshold increasing the false acceptance rate is decreasing and false rejection rate increasing.

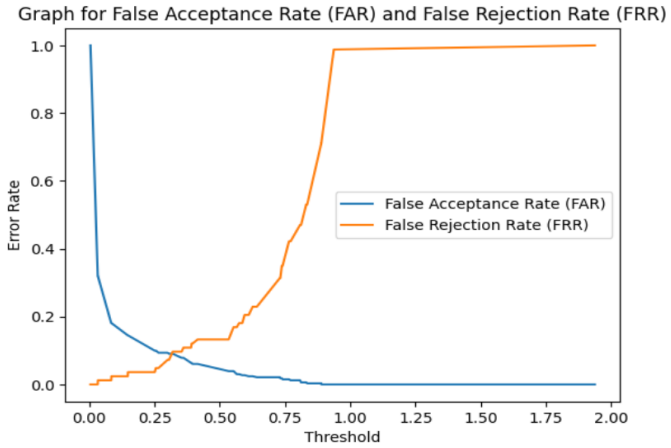

Figure 4. EER Graph plot

As we told, we are now testing model efficiency on the unseen 13 test users data. For that we have built a tkinter GUI which resembles our verification system. So here, from the test users we are selecting two inputs and verifying whether they are genuine or forgery as you can see in both figure 5 and 6. The model was efficient on almost 90% of test users data.

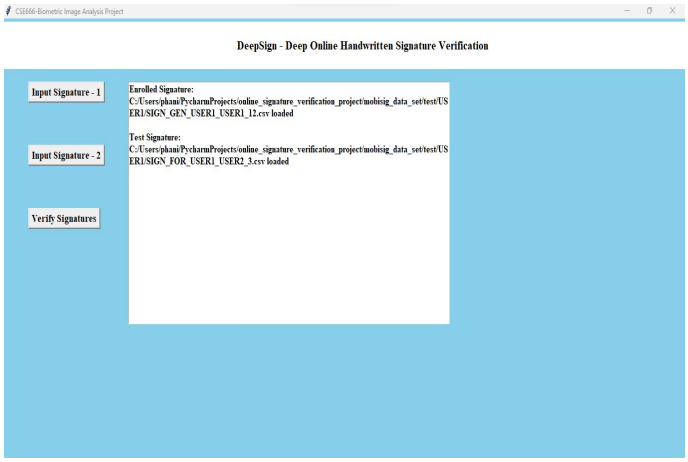

Figure 5. TestCaseScenario Comparing Genuine and Forgery

6. Conclusion

In our work, we have used TA-RNN approach which authors told an efficient algorithm for online signature verification using deep learning approach to test on mobisig dataset.

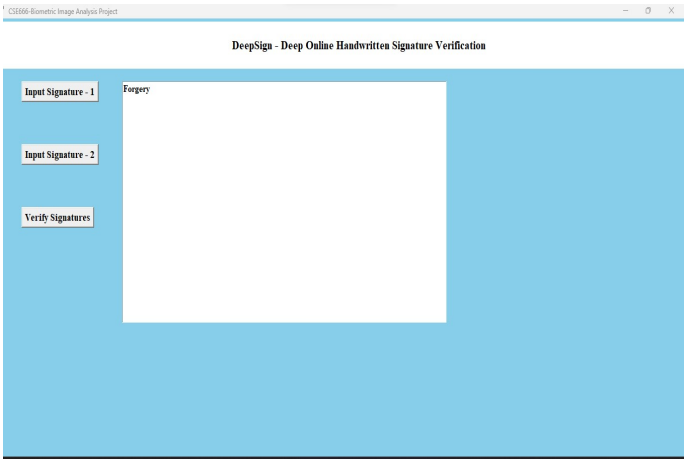

Figure 6. TestCaseScenario Output

Our primary goal was to train a model to make them robust against finger case scenario skilled forgeries. We were able to achieve the lower EER than the previous state of art approaches proving TA-RNN to be an efficient deep learning algorithm to perform signature verification tasks.

Also, Our TA-RNN model was able to predict whether the query signature is genuine or forgery accurately on the unseen data in the finger case scenarios.

7. Future Work

In future, we would like to use this pre-trained model on MOBISIG database to work on authors DeepSignDB and implement Transfer learning so that we can improve the results what authors received in case of finger scenarios.

Also, we can use this TA-RNN approach in other biometrics such as key-stroke,neuromotor related aspected tasks.

References

[1] R.Tolosana , R.Vera-Rodriguez , J. Fierrez , and J. Ortega-Garcia, "DeepSign: Deep On-Line Signature Verification",IEEE TRANSACTIONS ON BIOMETRICS, BEHAVIOR, AND IDENTITY SCIENCE, VOL. 3, NO. 2, APRIL 2021

[2] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Exploring recurrent neural networks for on-line handwritten signature biometrics," IEEE Access, vol. 6, pp. 5128–5138, 2018.

[3] S. Lai and L. Jin, "Recurrent adaptation networks for online signature verification," IEEE Trans. Inf. Forensics Security, vol. 14, pp. 1624–1637, 2019.

[4] A. Hefny and M. Moustafa, "Online signature verification using deep learning and feature representation using legendre polynomial coefficients," in Proc. Int. Conf. Adv. Mach. Learn. Technol. Appl., 2019, pp. 689–697.

[5] C. Li et al., "A stroke-based RNN for writer-independent online signature verification," in Proc. Int. Conf. Doc. Anal. Recognit. (ICDAR), 2019, pp. 526–532.

[6] R. Tolosana, R. Vera-Rodriguez, J. Fiérrez, A. Morales, and J. Ortega-Garcia, "Do you need more data? the DeepSignDB on-line handwritten signature biometric database," in Proc. Int. Conf. Doc. Anal. Recognit. (ICDAR), 2019, pp. 1143–1148.

[7] A. Fischer and R. Plamondon, "Signature verification based on the kinematic theory of rapid human movements," IEEE Trans. Human-Mach. Syst., vol. 47, no. 2, pp. 169–180, Apr. 2017.

[8] R. Blanco-Gonzalo, R. Sanchez-Reillo, O. Miguel-Hurtado, and J. Liu-Jimenez, "Performance evaluation of handwritten signature recognition in mobile environments," IET Biometr., vol. 3, no. 3, pp. 139–146, Sep. 2014.

[9] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Presentation attacks in signature biometrics: Types and introduc- tion to attack detection," in Handbook of Biometric Anti-Spoofing, 2nd ed., S. Marcel, M. S. Nixon, J. Fierrez, and N. Evans, Eds. Cham, Switzerland: Springer, 2019.

[10] J. Ortega-Garcia et al., "The multiscenario multienvironment biosecure multimodal database (BMDB)," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 6, pp. 1097–1111, Jun. 2010.

[11] R. Tolosana, P. Delgado-Santos, A. Perez-Uribe, R. Vera-Rodriguez, J. Fierrez, and A. Morales, "DeepWriteSYN: On-line handwriting synthesis via deep short-term representations," in Proc. 35th AAAI Conf. Artif. Intell., 2021.

[12] J. Galbally, J. Fierrez, M. Martinez-Diaz, and J. Ortega-Garcia, "Improving the enrollment in dynamic signature verfication with synthetic samples," in Proc. IAPR Int. Conf. Doc. Anal. Recognit. (ICDAR), 2009, pp. 1295–1299.

[13] A. Morales et al., "Keystroke biometrics in response to fake news propa- gation in a global pandemic," in Proc. IEEE Comput. Softw. Appl. Conf., 2020, pp. 1604–1609.