International Journal of Engineering, Business and Management (IJEBM)
ISSN: 2456-7817
[Vol-7, Issue-2, Mar-Apr, 2023]
Issue DOI: https://dx.doi.org/10.22161/ijebm.7.2
Article DOI: https://dx.doi.org/10.22161/ijebm.7.2.1

Offline Signature Recognition and It's Forgery Detection using Machine Learning Technique

Malay Karmakar

Department of Computer Science, IIT Kharagpur, India

Received: 25 Jan 2023; Received in revised form: 25 Feb 2023; Accepted: 05 Mar 2023; Available online: 13 Mar 2023

Abstract— *Signature verification is an important aspect in today's World. Signature has been verified in Banks, Government Agencies, Universities (Degree Verification) etc. Signature can involve in its shape, size, pressure, speed and angle. From a population of Signatures an original signature can be found out and distinguished. In this Paper for Forgery signature Detection we use two algorithm viz.Harris Algorithm and Surf Algorithm. We have also discussed about CNN Algorithm. Moreover in this paper we take the x-y co-ordinate of the real signature and also the x-y co-ordinate of the forged signature and compare among the two. We have used Python Programming for plotting the graphs whereas the graph can be plot using Matlab, R, Microsoft Excel and Python.*

Keywords— ***Python, R, CNN Algorithm, Harris Algorithm, Surf Algorithm, Matlab, Microsoft Excel***

INTRODUCTION

Biometrics can literally be defined as the biological characteristics of aperson that can be used for identification purposes. The development of a biometric system has two main objectives: the identification of a person and the verification of his identity. Generally, applications of biometrics have been deployed for access control and monitoring. These applications are not yet widespread in various organizations such as banks and financialinstitutions. Biometric systems can classifypeople based on their physical or behavioral characteristics.

Physical characteristics refer to a person's biological characteristics, such as finger prints, deoxyribonucleic acid (DNA), iris, and facial features. These biological characteristics are unique to each individual and remain constant over a long periodoftime. Therefore, biometric systems that rely on physical characteristics are mostly sufficientlyaccurate and reliable for identification purposes, including one-to-many comparisons. Behavioral traits, on the other hand, refer to an individual's behavior, such as signature, gait, and voice. These traits change over time, making iteasierfor a skilled Impostertopass. Therefore, designing an accurate biometric system based on behavioral characteristicswillbeadifficulttask. Even as technology advances, handwritten signatures remain the most widely accepted form of authentication for legal documents, financial transactions, checks, loan and mortgage documents, insurance and compliance documents, commercial contracts, etc. The purpose of signature verification is to identify forged signatures to reduce the risk of hacking and crime. Signature verification systems should automatically differentiate between biometric samples.

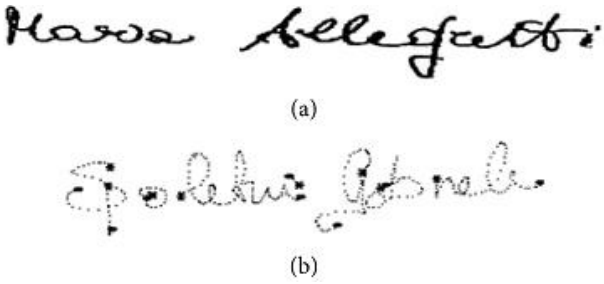
Type of Signature Forgeries:

- Unskilled/Trace-Over Forgery: It appears as a faint indentation on a piece of paper underneath.
- Skilled Forgery: These are done by perpetrators that has accesses to one or more samples of the authentic signatures.

The different types of signature traits are given below:

- Shaky handwriting (static).
- Lifting Pen (Dynamic).
- Letter Proportions (static).
- Signature Dimensions (static).
- Signature Angle (static).
- Close Similarity between two or more signatures (static).
- Speed of the signature (Dynamic).
- Pressure of the Pen (Dynamic).

- Pressure that changed its pattern (Dynamic).
- The Pattern of Acceleration (Dynamic).
- The Smoothness of Curve (Static).



(a)

(b)

Fig A. Sample of Signatures (a) Offline Signatures. (b) Online Signatures

What is CNN (Convolutional Neural Network) Algorithm?

A CNN Algorithm is a subset of Machine Learning Algorithm. It is one of the various types of Artificial Neural Networks which is used for different applications and data types. It is a kind of network architecture for Deep learning and is specifically used for Image recognition. It is suitable for application involving natural language Processing (NLP), Language translation and speech recognition.

CNN Layers:

There are three kind of CNN Layers viz.

- Convolutional layer.
- Pooling layer.
- Fully Connected Layer.

How Convolutional Neural Networks Works:

A convolutional neural network (CNN) is an artificial network commonly used inimage recognition and processing. This type of network extractsfeaturesfromimagesorotherinputdata using a series of filters called convolutional layers. These layers are designed to identify characteristic patterns in the input data that are important for the task at hand. For example, in this article, the first convolutional layer canrecognize basic shapes and edges, while subsequent layerscanrecognize more complex featuressuchas corners or textures. Once featuresareidentified,a CNN uses a process called clustering to reduce the dimensionality of the data and extract the most important features. This is usually done using a technique called max pooling. After the data has been processed throughconvolution and pooling layers, it is typically passed through one or more fully connected layers, similar to traditional neural networks. These layers use the extracted features to make predictions or classifications based on the input data.

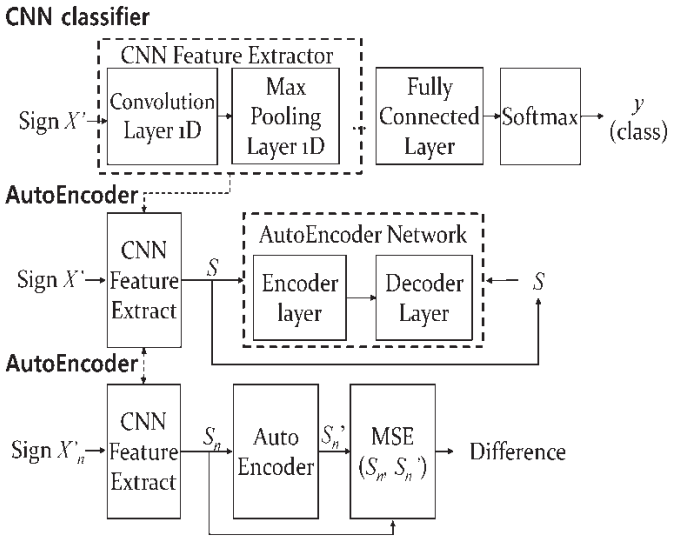


Fig B. Proposed Architecture of the CNN-AE model

Fig C. Example of Forgery Signatures.

Python Program to distinguish an original signature from a forged One using Graph:

#real signature

X1= [0, 1, 2, 3, 4, 5]

y1= [0, 2, 4, 6, 8, 10]

#forged signature

X2= [2, 3, 4, 5, 6, 7]

y2= [2, 4, 5, 7, 10, 12]

#plotting the graph

Plt.plot (x1, y1, color='blue',

Label="Real signature")

Plt.plot (x2, y2, color='red',

Label="Forged Signature")

#adding labels

Plt.xlabel ('x-axis')

Plt.xlabel ('y-axis')

Plt.title ('Real vs Forged Signature')

#showing the legend

Plt.legend ()

#displaying the graph

Plt.show ()

Steps to plot graphs using python Programming:

Plotting graphs in Python is easy and convenient with the help of various libraries such as Matplotlib, Seaborn,Pandas and others.Here are the steps to plot a graph using python.

- Install the libraries:
  Before plotting the graph, I need to install the libraries such as Matplotlib,Seaborn,and Pandasetc.I can install them using the pip command.
- Import the libraries:
  Once the libraries are installed, I need to importthem in my program. I can do this by using the import keyword.
- Data Preparation:
  The next step is data preparation. I can use various methods such as reading from a CSV file, creating a data frame, etc.
- Create the graph:
  Once the data is ready, I can create the graph. I can do this by using the plot () function of the Matplotlib library.
- Customize the graph:
  I can customize the graph by adding labels, colors, and other features. I can do this by using the various functions of the Matplotlib library.
- Display the graph:
- Finally, I can display the graph by using Plt.show () command.

Support Vector Machines (SVM)?

Support Vector Machines (SVM) is a popular Supervised Machine Learning algorithm that is based on the concept of finding a hyperplane in a high-dimensional space that can separate the data into different classes. The goal of SVM is to find the best hyperplane that maximizes the distance between the closest data points of different classes, which is called the margin. The points of data or the data points that lie close to the hyperplane are known as support vectors.SVM can be used both in regression and classification problems, and it is particularly useful when dealing with high dimensional data that cannot be easily visualized.SVM has several advantages ,including its ability to handle both linear and non-linear data, its robustness to outliers, and its capacity to work on small or large datasets.SVMs have found several applications in various domains, including image classification, text classification, and bioinformatics.

- Types of SVM?
  1) Linear SVM
     Linear SVM is used for linearly separable data, which means if a setofdata can be divided into two classes with a straight line, this data is called linearly separable data and the classifier is used as alinear SVM classifier.
  2) Non-Linear SVM.

Nonlinear SVM is used for data separated inanonlinearway, which means if thedataset cannot be classified using a straight line, that data is callednonlinear data and the classifier used is called a nonlinear SVM classifier.

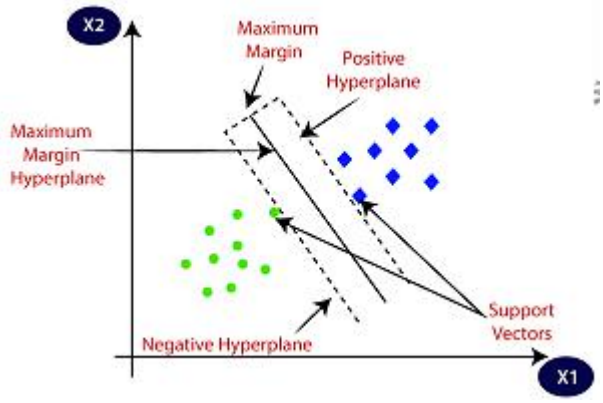


Fig D.A Support Vector Machine

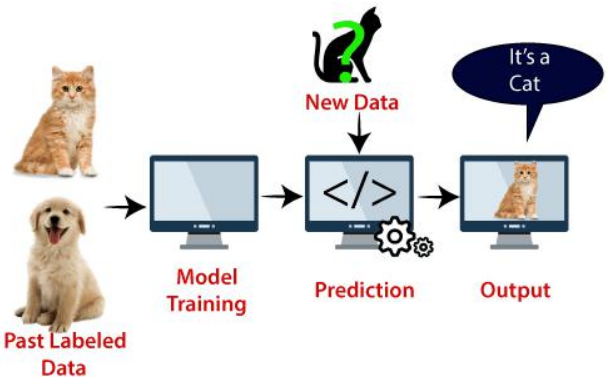


Fig F. SVM Algorithm can be used in Face detection, image classification and text categorization

How does SVM Technique works?

SVMswork by mapping an input space to a higher-dimensional feature space.It uses kernel functions to transform input data into a high-dimensional space where classes become separable.The hyperplane is then used to

separate the classes in the transformed feature space.The hyperplane with the largest margin or distance between itself and the closest data points of thetwo classes is chosen as the final classifier. SVM thus uses a cost function to know about the optimal hyperplane. It tries to minimize the error rate or the number of misclassified data points by adjusting the hyperplane.The regularization parameterC controls the trade-off between reachingthe minimum error rate and maximizing the margin.

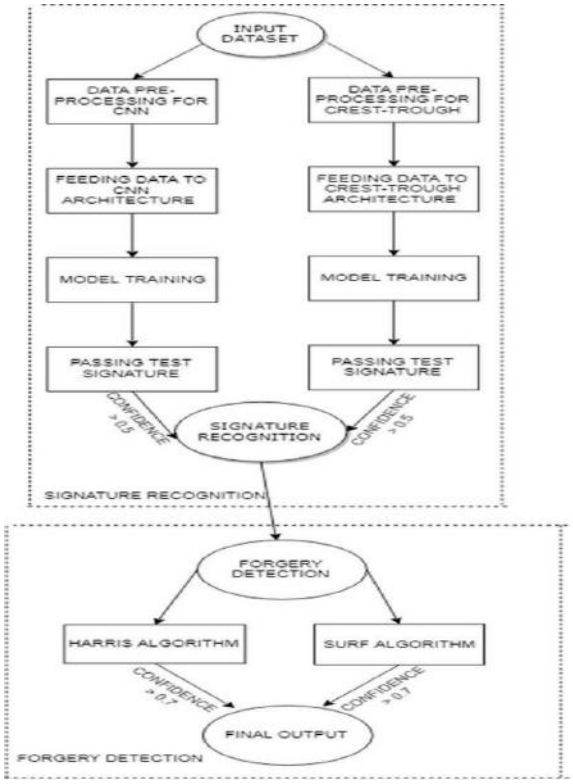


Fig G. Proposed Diagrams for Signature recognition and forgery detection

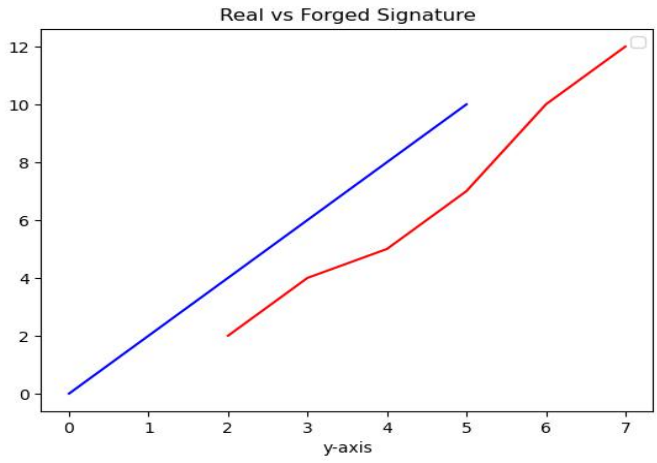


Fig H. Plot of the Python Program using the data points

For Forgery Detection two algorithms are used a) Harris Algorithm. B) Surf Algorithm.

Harris Algorithm:

Harris' algorithm is a popular feature extraction algorithm used in computer vision and machine learning todetect corners in images. It determines the verticesof an image by analyzing the change in brightness between adjacent pixels in different directions.

- The Harris algorithm works as follows:
  1) Compute the gradient of the image in both x and y directions.
  2) Calculate the products of the gradient for each pixels.
  3) Calculate a sum of squares of products for a window surrounding each pixel.
  4) Calculate the Harris response function value of each pixel using the determinant and trace of the windows corresponding covariance matrix.
  5) Apply the threshold to filter out low-quality corner points.

  For Computing the corner response function R of each pixel location using the following equation:

  R= determinant (M)-k*trace^2(M).

  Where M is the squared gradient Matrix, trace (M) is the trace of M, and k is the empirically determined constant.

Surf Algorithm:

The SURF (Speeded up Robust Features) algorithm is a feature detection and description algorithm for object recognition and computer vision tasks. It is used to identify and describe localfeatures such as edges, corners and nodes.

- The Surf Algorithm works by analyzing the intensity, contrast, and orientation of image pixels in small regions or patches. The algorithm scales the image, detects the features, and then computes the features descriptors using the sum of Haar wavelet response in a surrounding region.
- The key advantage of the surf algorithm over other feature detection and description algorithms are its robustness to scale, rotation, and illumination changes, as well as its ability to compute features quickly.
- The Surf Algorithm is widely used in various applications, such as image registration, object recognition,3D reconstruction, and robotics. It is implemented in many popular computer vision libraries, such as Open CV and MATLAB.

PROPOSED WORK

In this paper I have used Support Vector Machine (SVM) Algorithm to find out the data points of the signature and also for the forged signature. Then used the Python Code to plot the Graph between Original Signature and Forged Signature.

FUTURE WORK AND SUGGESTIONS

Many aspects can be considered in future work, which can be increasing the number of reference signature Images in case of both Offline and Online signatures. We can further use multilingual signature datasets for a large number of user. Also deep learning can be used for further verification of the signatures.

REFERENCES

[1] Seungsoo Nam;Hosung Park; Changho Seo; Daeseon Choi; Forged Signature Distinction Using Convolutional Neural Network for Feature Extraction; Appl. Sci. 2018,8(2), 153; 23rd January 2018.

[2] H.Srinivasan; S. N. Srihari; Matthew J.Beal; Machine Learning for Signature Verification; Center of Excellence for Document Analysis and Recognition (CEDAR);Buffalo NY.

[3] Guo, J.K., Doermann, D., Rosenfield, A.: Local correspondences for detecting random forgeries, Proceedings of the International Conference on Document Analysis and Recognition (1997) 319–323.

[4] Z. Hashim; H. M.Ahmed;Ahmed Hussein Alkhayyat; A Comparative study Among Handwritten Signature Verification Methods Using Machine Learning Techniques;15th October 2022.

[5] M. A. Taha and H. M. Ahmed, "Iris features extraction and recognition based on the local binary pattern technique," in *Proceedings of the 2021 International Conference on Advanced Computer Applications (ACA)*, pp. 16–21, Maysan, Iraq, July 2021.

[6] K. Radhika and S B. Gopika, "Online and offline signature verification: a combined approach," *Procedia Computer Science*, vol. 46, pp. 1593–1600, 2015.

[7] A. Kumar and K. Bhatia, "A survey on offline handwritten signature verification system using writer dependent and independent approaches," in *Proceedings of the 2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA) (Fall)*, pp. 1–6,at Bareilly, India, September 2016.

[8] Jivesh Poddar; Vinanti Parikh; Santosh Kumar Bharti; Online Signature Recognition and Forgery Detection using Deep Learning,6-9th April 2022.