SIGNATURE FORGERY DETECTION USING MACHINE LEARNING

Ms. Manjula Subramaniam*¹, Teja E*², N Arpith Mathew*³

*¹Guide, Department Of Computer Science & Engineering, Sir M Visvesvaraya Institute Of Technology, Bengaluru, Karnataka, India.
*²,³Student, Department Of Computer Science & Engineering, Sir M Visvesvaraya Institute Of Technology, Bengaluru, Karnataka, India.

ABSTRACT

In today's society, signatures are used in many important documents such as bank cheques, passports, driving licenses etc. and can be faked in multiple ways. This creates many problems such as false identifications, identity theft, hacking etc. To reduce this issue, our project is focused on developing a system for detecting whether a signature is real or fake from a dataset of signatures using CNN and Deep learning. The reason we are using CNN and Deep learning is because signature change over a period of time based on multiple behavioral changes such as age, state of mind, physical health etc. We require a system that can learn from multiple training datasets and increase its accuracy of detection. There are two types of signatures authentication methods, which are online signature and offline signature verification methods. Our project is based on offline signatures forgery detection methods. This type of signatures are handwritten on the documents and require an image of the signature. This is why we also should consider image processing for this project. We are referencing a few papers which implement the project using a few methods for both online and offline signature forgery detection methods based on deep learning models, we plan on implementing the offline methods and try to achieve a better accuracy.

Keywords: Signature, CNN, Forgery, Authentication, Deep Learning.

I. INTRODUCTION

A signature is a distinct form of a person’s name, which mostly tries to identify itself as a unique form representing the person. But signatures can be identical, since many people can have the same name. Hence it tries to act as a unique identifier to authenticate that person’s authenticity, usually in legal documents. It is a very important means of identification as it is used in many things such as legal documents, identification cards, cheque’s, etc. It can be considered to be a legal mark which represents a particular person. Since a signature is so widely used, there are many malicious actors trying to forge the signature for some personal gain; hence a very good signature forgery detection methods are necessary. Forgery is a legal term which is used when the purpose of is to defraud anyone. A forged signature is a signature than has been determined to not be genuine. Nowadays, a handwritten signature is widely accepted as a form of confirmation form a person in any transactions or legal documents which require authentication from the person. Since a signature acts as a confirmation in any binding legal or financial process, its authenticity has to be established and it needs to be verified. Signature forgery detection can be done mainly using online and offline methods. Online signatures are acquired using a touch screen device usually, and the trajectory and variations are recorded with time as the signature is being done and extracts features from the signature and performs several checks by comparing with a predefined database containing the signature. Hence, online signature forgery detection usually gives a very high rate of successful signature forgery detections. Offline Signature verification is an authentication method that usually uses the dynamics of a person’s handwritten signature to measure and analyze the physical activity of signing. It tries to detect forgeries by using feature extraction and checking for any discrepancies in the signature. It is less accurate than an online signature verification system.

II. LITERATURE SURVEY

a) HANDWRITTEN SIGNATURES FORGERY DETECTION

Kshitij Swapnil Jain, et al. (2021) defined the objective of the paper to verify if the signature is original or forged and to understand the characteristics of the signatures and implement a system to detect if the signature is forged. Although signature forgery can be detected by highly skilled experts, high accuracy cannot be achieved due to the many variations in handwriting styles and professionalism of forgers. Automatic
recognition system can be significantly more effective in verifying signatures with high accuracy and differentiate a genuine and forged signature. A CNN is used as a feature extractor and a classifier in the proposed method. The feature extractor extracts features from the input data via convolution filtering and down sampling. The authors assume that if a CNN is trained for classifying forged and genuine signatures, the trained CNN can extract effective features for distinguishing behavior characteristics of forgery, such as drawing the complicated sections of a signature with delays or hesitation. While deep networks can represent complex functions and learn features at different levels of abstraction, it will pose a problem when the gradient decreases exponentially and quickly reaches zero as the backpropagation continues from the final to the first layer. To overcome this, the authors use ResNet, which essentially skips certain steps or connections, which in turn allows the direct back propagation of the gradient and prevents the gradient from falling to a very small value quickly. The authors proposed method has successfully completed offline signature verification with increased efficiency and accuracy, as well as the ability to detect sophisticated forgeries. Signature forgery detection was done successfully using python and its libraries along with a solution based on Convolutional Neural Network (CNN) by the authors.

b) Digital signature Forgery Detection using CNN

Kiran, Lakkoju Chandra, et al. (2021) aim to present an analysis of the methods for verifying the integrity of visual media, i.e., the detection of manipulated images. Digital image forgery detection is commonly divided into active and passive techniques. The authors have studied active approaches to digital image forgery detection. Digital signature and watermarking are significant methods for digital image authenticity using active techniques. A signature consists of a combination of individual features and is the most significant means to verify the authenticity of the signature. It can work based on online or offline methods. Offline methods verify the identity of an individual by comparing the signature of a current query with reference signatures that have already been written. Online methods use dynamic knowledge related to the script and are more stable and have higher accuracy. Offline verification method verifies a signature after it has been produced. Online verification searches a database and compares it with previous samples to authenticate. The authors are implementing an online verification system and a dataset of 2000 images of forged and original images of signatures are divided and used using the RGB format. The RGB image is converted into a grayscale image and then into a binary image; then noise is removed and the image is resized to obtain the signature. After the preprocessing features are extracted, forming a .CSV file for each individual. A CNN is trained using the data with three input layers with different weights and biases and 3 hidden layers with a set of neurons and an output layer to show the final output (Genuine or Forged). They have also defined loss and optimizer to minimize the loss. And applied softmax to calculate the accuracy. And finally, after evaluating the model we will receive an output, as to whether the image is Genuine or Forged. The highest accuracy the authors obtained was 99.7%. The average accuracy is about 97.8%.

c) OFFLINE SIGNATURE FORGERY DETECTION USING CONVOLUTIONAL NEURAL NETWORK

Raj Balsekar, et al. (2020) proposed a system for signature recognition and verification based on CNN (Convolutional Neural Network). This paper has literature surveys on other papers for this system. The first author covers how signature verification requires help from both static and dynamic signatures. The second author covers the exploration of multiple explanations for offline signature verification. The third author covers how an individual signature can change multiple times over a period of time-based on behavioral changes. The fourth author covers automatic signature verification systems based on current advancements in image processing and machine learning. The fifth author covers how to extract recent geometric parameters from pre-processed signatures for comparing among different people. This paper describes the proposed system that can distinguish whether a signature is real or fake. First, this system collects handwritten signatures to extract unique features for creating a knowledgebase for all the signatures. These handwritten signatures are from a dataset containing 750 signatures from 150 individuals providing 5 signatures each. Second, these signatures are stored in the form of a matrix, then converted from the RGB image to a black and white image with the help of a grayscale algorithm, and finally undergoing geometric transformation. Third, these pre-processed signatures features are extracted and compared to the signatures features stored in the system. Finally, these signature features are compared with the real signature to determine whether the signature is real or fake, and
this process is done using CNN. In CNN, this signature undergoes multiple layers which are the convolutional layer, pooling layer and fully connected layer to get the results. We can conclude that this proposed system works very well.

d) Handwritten Signature Verification using Deep Learning

Alajrami, Eman, et al. (2019) proposed a system for signature verification and forgery detection using deep learning. This paper discusses two types of signature verification, namely offline (or) static and online (or) dynamic verification. The related techniques and methods used by this system in the paper are template matching and Hidden-Markov model for signature verification. Whereas for forgery detection, we use CNN(Convolutional Neural Network). The proposed methodology in this paper based on behavioral changes in handwritten signatures and not psychological characteristics. In this methodology, at first signatures and their certain amount of unique characteristics are collected and used for creating a knowledge base for this signatures. Then, they undergo preprocessing where this signatures are resized and distributed in two subdirectories. Finally, these images are loaded into CNN, which uses Keras with the TensorFlow backend and performance is evaluated. This paper also discusses CNN operations which are convolution, max pooling, fully-connected layer, ReLu(Rectifier) and softmax. The dataset used for this methodology is 300 images, of which 150 were real and forged, and are from 30 people with real signatures. The result of this methodology is a dataset with the split ratio of 8:2 that has the highest accuracy from 99.7 to 99.9%. We can conclude that this methodology has been implemented successfully.

e) Offline Signature Recognition and Forgery Detection using Deep Learning

Poddar, Jivesh, et al. (2020) proposed a system for signature recognition, forgery detection and verification based on CNN(Convolutional Neural Network), Crest-Through Method, SURF algorithm and Harris corner detection algorithm. In this system, CNN and Crest-Through Method are used for recognition and verification. Whereas, SURF and Harris corner detection algorithms are used forgery detection. This paper has literature surveys on other papers also. These signatures undergo pre-processing using CNN and Crest-Through Method. In CNN, the signatures go through noise removal, scaling, centralization and rotation. Whereas in Crest-Through Method, the signatures go through length to space ratio, width to space ratio and Crest-Through parameter. Both CNN and Crest-Through Method are also used for signature recognition. Then, this signatures go through forgery detection. In forgery detection, Harris corner detection algorithm is used detect and extract corner points of the signature for comparison with real signature. Whereas, SURF algorithm is used to detect and extract index points of the signature for comparison with real signature. The result of this system is that it has 94% and 85-89% accuracy for signature recognition and forgery detection, respectively. We can conclude that this system may have some flaws, but it is a useful one.

III. METHODOLOGY

We intend to use a Convolutional Neural Network (CNN) to implement the offline authentication methods, since all the offline methods exploit the content based features and the visual information of the signature, it is better to use a CNN since a CNN can classify the extracted features from the signature. We will use the CEDAR Signature dataset to train the neural network. The CEDAR Signature dataset is a signature verification database. 55 individuals contributed 24 signatures each and hence the dataset consists of 1320 genuine signatures. People were asked to forge the three other writers signatures, eight times every subject creating a total of 1320 forged signatures. The obtained dataset consists of 24 genuine and 24 forged signatures for each writer.
The data we are using has already been preprocessed, salt pepper noise removal and slant normalization were used on the data to remove some noise and straighten the slanted signatures, and we will further preprocess it to fit with our training model and make the signature verification easier to detect signature forgeries using a CNN. We will normalize the signature data while preprocessing, it will help in the feature extraction process. The data is used as training data and a part of it is used as test data. The next step to perform is feature extraction, it is a dimensionality reduction process, and it derives data using the signature images which can be used for processing and train the neural network. Feature Extraction aims to reduce the number of features in a dataset by creating new features from the existing ones. In feature extraction the rich features of the signature are extracted and converted into attributes for training the model. It will basically take the signature images and generates data like angles, curvature, and arc length, etc. of the signature images. This data is used to train our CNN model. We then train the CNN model, the data from feature extraction is used to train the CNN model, the trained CNN model will iterate over all the data from the signatures extracted during the feature extraction process and the model will learn using the features during the training process. In the next step the trained model is used to predict the signature threshold which represents the authenticity of the signature. The signature threshold ranges from 0 to 1, which represents the authenticity of the signature. If the signature threshold is higher than a given number say 0.5, then the confidence of the CNN that the signature is authentic is 50%. The signature threshold is used as the judging criteria for the authenticity of the image.

We can expect the CNN model to be capable of identifying forged signatures with great accuracy after the CNN has been trained, the accuracy can be further improved by training the CNN in multiple epochs and by improving the quantity of the training dataset.

**IV. CONCLUSION**

By following the above mentioned steps and performing signature verification, we can achieve the following results:

1. We can use above mentioned process to simulate an offline signature forgery detection method which is viable and useful.
2. We can detect forged signatures using the above mentioned method.
3. It is expected that the above mentioned method is more accurate and faster in detection of forged signatures than a trained professional to identify the signatures.

4. It is expected that the accuracy of detection of forged signatures is very high and less time and resource consuming.

This method can be implemented in places that require signature authentication and verification, it can be used faster hence processing a large number of signatures and can help save resources and time while detecting forged signatures.

V. FUTURE WORK

The project can progress further and can be improved by the following methods to improve the accuracy of offline signature forgery detection:

1. It can be improved further by improving the dataset quantity and quality to train the model.
2. The training model can be trained in multiple epochs to produce more accurate results.
3. It can use more types of algorithms when training the model while including features extracted in different scenarios.
4. It can be used in combination with online signature authentication methods to produce a higher accuracy of signature forgery detection if it involves crucial signatures of important dignitaries.

VI. REFERENCES


