Signature Verification Using CNN with Information Retrieval

Utkarsh Shukla, Srishti Verma, Tushar Gwal, Atul Kumar Verma

Department of Computer Science and Engineering, Shri Ramswaroop Memorial College of Engineering and Management, Lucknow, Uttar Pradesh, India

ABSTRACT

When it comes to information security, biometric systems play a significant role in it. Signature verification is a popular research area in field of pattern recognition and image processing. It is a technique used by banks, intelligence agencies and high-profile institutions to validate the identity of an individual by comparing signatures and checking for authenticity. In this paper, the approach for the verification of signatures is based on Conventional Neural Network (CNN). This method saves time and energy and also helps to prevent human error during the signature process and lowers chances of fraud in the process of authentication. We achieved test accuracy of 89% and validation accuracy of 93%.

Keywords : Signature Verification, Pattern Recognition, Image Processing, CNN

I. INTRODUCTION

Since we all know that still offline signature is the main means of authentication in the government and as well as private organization [4]. Signature recognition and verification involves two tasks: one of them is identification of the signature owner, and the other is the decision about whether the signature is genuine or forged [5].

There exist two types of signature verification: online and offline. Online verification requires an electronic signing system which provides data such as the pen's position, azimuth/altitude angle, and pressure at each time-step of the signing. While offline verification uses solely 2D visual (pixel) data acquired from scanning signed documents. While online systems provide more information to verify identity, they are less versatile and can only be used in certain contexts (e.g. transaction authorization) because they require specific input systems [1].

When any person does his/her signature, then instantly the signature is fed into the signature verification system and then image is compared with one on one signature file. [3] Signature verification is essential in preventing falsification of documents in numerous financial, legal, and other commercial settings. Fig. 1 shows the acquisition that creates a digitally encoded representation of the visual characteristics of the image. Since the signatures to be processed by the system should be in the digital image format. After that we have to normalize the signature, resize it to proper dimensions, remove the background noise, and thin the signature. [7].

There various problems faced by biometric systems regarding frauds, one of them is forgery.

Forgery is also known as fake signature, it can be classified as follows: random forgeries, produced without knowing either the name of the signer nor the shape of its signature; simple forgeries, produced knowing the name of the signer but without having an example of his signature; and skilled forgeries, produced by people who, after studying an original instance of the signature, attempt to imitate it as closely as possible. Since offline signature is one of the most used authentication tool, hence any type of forgery in it may result in a huge loss. Many times human tester also make mistakes, hence machine helps us in recognizing the forgeries at every level.

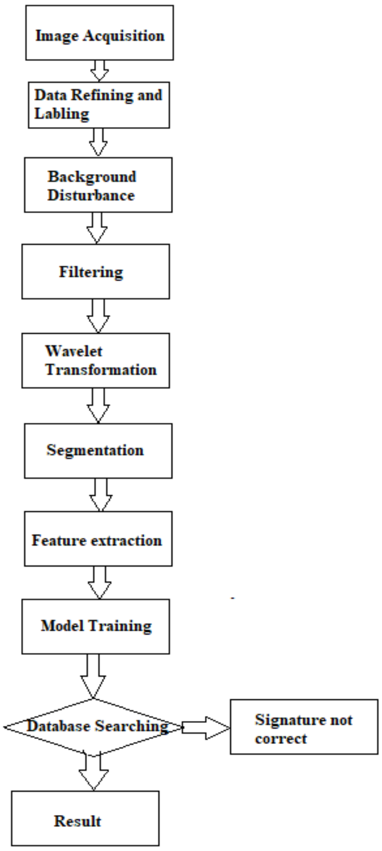


Figure 1. Flowchart of signature verification

II. LITERATURE REVIEW

In 1977, first algorithm regarding offline/online signature was published. Much research has followed, attempting various methods for both feature extraction and matching. [2]

"Offline Signature Verification with Convolutional Neural Networks" this paper aims to build offline signature verification system using CNN based on the dataset from the International Conference on Document Analysis and Recognition (ICDAR). [1]

"Signature Recognition and Verification with ANN" this paper presents an off-line signature recognition and verification system which is based on moment invariant method and ANN. Two different types of neural networks are developer with distinct task, one of signature recognition and the latter one for verification. The system achieved a success rate of 100% for 30 trained signatures and failed for untrained signatures. System must be scalable because if we want more accuracy in recognition and verification, we must add extra feature. [5]

"Offline signature recognition using neural networks approach" presents Offline Verification method of signatures using a set of simple shape based geometric features. Three features that matters the most area, centre of gravity, eccentricity, kurtosis and skewness. [8]

"Offline signature recognition system using histogram of oriented gradients" in this paper, an offline signature recognition system which uses histogram of oriented gradients is presented. Feedforward backpropagation neural network is used for classification. The system gives recognition rate of 96.87% with 4 training sample per individual. [10]

"Hindi and English Off-line Signature Identification and Verification" in this paper multi-script (Hindi and English) signature verification is done. The system will classify the signature and tell whether it belongs to the English category or the Hindi one. This paper used SVM classifier that were employed for identification and verification purpose. [9]

III. SIGNATURE VERIFICATION

Since we all know signature verification is one of the most important process in banking field. Hence the process of signature verification starts with image acquisition and ends at database searching of the user.

A. Dataset refining and Dataset labelling

We have chosen TC11 dataset which contained 1600 English signatures, this was used for the profile verification and data rendering of the person. There were 20 sets being used for the experiment and we made classes accordingly. Each set contained 25 signatures, it means total 500 signatures were considered. Through Google open refined tool we refined the data and labelled it.

B. Filtering

We basically filtered the data using mean filtering and spatial filtering. Here, we grey scaled the image up to 6 steps and obtained the input image. Then we started to segment the image into 15*5 blocks, after it we used the convolutional layer for the filtering of the image. It removed the background disturbance and noise of the image. Through convolutional layer we also enhanced the image through edge detection filter. Here we used 4*4 filters for more accurate values. In Fig. 2 we can see the filtered image, comprising of the below formula.

$F_R (x', y') = \int Cp (x', y', \sigma_1) VR (\sigma_1) d\sigma_1$

$F_G (x', y') = \int Cp (x', y', \sigma_2) VR (\sigma_2) d\sigma_2$

$F_B (x', y') = \int Cp (x', y', \sigma_3) VR (\sigma_3) d\sigma_3$

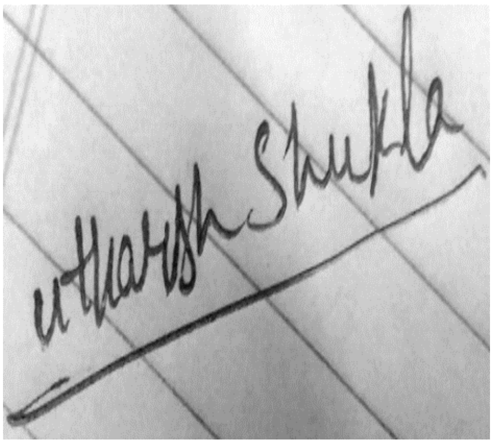


Figure 2. Signature image after filtering

$\theta (x , y) = \tan^{-1} \sigma v / \sigma u$ Where $\sigma u = g(x+1, y+1) - g(x, y)$

$\Delta v = g(x+1, y) - g(x, y+1)$ and $g(x, y)$ is the gray level of (x, y) point.

Hence through the edge Detection and spatial filter a value of 10 * 4 * 16 = 640 dimensional feature vector is obtained. Now, taking this vector we basically start our wavelet transformation.

C. Wavelet Transformation

Image comprises of 2 dimensions that are (x,y), when we grey scale the image we basically change the intensity of the image that is co-ordinates of the image. This transformation is needed for the less memory being required by the database to store the signature of the image and searching of the database could be fast. We have done wavelet based compression. We have used MATLAB for the wavelet transformation.

Here, we are using 4 x 4 matrix for the wavelet transformation. When we start the recognition part in the wavelet transformation we use $e_l^{\cdot d2i/\sigma}$. Following Fig. 3 and Fig. 4 shows the transformation according to the wavelet's value.

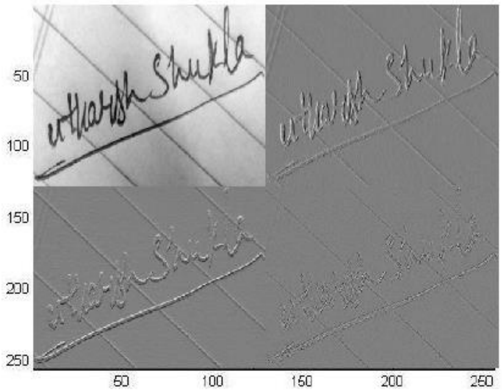


Figure 3. Image of signature after 1 wavelet transformation

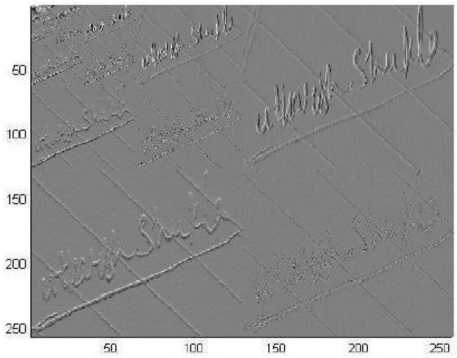


Figure 4. Image of signature after 4 wavelet transformation

Here as you can see that we have used 4 x 4 matrix, this we had done with the help of time adjacency matrix. This we had achieved by temporal groups.

Hence from the above two images it is pretty much clear that how much memory and loading time we can save through compression. Now, this quantization leads us to the important part and that is segmentation.

D. Segmentation

In segmentation part we started dividing the signature images accordingly with edge detection algorithm. This algorithm is used for edge detection in the images and with the help of canny method we divided the image into two threshold. These threshold detected the strong and the week part of the image. This would result in detection of forgeries. It comprised of the following pseudo code-:

```
I = imread('signature.jpeg');
imshow(I)
BW1 = edge (I,'Canny');
```

Here we took input the grey scaled image and started the function. By canny method, we looked for the local maxima of the gradient. This was all done by the Gaussian Filter. Here, we used two thresholds to detect the wee and the strong edges. Threshold is the sensitivity and we took the following readings-: Sigma would be 5, Low threshold is 40% and High threshold is 50%. Hence following Fig. 5 and Fig. 6 shows the input and output of the segmentation-:

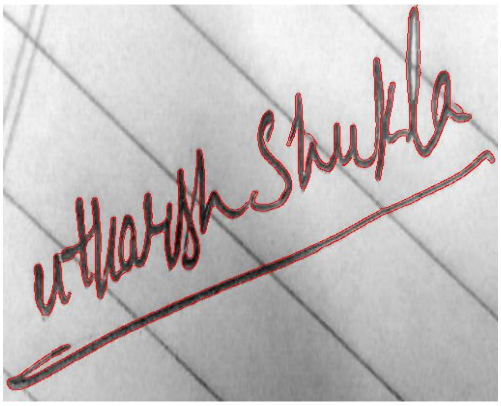


Figure 5. Input image to the segmentation.

From this input image we check all the results in edge detection algorithm and then we get the most 80% accuracy with canny method.

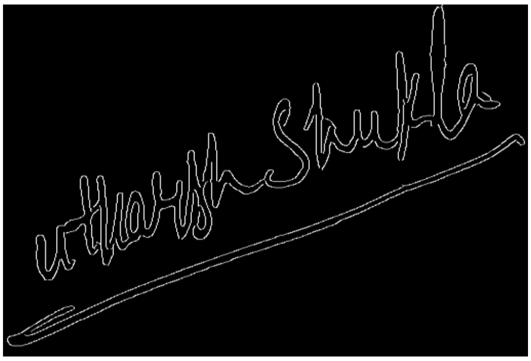


Figure 6. Output with the canny method.

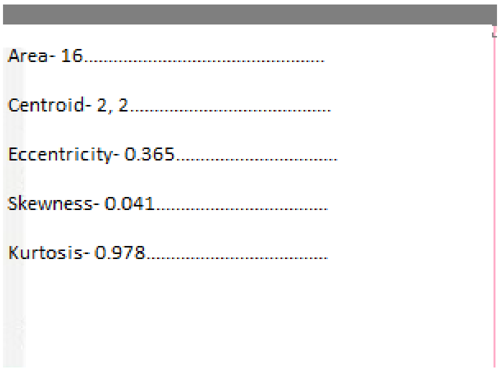
Now, with this image we started feature extraction.

E. Feature Extraction

Since feature extraction is one of the most important method for forgery detection, database rendering result and signature recognition. Hence, if any type of noise or shift to the factors like eccentricity, skewness etc. could cause temporal shift. Here, in the results we got-:

Area of the image 16 pixels, centroid 2, 2, eccentricity 0.365, skewness is 0.041 and kurtosis is 0.978. We basically extracted the wavelets, shadow, texture of the signature and graphometric and geometric of the signature. These results can be seen in the Fig. 7.

From feature extraction done, we finally started training the model of our experiment through CNN.



Area- 16..

Centroid- 2, 2..

Eccentricity- 0.365..................................

Skewness- 0.041....................................

Kurtosis- 0.978......................................

Figure 7. Feature extraction

F. Model Training

We had chosen CNN for our model training, since in the past years CNN has been a good and reliable source for image processing or machine learning experiments. Going through CNN, CNN consists of the layers. Here, each layer plays a significant role. Since, when data is provided to the first layer and when it reaches a conclusion or finds a result, then it passes its result to the next layer. And that result is the input to the next layer.

Now, coming on the training part. We had chosen regression loss function, it is one of the most common methods for the loss between predicted quantity and true answer.

$L_i=\|f-y_i\|_1=\sum_j|f_j-(y_i)_j|$

Though this loss gradient function we updated all the parameters in the network using the Nestrov Momentum update. This we had chosen because it was the most accurate method and also keeps an eye on the changing of parameters in the network. Now, since we had a huge network we also provided a huge dataset of images to train our CNN, hence we used online handwriting database, since it was huge one, source is-https://www.gavo.t.u-tokyo.ac.jp/~qiao/database.html.

Through this step we had got 3 patched output, this gave us that our trained model gives us 87% accuracy (somewhat less than expected).

We had taken 420 signatures for training and testing of the model. For the better forgery detection we took 4 different signatures from the users, some of them also gave 6 signatures (15). For each person 1 forgery signature of each type is signed. In the training set the total number of signatures is 190 = (original) = (25x4 + 15x6). For testing 22 forgeries of each type were taken and to test original signatures we took total 13 signatures.

Following is the design of the CNN-:

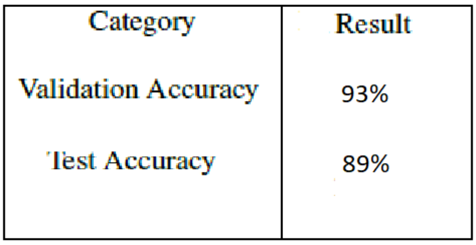
Network layers-7
Neurons in layer-8
Learning rate-0.25
Initial weight-1

G. Database Rendering

In database rendering we do the process of pattern matching of the signature. Through histogram based unique code we get the fast database matching and information rendering. Here, in the database of the images we make primary key as the signature, which is also the recognition key for the user information. We start connecting this with our database, since project was developed in MATLAB hence it is one of the toughest part. We hadn't achieved the database storage and information rendering part, but still we know the pseudo codes to how to manage the database and safely store the information with the help of histogram graphs. Database storing the information could be one of the most important part for the fraud/forgery detection.

IV. RESULT

The data base of 400 signatures was tested. The accuracy and precision that we achieved is 89% and we expected it of 93%, as shown below in the Fig. 8.



Category	Result
Validation Accuracy	93%
Test Accuracy	89%

Figure 8. Result

There was some error as the rejection ratio was also detected, which effected the accuracy of the result. We may achieve the better results by extra factors and minimizing the errors. We have also taken a good leap and are working on the proof of concept of data base searching. Accuracy of histogram searching is still 0.43, which we are working on.

V. CONCLUSION

Since, Convolutional Neural Networks are used in each almost every part of image processing, hence this could be the effective part for the bodies like banks, government offices. Since, the offline signature verification and retrieval of the information is no doubt one of the biggest task now a days, hence CNN could play an important role in it.

Since, CNN could be effectively be used in the more powerful and the next generation computers. And this above discussed method could be one of the best feature for all the bodies, working in the offline signature field. Since, we only achieved the accuracy of 89% only, but with minimization of the errors and increased CNN layers we could achieve an accuracy of above 95%. This method could also save the bank frauds and forgeries. After the use of this method as a whole pack of software model could get well trained and accuracy may increase up to 98%.

VI. ACKNOWLEDGEMENT

I would like to pay special thankfulness and warmth to my Project Guide, Dr. Atul Kumar Verma, who gave me this opportunity to do the research paper. His vital support, assistance and encouragement made it possible to achieve the goal. Last I would like to thank various authors of reference materials mentioned in the reference section for their commendable research.

VII. REFERENCES

[1]. Gabe Alvarez, Blue Sheffer, Morgan Bryant, "Offline Signature Verification with Convolutional Neural Networks"

[2]. Manoj Kumar"Signature Verification Using Neural Network", IJCSE, 0975-3397, 09/09/2012

[3]. Vahab Iranmanesh, Sharifah Mumtaz Syed Ahmad, Wan Azizun Wan Adnan, Fahad Layth Malallah, Salman Yussof, "Online signature verification using neural network and person correlation features", 02-04/12/2013, IEEE conference on Open System, 14095766

[4]. Saba Mushtaq, A. H. Mir, "Signature verification: A study" International Conference on Computer and Communication Technology (ICCCT), 27 February 2014, DOI: 10.1109/ICCCT.2013.6749637

[5]. Cemil OZ, Fikret Ercal, Zafer Demir, "Signature Recognition and Verification with ANN" Oct 16, 2014

[6]. Julita A., Fauziyah S., Azlina O., Mardiana B., Hazura H., Zahariah A.M. , "Online signature Verification System", 5th International Colloquium on Signal Processing & Its Applications (CSPA), 978-1-4244-4152-5/09 2009

[7]. V A Bharadi, H B Kekre, "Off-Line Signature Recognition Systems", International Journal of Computer Applications (0975 - 8887) Volume 1 – No. 27 2010

[8]. Ali Karounia , Bassam Daya, Samia Bahlak, "Offline signature recognition using neural networks approach" Procedia Computer Science Volume 3 155–161 doi:10.1016/j.procs.2010.12.027, 2010

[9]. Srikanta Pal, Umapada Pal, and Michael Blumenstein, "Hindi and English Off-line Signature Identification and Verification"

[10]. Pallavi Patil, Bryan Almeida, Niketa Chettiar, Joyal Babu, "Offline signature recognition system using histogram of oriented gradients" International Conference on Advances in Computing, Communication and Control (ICAC3), 10.1109/ICAC3.2017.8318766, 2017