Verifying Digital Signature Via Deep Neural Network

Jiani Tao

University College London, London, Uk, China

Keywords: Psychological neural network, Digital signature, Verification

Abstract: Signature verification is becoming more and more popular in the recent years with growing realistic applications. It serves as an auxiliary personal identity verification method. However, how to establish an effective signature verification system is still an open problem. In this paper, we study utilizing deep learning methods to build such verification system by considering varying feature extractions and different model architectures, ranging from Recurrent Neural Network, Convolutional Neural Network to Transformer. We numerically demonstrate the effectiveness of the built system on benchmark Signature Verification Contest 2004 to get satisfactory evaluation results.

1. Introduction

Signature verification is receiving more and more attentions from various applications nowadays. It is widely used for the scenarios which requires efficient identity verification, such as personal authorization, credit card daily usage, various banking scenarios, contract agreement etc [1,2,3,17]. However, how to build an effective signature verification system is still an open problem to employ it into realistic applications.

Comparing with other reliable biomarkers, such as fingerprint and DNA, using signature to identity verification is somewhat more challenging because of the technical perspectives such as low interpersonal variance [4] and high intrapersonal variance [5] and the challenges from the society. Particularly, from the technical perspective, low interpersonal variance means that different people may have similar writing style, and high intrapersonal variance refers that one single person may sign very differently in various places and time periods. On the other hand, along with the increasing concern regarding users' personal data, since signature is related to the personal privacy to some extent, then consequently it increases the difficulty to establish a sufficient database to collect enough information of each person in order to achieve effective verification.

To establish an effective signature verification system, we utilize various modern deep learning technologies. Specifically, we explore to use Recurrent Neural Networks (RNNs) [6], Compact Convolutional Neural Network (CNNs) – Recurrent Neural Networks [7] with different featurization, and the popular transformer architecture [8]. The features that we used origins from the well-known Path Signature Feature for Handwritten recognition, but are enhanced to incorporate temporal information to amplify the interpersonal variance and eliminate the intrapersonal variance. Therefore, our established signature verification system can effectively overcome the above challenges.

Our main contributions can be summarized as follows:

- We explore various modern deep learning techniques for building effective signature verification system.
- We extend path signature features to incorporate it with temporal information.
- We numerically demonstrate the effectiveness of verification system.

2. Related Work

DTW (dynamic time warping) is another relevant method for feature extraction. By align vectors into same lengths can compare and compute the distance between signatures of different lengths. After normalization, the accuracy of recognition will increase. However, compared with PSF, DTW

  DOI: 10.25236/iccem.2021.062

only considers 1st darn FPS-the difficult distance between two sequential points in X, Y coordinates-which describes the local distribution of features. Since we are solving SVC2004 problem[9] which only obtained a small amount data, increasing the number of features as much as possible by intervention is the object to ensure the verification accuracy. Therefore, the implementation of DWT does not seem to be the best fit, so we use PSF to extract feature.

3. Method

3.1 Rnn.

As diagrammed in Figure 1, we establish a Recurrent Neural Network architecture for signature verification. In particular, given an input ink data, we at first transform it into a sequence of feature vectors, then feed this sequence of feature vectors into a RNN model to generate a sequence of probability distribution prediction on each time stamp, wherein the prediction on the last time stamp is used for determining the predicted label of the current input ink data.

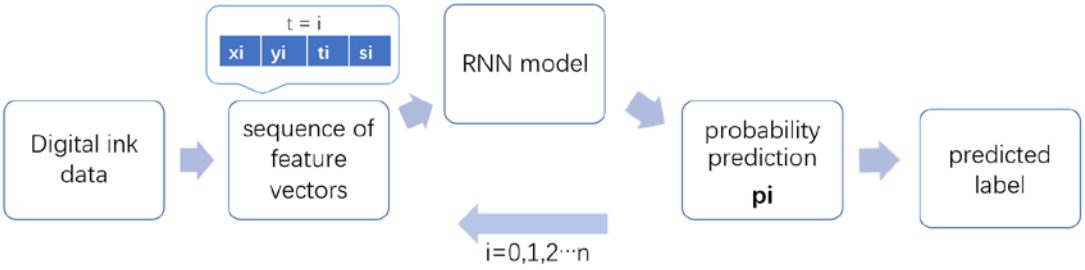

Fig.1 Architecture of Rnns in Signature Verification System

In particular, to train such RNN model. The input features at each timestamp include x-axis, y-axis, timestamps and penups status. By extracting the features from data paths, transforming them into 2D matrix and labelling each piece of data with zero (forged signature) or one (valid signature), we created a SVD dataset containing all the above such features and the corresponding label, and put the dataset into a data loader. Additionally, as the number of stroke points are relatively large, in order to avoid the information exploding issue, Long-Short Term Memory (LSTMs) [10] is the ultimate RNN variant in our consideration.

3.2 Compact Cnn+ Rnn

As diagrammed in Figure 2, we then establish a compact CNN-RNN architecture for signature verification. Given an input digital ink data, we firstly transformed it into a path signature feature [11] and then construct a 3D tensor with fixed height but the varying widths. PSF features encodes the geometrical spatial information of the stroke data via varying orders.

In particular, its zero-th order encodes the existence of the stroke point, the first-order describes the spatial transition between consecutive points, and the second order represents the curvature information. To be more specific, the zero ordered feature describes the location of points in the strokes via binary values. The 1st ordered feature contains two matrixes. It describes the changed value of x-coordinate position and y-coordinate position between the current point and the previous point. The 2nd ordered feature contains four matrixes totally. It is computed by computing the Hadmard tensor product of the 1-st order features, which explicitly take the square and the cross product of x-axis and y-axis difference. We stack the computed the tensors as the whole PSF tensor with 7 channels, which is then fed into a subsequent Convolutional Neural Network to get a 3D tensor output.

Because of the varying width in the input tensor, this output 3D tensor also has varying widths, but with fixed height and fixed channel. We then reduced the dimension of the 3d tensor into a matrix by the averaging max pooling operator with varying number of columns but the fixed number of rows. This matrix is next utilized to be processed by a RNN model.

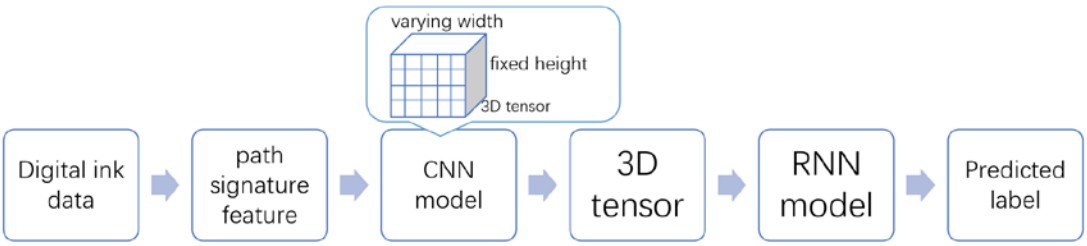

Fig.2 Architecture of Cnn-Rnns in Signature Verification System

To use path signature feature properly, we need to scale the axes of stroke points into the target height as 128. To ensure the smoothness of the generated features, processing uniform distance sampling to interpolate points is necessary. Additionally, we normalize the computed PSF tensor into an interval between -1 and 1.

However, the path signature feature only considers the spatial information, but lacks the consideration of temporal information which is important to distinguish different signatures as the forgery signature may requires longer time to be finished. Therefore, we further enhance the original path signature feature with additional temporal information and proposed temporal-enhanced path signature features. Specifically, besides the current seven matrices from spatial PSF, we concatenate one additional channels which are computed by averaging the relative timestamp information. Since the value of timestamp in the original dataset is too large, we reduce each timestamp by the first value of timestamp in a input digital link data and scale them by the maximum value of timestamp. We further transfer them into a matrix, which is the 8th channel in the tensor.

3.3 The Transformer

Transformer [12] is a popular model architecture in the recent years, which has been demonstrated as a very effective structure to reach the state-of-the-art results in various applications, ranging from natural language processing [13] to computer vision tasks [14]. The transformer consists of fully connected layers and multi-head attention layers with some residual connection. As the transformer fits sequential data naturally, we used the same CNN feature extractor as the above section, then feed the extracted features into the transformer architecture to do the binary classification.

4. Experimental Results

In this section, we numerically demonstrate the effectiveness of the above methods described in Section 4. We implement various deep learning algorithms via Pytorch and conduct all the experiments an Nvidia RTX 1080 GPU. The training and testing datasets are constructed by splitting the SVC2004 into 80% and 20% ration respectively. For each model backend, we train via the standard stochastic gradient descent method (SGD) till the convergence is reached. The overall numerical results are displayed in Table 1. The RNN structure we selected has two RNN layers, with hidden size as 256. All CNNs used in this paper is coming from the benchmark VGG16 [15] but removing the tail fully connected layers. The transformer architecture that we used is based on the transformer structure by huggingface [16] while replacing the foremost embedding layers.

Table 1 Experimental Result

	RNN	CNN+RNN	CNN+RNN Temporal Enhanced PSF	Transformer	Transformer Temporal Enhanced PSF
Precision	61.6%	66.8%	79.5%	71.2%	83.1%
Recall	65.4%	67.9%	82.3%	73.5%	88.4%
Accuracy	63.3%	67.1%	80.0%	72.9%	86.7%
F1 score	63.4%	67.3%	80.8%	72.3%	85.7%

As shown in Table 1, the RNN on the original features (the x-axis, y-axis, time-stamp and pen

up/down) is no doubt the worst method as it only reaches F1 score as 63.4%, which implies these features are either not sufficient or RNN lacks the capability to extract enough information to achieve effective signature verification. Compared to RNN, the compact CNN and RNN with path signature features performs relatively better, which is a strong evidence that the path signature features indeed provide more insightful information, but the metric is still not good enough. One reason might be that the path signature feature only considers the spatial information, while for distinguishing the genuine and forgery, the temporal information is crucial as well. Therefore, after incorporating the temporal enhanced information, the results are improved significantly from F1 score as 67.3% to 80.8%. Transformer is a new but has been demonstrated as a remarkable successful algorithm. It only equipped with PSF can achieve better result than the compact CNN+RNN. Similarly, after adding temporal enhanced information, the transformer boosts the results to a higher level with the highest F1 score as 85.7%.

5. Conclusion

We studied utilizing the popular deep neural network to establish system to verify personal signatures. The included DNN architectures include RNN to transformer. The built system can reach accuracy about 87% on benchmark dataset which demonstrate the validness of the our method.

References

[1] Rachel Sugar ,"Why are we still signing credit card receipts",Voxmedia, Dec 5, 2018,from https://www.vox.com/the-goods/2018/12/5/18092092/credit-card-signatures-receipt-explained

[2] Muhammad Imran Malik, Sheraz Ahmed, Andreas Dengel and Marcus Liwicki. 2012. A Signature Verification Framework for Digital Pen Applications. IEEE.DOI:10.1109/ DAS.2012.10

[3] Emir Husni, Bramanto Leksono, Muhammad Ridho Rosa. 2015. Digital signature for contract signing in service commerce. IEEE. DOI: 10.1109/TIME-E.2015.7389757, 25 January 2016

[4] Marc A.Fournier, D.S.Moskowitz, David C.Zuroff. 2009, The interpersonal signature, Journal of Research in Personality, Volume 43, Issue 2, April 2009, Pages 155-162

[5] Maged M.M.Fahmy, 2010, Online handwritten signature verification system based on DWT features extraction and neural network classification, Ain Shams Engineering JournalVolume 1, Issue 1, September 2010, Pages 59-70

[6] "Recurrent neural network.", Wikipedia, Wikimedia Foundation, 3 July 2021, https://en.wikipedia.org/wiki/Recurrent_neural_network

[7] "Convolutional neural network", Wikipedia, Wikimedia Foundation, 8 July 2021, https://en.wikipedia.org/wiki/Convolutional_neural_network

[8] "Transformer (machine learning model)", Wikipedia, Wikimedia Foundation, 12 July 2021, https://en.wikipedia.org/wiki/Transformer_(machine_learning_model)

[9] [Dataset creator's name]. ([2003; April]). [Database for Task1], [Version of the dataset]. Retrieved [October 2003] from [https://www.cse.ust.hk/svc2004/download.html]

[10] "Long short-term memory", Wikipedia, Wikimedia Foundation, 12 July 2021, https://en.wikipedia.org/wiki/Long_short-term_memory

[11] Songxuan Lai, Lianwen Jin, Weixin Yang. 2017. Online Signature Verification Using Recurrent Neural Network and Length-Normalized Path Signature Descriptor. IEEE. DOI: 10.1109/ICDAR.2017.73

[12] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. "Attention is all you need." In Advances in neural information

processing systems, pp. 5998-6008. 2017.

[13] Tetko, Igor V., Pavel Karpov, Ruud Van Deursen, and Guillaume Godin. “State-of-the-art augmented NLP transformer models for direct and single-step retrosynthesis.” Nature communications 11, no. 1 (2020): 1-11.

[14] Arnab, Anurag, Mostafa Dehghani, Georg Heigold, Chen Sun, Mario Lučić, and Cordelia Schmid. “Vivit: A video vision transformer.” arXiv preprint arXiv:2103.15691 (2021).

[15] Chen, Tianyi, Bo Ji, Tianyu Ding, Biyi Fang, Guanyi Wang, Zhihui Zhu, Luming Liang, Yixin Shi, Sheng Yi, and Xiao Tu. “Only Train Once: A One-Shot Neural Network Training And Pruning Framework.” arXiv preprint arXiv:2107.07467 (2021).

[16] Wolf, Thomas, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac et al. “Huggingface's transformers: State-of-the-art natural language processing.” arXiv preprint arXiv:1910.03771 (2019).

[17] Ji, Bo, and Tianyi Chen. “Generative adversarial network for handwritten text.” arXiv preprint arXiv:1907.11845 (2019).