**“COMPARATIVE ANALYSIS ON MACHINE LEARNING APPROACHES FOR SIGNATURE DETECTION FROM DOCUMENTS”**

A Synopsis submitted in Partial Fulfillment of the Requirements for the Degree of

# Master of Science in Computer Science

**By,**

# VIJAY S N

**NITHIN KUMAR D R**

# DEPARTMENT OF STUDIES IN COMPUTER SCIENCE

# UNIVERSITY OF MYSORE, MANASGANGOTHRI

# MYSORE –570006

# April 2022

**COMPARATIVE ANALYSIS ON MACHINE LEARNING APPROACHES FOR SIGNATURE DETECTION FROM DOCUMENTS**

# Abstract:

Handwritten signatures can be employed as a sign of confirmation in a wide variety of documents, namely, bank checks, identification of documents, and a variety of business certificates, and contracts. Since those documents are present in complex backgrounds, the automatic extraction of handwritten signatures from documents remains an open task in the Offline Signature Verification field. Other factors that make it challenging are the various variations present in handwritten signatures such as location, type of ink, color and type of pen, and the type of stroke. In this work, we propose an approach to locate and extract the pixels of handwritten signatures on identification documents, without any prior information on the location of the signatures. This proposed work presents techniques and methods evolved for signature detection and extraction.

# Introduction:

**About the Problem:**

The document images captured by smartphone cameras are usually presented with distortions and background noise. Therefore, treating these images in such a way that only handwritten signatures can be extracted for analysis of their characteristics becomes a challenging task in image processing. These images do not always present the desired features or the expected quality, negatively influencing the process of recognition and classification of these handwritten signatures. The situation may become more critical if the image of the source document has unwanted characteristics, such as imperfections, backgrounds, printed text, shape, and variations in size.

Another condition that can affect the quality of the attributes of handwritten signatures occurs when the image presents some distortion, such as perspective, inclination, scale, or unexpected resolution, all of them when scanning photos. All these interference can also harm the verification systems of handwritten signatures with the increase of false positives or false negatives in the classification process.

**Existing System:**

Several approaches have been developed in the field of machine learning and statistical methods to perform the signature detection and verification tasks automatically. Among these approaches, we can mention techniques based on Artificial Neural Networks (ANN), Hidden Markov Models (HMM), Support Vector Machines (SVM), and Fuzzy Logic. Among the neural network techniques, it is important to mention Faster Region-based Convolutional Neural Networks (RCNN) and YOLOv2. Both models were adapted for logos and signature localization in noisy documents.

Many of the techniques that address signature verification use public databases. However, such bases as GPDS, and MCYT present images with a light background and dark signatures. This characteristic does not present an environment of great complexity for segmenting the signature pixels. Besides, due to the insertion of mobile devices and their growing popularity.

**Motivation for the Proposed System:**

A handwritten signature is a biometric authentication method widely used for personal documents and legal contract validations. Besides, experts in forensic analysis examine handwritten signatures to certify the authenticity of the writing and reveal possible fraud, which in some cases can mean high-value financial losses. Several commercial and banking applications, for example, use images captured by smartphones for transactions, payments, account opening, and copies of documents.

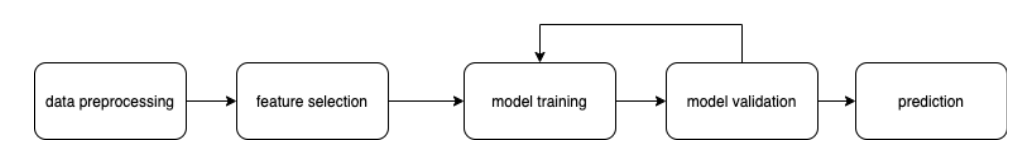
**Proposed System:**

In this work, we propose an approach to the pixel-level segmentation of handwritten signatures on images. Our model will be trained with document images with the same characteristics and interference that can arise in a real-world scenario. Our model will be able to get around the problems presented during the capture of signature images in different identification documents in noisy environments. Our proposal will also enable the acquisition of signatures with greater fidelity in the strokes regardless of the types of pen, ink, and background, preserving the graphic characteristics. These features are used by forensic experts and may be applied in future systems for verifying handwritten signatures with a bias in forensic science. We use a Fully Convolutional Network (FCN) for signature segmentation on identity document images with refinement layers for the alpha channel of the image.

**The Objective of the Proposed System:**

Extracting handwritten signatures from images with a complex background and noise interference, imperfections, backgrounds, printed text, shape, and variations in size from the documents. The signature verification systems can be used on handwritten signatures extracted from different images of ID documents captured by various computing devices such as smartphones and scanners. Achieving the maximum fidelity of a signature’s characteristics can have a positive impact on the classification results, including guaranteeing the graphotechnical characteristics used by forensic specialists.

**Probable Architecture of the Model:**



The analysis of the literature shows, that many algorithms involve image preprocessing techniques (e.g. document rectification), deal with limited forms of documents (e.g. bank cheques), and employ handcrafted features together with multi-tier approaches. The other observation is that there is a lack of methods aimed at the detection of all possible classes of interesting objects in scanned documents. It is caused by the characteristics of objects that have a lot of variability in how they appear.

**Proposed Methodology:**

**Datasets**

The handwritten signatures blended in the document image were captured by smartphones under different backgrounds, distances and Another aspect of the dataset is the type of pen used, in which different types and different colors were used to avoid possible bias in the learning of the network regarding the color and type of ink. A small area of the target handwritten signature is highlighted to give an idea of the challenge to detect the signature pixels in such conditions.

**System Requirements:**

**Hardware Requirements:**

ROM: 256 GB SSD/512 GB Hard disk

RAM: Minimum 4 GB

Processor: Intel Core i3/i5/i7 or AMD Ryzen 3/5

**Software Requirements:**

Language: Python

OS: Windows/8/9/10/11

**Conclusion:**

Extracting handwritten signatures from images with a complex background and noise interference, such as identification documents, is a complex but promising task for the application of signature verification systems. As possible future works, it is also possible to point out the implementation of a semantic segmentation solution based on the proposed architecture for the application domain relevant to the structure and characteristics of documents. An example is the detection of anomalies in documents forged or created by gross forgeries—an interesting area of forensic science. In addition to the implementation, as mentioned earlier, instantiating the model to other application domains would allow a way to assess the generalization capacity of the proposed model, as well as to detect possible adjustable points in the architecture.

**References**

[1] V. K. S. L. Melo and B. B. L. Dantas, “A fully convolutional network for signature segmentation from document images,” in 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR). IEEE, 2018, pp. 540–545.

[2] A. Pansare and S. Bhatia, “Handwritten Signature verification using neural network,” International Journal of Applied Information Systems, vol. 1, no. 2, pp. 44–49, 2012.

[3] M. Diaz, M. A. Ferrer, G. S. Eskander, and R. Sabourin, “Generation of duplicated off-line signature images for verification systems,” IEEE transactions on pattern analysis and machine intelligence, vol. 39, no. 5, pp. 951–964, 2016. [4] C. A. B. MELLO, W. P. SANTOS, and A. L. I. OLIVEIRA, Digital Document Analysis and Processing. New York: Nova Science Publishers Inc, 2012.

[5] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.

[6] P. G. S. Silva, C. A. M. Lopes Junior, E. Lima, B. Leite Dantas Bezerra, and C. Zanchettin, “Speeding-up the handwritten signature segmentation process through an optimized fully convolutional neural network,” in 2019 15th International Conference on Document Analysis and Recognition (ICDAR), Sep 2019.

[7] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE transactions on image processing, vol. 13, no. 4, pp. 600–612, 2004.

[8] J. Coetzer, B. M. Herbst, and J. A. du Preez, “Offline signature verification using the discrete radon transform and a hidden Markov model,” EURASIP Journal on applied signal processing, vol. 2004, pp. 559–571, 2004.

[9] I. V. Anikin and E. S. Anisimova, “Handwritten signature recognition method based on fuzzy logic,” in 2016 Dynamics of Systems, Mechanisms, and Machines (Dynamics). IEEE, 2016, pp. 1–5.