IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES
ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

A Siamese Network based Writer Independent Offline Signature Verification

DOI:10.48047/IJFANS/V11/I12/204

P. R. Krishna Prasad[1], Associate Professor, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.

Kurri Renuka Reddy[2], **Maddukuri Harshitha Sai**[3], **Lakshmisetty Thirumala Gopi**[4], **Kanneti Vedakshari**[5]
[2,3,4,5] UG Students, Department of CSE,
Vasireddy Venkatadri Institute of Technology, Nambur, Guntur Dt., Andhra Pradesh.
[1]prkrishnaprasad@vvit.net, [2]renukareddy.kurri@gmail.com,
[3]harshithamaddukuri01@gmail.com, [4]gopilakshmisetty5752@gmail.com,
[5]kannetivedakshari@gmail.com.

Abstract

Signatures are frequently used for both personal authorization and verification. The signatures on numerous papers, including legal agreements and bank checks, must be verified. Offline signature verification is a type of authentication that examines the physical action of signing while measuring the characteristics of an user's handwriting. In the present paper, an offline signature verification method was developed utilising Convolutional Siamese Network of Deep Learning [4]. Convolutional Siamese networks, which are dual networks that shares parameters, can be developed to acquire a feature map. Thus, Convolutional Siamese network is employed to verify the person's signature to a input signature that is saved in the database. Research has been conducted with a dataset of signatures that includes 750 signatures from 30 different individuals. An Accuracy of 91.6% was attained by the proposed solution.

Keywords: Signature verification, Writer Independent, Siamese Network, Deep learning.

Introduction

From ancient times, signatures are being utilised to authenticate a variety of human-related items, which includes documents, bank checks and individuals. Signatures are among the most well-known and widely acknowledged biometric trademarks. Thus, offline signature verification is a crucial task, many numerous attempts were made to eliminate the unpredictability associated with the human validation methodology, making signature verification a significant study in area of the machine learning along with the pattern recognition. The two forms of signature verification are (1) online and (2) offline, based on the input format.

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

An digital writing pad and styles, which may primarily record a series of dimensions of the electric pen tip when signing, are required for the online signature capture process. The writing speed, pressure, etc., can also be retrieved by these devices in addition to the signature's writing coordinates, and this additional information is employed in the online verification process.

The offline signature, on the other hand, is typically recorded by a scanner or other imaging device, which essentially results in two dimensional signature images. Since signature verification has been an incredibly popular place for research for many years, both offline and online signature verification are given a lot of attention.

Offline signature verification could be approached in two ways: writer dependent and writer independent. The writer independent scenario is preferable to writer dependent techniques because a writer dependent system must be retrained with every new writer in order for it to work [7] . It costs a lot for a system that depends on people, like a bank, where new customers might open accounts every day. In contrast, a generic system is created in writer independent cases to simulate the difference between authentic and fake signatures. The available signers are divided into train and test sets when a signature verification system is being trained under a writer independent approach [5].

Signatures are paired as comparable (real, real) or dissimilar (real, forged) pairs for a specific signer. For the purpose of balancing the number of instances, a random selection process selects an equal number of similar and dissimilar pairs from all the tuples of a single signer. To create the training and test examples for the classifier, this method is applied to each user in the train and the test samples. In this regard, a Siamese network, which comprises of twin convolutional networks accepting two unique signature pictures coming from the tuples that are either similar or dissimilar, can be used to efficiently model a signature verification system.

The convolutional neural networks that make up the network are joined at the top by a cost function that calculate similarity score seen between the highest level of feature map on each region of the network and distributes the weights between these dual networks.

Literature Survey

1. The research paper "Offline Signature Verification on Real-World Documents" is able to solve the real-world writer independent offline signature verification problem. To accomplish this, VGG-16 and ResNet-50 network models were used. During training, these models achieved accuracies of 76.38% and 75%, respectively, when tested with signatures from the training set [3].

2. The research work titled "Offline Signature Verification Using Convolutional Neural Network" proposes a new deep learning method for computing features in handwritten signature verification. In this study, the fine-tuned CNN model was employed to extract features from the signature for writer-dependent approach, which led to an accuracy of 89% [2].
3. The paper "Off-line Signature Verification through Machine Learning" proposes the use of a neural network as a subnetwork in a Siamese network for signature verification. The Siamese network was tested on three datasets, namely CEDAR, GPDS Synthetic Signature, and BHSig260, for both writer-independent (WI) and writer-dependent (WD) verification. The results showed an accuracy of 89.3% for the CEDAR dataset using the WI approach [1].

From these, it is observed that Convolutional Siamese Network yields better accuracy and performance in Verifying the offline signature of Writer Independent case.

Problem Identification

Fingerprints, voices, irises, faces and handwritten signatures are five common biometric recognition methods used in a variety of practical applications such as computer vision, contract signature and attendance. Subsequently, due to time and necessity, various biometric features connected to the human body were used to enable authentication.

Verifying a person's identification through handwritten signatures is difficult since a forger can acquire a user's handwritten signature and purposefully tries to copy it.

Methodology

The proposed offline signature verification system uses Convolutional Siamese Network for classification and feature extraction. The CNN design consists of a number of layers, each of which conducts a straightforward computation beginning with the original image pixels and feeds the outcome to the following layer.

Initially, the dataset will be gathered; this is known as data collection. Later, some preparation operations such as data pre-processing and data augmentation will be carried out. The data set will next be divided into training and testing data sets. Following that, a Convolutional Siamese model will be used. The outputs of two networks are used to calculate which yields whether the signature is genuine or forged on comparing with threshold. As seen in the figure below.

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

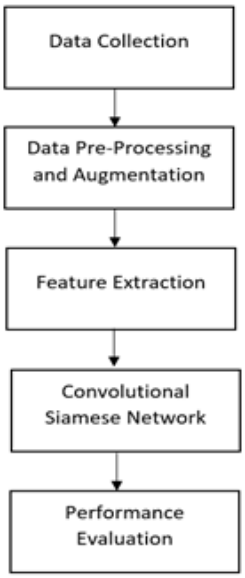

Fig:1 Methodology

Convolutional Siamese Network

Siamese Neural Networks (SNNs) are a group of the neural networks that contains two or more identical subnetworks. The sub-networks are configured identically, meaning that they share the same parameters and weights. During training, the sub-networks mirror each other in terms of parameter updating.

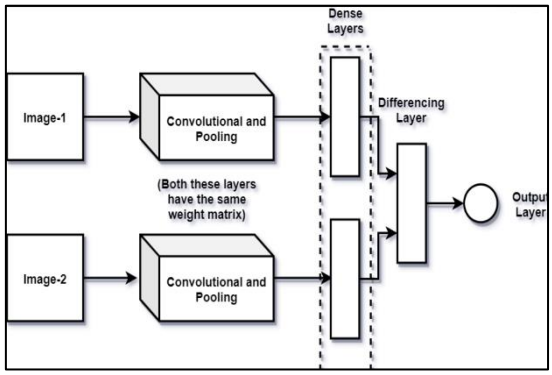

Fig:2 Convolutional Siamese Network

When a neural network is trained to predict different classes, adding or removing new classes can be challenging. This requires retraining the entire network using the entire dataset. However, SNNs are trained using a similarity function, which enables them to classify new types of data without needing to be retrained. This is because they learn to determine whether two photos are similar, rather than classifying them as belonging to a specific category. SNNs are particularly useful when dealing with deep neural networks that

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

require a large amount of data to be trained on. By using a similarity function, SNNs can be trained with less data and still be effective in classifying new types of data [6].

Implementation

A. Data Collection

In data collection phase, a dataset that contains 750 signatures of 30 individual users is employed.

For each user each of 5 real and forged signatures are collected.

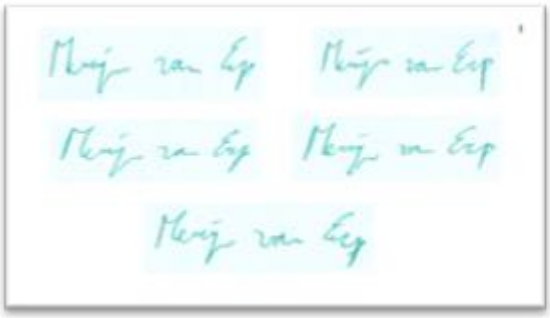

Fig:3 Five Real Signatures of a User

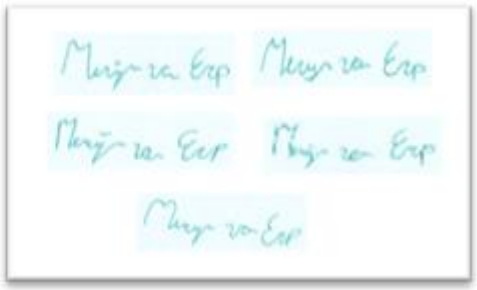

Fig:4 Five Forged Signatures of a User

B. Data Pre-Processing

All training and test photos are pre-processed for purpose of getting the signatures prepared for verification . The goal of this pre-processing is to perform Gray-scaling of images using OpenCV library of python and resizing the signatures to a specific size of 150 x 300 pixels.

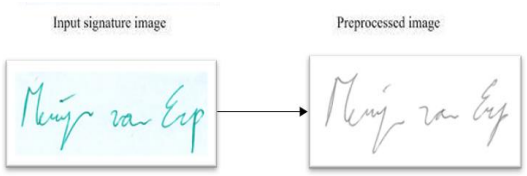

Fig:5 Signature after Pre-Processing

C. Data Augmentation

Data Augmentation technique yields pairs of real and forged signatures. The model is then trained using a few compositions of transformations, which results in the creation of an almost unlimited number of signature samples.

D. Feature Extraction

The main aim of feature extraction is to convert signature images into numerical values, which the Siamese network can use to learn the signature verification task. In order to train the CNN, a vast dataset of signature pairs labelled as either genuine or forged is necessary. The model is trained using the Adam optimizer for 30 epochs, with a batch size of 128 and image size of 150x300.

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

After pre-processing, the images are then passed through the CNN, which extracts a set of features from each image. These extracted features are then compared using a distance metric, such as the Euclidean distance, to determine the similarity between the two signatures. This process of comparing the extracted features plays a critical role in accurately verifying the signatures.

```
Model: "model"
__________________________________________________________________________________________________
 Layer (type)                   Output Shape         Param #     Connected to
==================================================================================================
 input_1 (InputLayer)           [(None, 300, 150, 1  0           []
                                )]

 input_2 (InputLayer)           [(None, 300, 150, 1  0           []
                                )]

 sequential (Sequential)        (None, 1, 128)       853184      ['input_1[0][0]',
                                                                  'input_2[0][0]']

 subtract (Subtract)            (None, 1, 128)       0           ['sequential[0][0]',
                                                                  'sequential[1][0]']

 dense_1 (Dense)                (None, 1, 1)         129         ['subtract[0][0]']

==================================================================================================
Total params: 853,313
Trainable params: 852,321
Non-trainable params: 992
```

E. Model Creation

For creation of model, “keras” library of python is used. Keras module of python consists of many pretrained models regarding to neural networks. Data is separated such that 75% of data is used for training and 25% of data is being used for testing the model. The below figure gives summary of created model.

F. Binary Cross Entropy Loss Function

A loss function applied in binary classification problems is binary cross-entropy. The fundamental goal of these problems is to answer a question with only two options, such as a genuine signature or a forgery.

The formula of this loss function can be given by:

$$H_p(q) = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i))$$

Here, y represents the label / class (1 for the forged signature and 0 for the real signature).

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

p(y) represents the predicted probability of the data point being real or forge for all N data points.

Euclidean distance

Euclidean distance, to determine the similarity between the signatures whose score is attained after loss function. Formula to calculate Euclidean distance is

$$sqrt(sum(r1[i] - r2[i])^2))$$

where r1,r2 are scores obtained from pair of signatures.

Results

The proposed system was created in Python. As indicated in the picture, a graphical user interface (GUI) is used to verify the signature using html, css and bootstrap.

Firstly, user must register by supplying credentials such as their name, ID, and at least one image of their signature. The user can submit up to 5 signatures, which can improve the results while verifying the signature. The user's details are stored and can be retrieved upon verification using the Python Django framework.

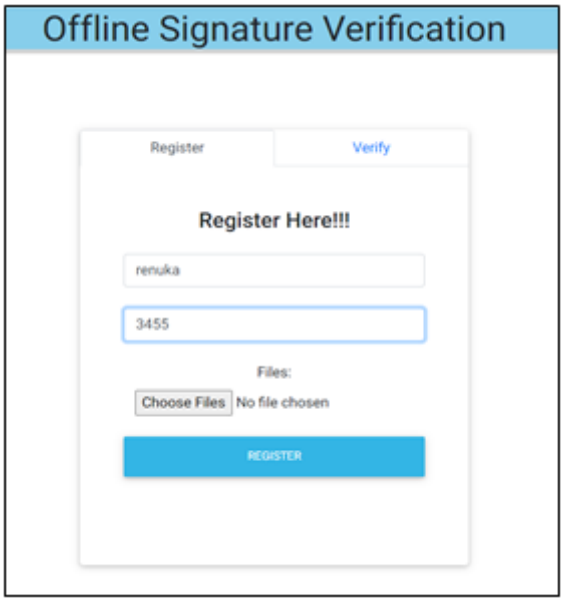

Fig:6 User registration Page

Following registration, for verification the user must provide the id in addition to the signature to be verified. This signature is compared towards the signature recorded in the database with respect to the id provided by the user.

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

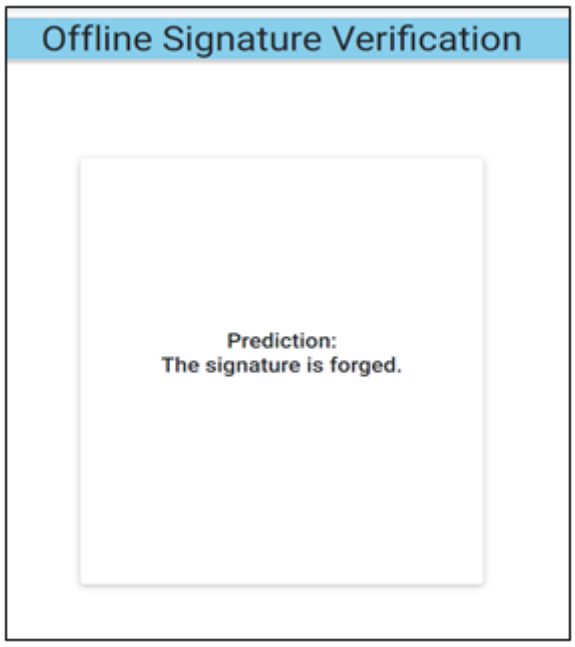

Fig:7 Output Page

The generated Siamese network model accepts these images as input and provides a similarity score, which is then compared to a threshold to yield the result.

A Siamese Neural Network-based model for offline signature verification has been developed. This model is designed to provide a final assessment of the authenticity of signatures. Our model performed best when trained using a blend of synthetic augmented and genuine signatures.

This trained model obtained an accuracy of **91.6%** on test samples and an accuracy of **99%** on trained images.

It is crucial to note that our method is writer-independent, which makes the improving of accuracy more challenging than in the writer-dependent scenario.

The **accuracy** of test images of dataset is calculated using the accuracy formula

False Acceptance Rate is calculated using formula

$$FAR(\%) = \left(\frac{False\ accepted}{All\ Tests}\right) * 100\ \%$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

False Rejection Rate is calculated using formula

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

$$FRR = \frac{Correct\ rejected}{All\ Tests} * 100$$

Metrics	Percentage
Accuracy	91.7
FAR	8.3
FRR	0

Table-I: Metrics along with Percentages calculated

Conclusion

The present paper presents a novel approach to offline signature verification using Siamese networks that utilizes writer-independent feature learning. Unlike its predecessors, this method does not depend on manually designed features. Rather, it extracts features from the data in a writer-independent manner. Our model has acquired an accuracy of 91.6%. Experimental results indicate that it can be applied to new signers without requiring further training.

Future Scope

One possible avenue for future work is to explore new approaches that can more effectively distinguish between genuine signatures and forgeries, while also decreasing the number of signatures utilised for the training of model. Another possible approach is to develop algorithms that are more robust to variations in handwriting styles, such as those caused by aging, injury, or intentional obfuscation. This can be achieved by collecting a diverse set of signature samples and training the model on a larger and more representative dataset.

Overall, there is a great deal of potential for future research and development in signature verification, and continued efforts in this area will be essential for ensuring the security and reliability of digital transactions and authentication systems.

References

[1] Rateria and S. Agarwal, "Off-line Signature Verification through Machine Learning," 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON), Gorakhpur, India, 2018, pp. 1-7, doi: 10.1109/UPCON.2018.8597090.

[2] S. V. Bonde, P. Narwade and R. Sawant, "Offline Signature Verification Using Convolutional Neural Network," 2020 6th International Conference on Signal Processing and Communication (ICSC), Noida, India, 2020, pp. 119-127, doi: 10.1109/ICSC48311.2020.9182727.

[3] D. Engin, A. Kantarcı, S. Arslan and H. K. Ekenel, "Offline Signature Verification on Real-World Documents," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 3518-3526, doi: 10.1109/CVPRW50498.2020.00412.

[4] M. Hanmandlu, A. B. Sronothara and S. Vasikarla, "Deep Learning based Offline Signature Verification," 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, USA, 2018, pp. 732-737, doi: 10.1109/UEMCON.2018.8796678.

[5] L. G. Hafemann, R. Sabourin and L. S. Oliveira, "Writer-independent feature learning for Offline Signature Verification using Deep Convolutional Neural Networks," 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 2016, pp. 2576-2583, doi: 10.1109/IJCNN.2016.7727521.

[6] M. Mutlu Yapici, A. Tekerek and N. Topaloglu, "Convolutional Neural Network Based Offline Signature Verification Application," 2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT), Ankara, Turkey, 2018, pp. 30-34, doi: 10.1109/IBIGDELFT.2018.8625290.

[7] A. Kumar and R. Rastogi, "Writer-Independent Offline Signature Verification using LBP and NN," 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), Faridabad, India, 2022, pp. 776-779, doi: 10.1109/COM-IT-CON54601.2022.9850831.