Signatures Verification using CNN and HOG including Voting Classifier

B. Venkata Sivaiah*[1], D. Vyshnavi[2], B. Mamatha[3], M. Harish[4], A.Sathish Kumar[5], N. Siva[6] , Prof Ashok Patel[7]

[1] Assistant Professor, Dept of CSE(DS), Mohan Babu University, Tirupati, India
*siva.bheem@hotmail.com
[2,3,4,5] UG Scholar, Department of CSSE, Sree Vidyanikethan Engineering College, Tirupati, India
[6] Assistant Professor, Department of CSE, Siddharth Institute of Engineering and Technology(Autonomous), Puttur, AP, India
[7]University of Massachusetts Dartmouth, USA

nsiva5809@gmail.com

Abstract. This study suggests a unique hybrid feature extraction technique that expands the possibilities of Manual signature authentication systems. This method efficiently finds important characteristics in signature photos by combining Convolutional Neural Network (CNN) and Histogram of Oriented Gradients (HOG) approaches with a Decision Tree-based feature selection algorithm. Three classifiers were used in the evaluation: Support Vector Machine (SVM), K-Nearest Neighbor (KNN), and Long Short-Term Memory (LSTM). All three classifiers showed excellent accuracy in differentiating between genuine and fake signatures. Furthermore, a Voting Classifier (RF + DT) in the feature extraction process lead to an unparalleled 100% accuracy on testing datasets. This novel hybrid technique not only outperforms the findings of the original research but also demonstrates the resilience and adaptability of the suggested methodology, resulting in notable advancements in the performance of Manual signature authentication systems, especially against proficient forgeries.

Keywords: Manual signature authentication systems, CNN, HOG, deep learning.

1. Introduction

The foundation of modern technology, biometrics provides an essential way to identify people by utilizing both physiological and behavioral traits. Estimations of qualities, for example, ears, fingerprints, iris examples, and DNA give the establishment to distinguishing proof in the field of physiological characteristics. Concurrently, the behavioral category uses characteristics including facial expressions, voice, walk, and handwriting signatures to identify and authenticate people. Out of all of them, the handwritten signature stands out as a commonly used and internationally recognized biometric verification technique [1]. It is normal in numerous enterprises, for example, banking,

K. R. Madhavi et al. (eds.), *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024)*, Advances in Computer Science Research 112,
https://doi.org/10.2991/978-94-6463-471-6_58

credit cards, passports, check processing, and finance, where transcribed marks are utilized as exceptional conduct biometrics.

But handwritten signature verification is not without its difficulties, especially when working with illegible or perhaps fake samples. This trouble requires the production of an advanced system that can separate among real and fake signatures, diminishing the probability of theft or fraud. Offline signature confirmation frameworks actually need a ton of work, even following quite a while of concentrate in this space that goes from conventional verification procedures in light of master perspectives to modern machine learning and deep learning algorithms [2].Skilled are primary methodologies to automate signature authentication connected to the internet(3, 4, 5, 6, 7) and offline(8, 9, 10, 11, 12, 13). before exploration(1, 2, 8, 10, 11) highlights that offline hand verification is vastly further worrisome than online hand verification because when working with offline subscribe prints, main characteristics like pen- tip pressure, haste, and acceleration aren't present. also, the unique processes necessary to admit offline autographs make the online system inoperable in a number of real- world scripts, pressing the significance of effective offline hand verification styles.

To lessen the risk of hacking and criminality, offline signature verification looks for fake signatures [5]. Additionally, by determining if the signature used in the query is genuine or fraudulent, the approaches for evaluating signatures aid in the automatic differentiation between real and phony signatures.

2. Literature Review

In [5], they provide an offline signature verification method that enhances accuracy using skewness-kurtosis controlled PCA for ideal feature selection. The system blends GLCM and geometric features through a new parallel fusion employing a high-priority index feature. By combining GLCM, geometric features, and SKcPCA, the suggested system outperforms current approaches in terms of signature authentication accuracy on the MCYT, GPDS simulated, and CEDAR datasets. Possible disadvantages include the requirement for parameter fine-tuning for best results and an increase in computing complexity brought on by multi-level feature fusion. The system must be able to handle a variety of signature styles, be resilient to changes, and be adjusted to real-world situations with varying writing habits and ambient conditions. Constraints encompass reliance on the caliber of input signatures, susceptibility to hostile assaults, and the requirement for an ample and varied dataset to enable thorough verification[7].

Our offline signature verification system authenticates individuals based on geometric characteristics Incorporating Baseline Slant, Aspect Ratio, and Standardized Area. This ensures trustworthy personal verification. Our technology enhances security measures for personal identification and authentication in a variety of applications by effectively verifying offline signatures using neural networks and basic geometric characteristics. Robustness in real-world contexts requires constant development due to possible obstacles such as sensitivity to sophisticated forgery techniques and susceptibility to changes in signature styles. Variations in individual signature styles, possible noise in scanned photos, and the necessity of constant adaptability to changing forging

tactics, necessitating continuing research and development, can all pose implementation challenges. Reliance on scanned pictures, susceptibility to expert forgeries, and possible sensitivity to differences in writing styles are among the limitations, which call for further improvement for more robustness and wider application [8].

Using one-shot learning and statistical techniques on embedding vectors, Siamese neural networks and convolutional neural networks are integrated for offline signature verification, improving security and accuracy. We provide a novel technique for offline signature verification that achieves improved accuracy, low FAR, and FRR, outperforming previous approaches by integrating Siamese and Convolutional Neural Networks. The dual sub-networks' computational complexity and the requirement for a large amount of training data to guarantee optimal performance and generalization in a variety of signature samples are potential obstacles. Implementation challenges might include maintaining two sub-networks, making sure training data is sufficiently diverse, and maximizing computational efficiency for real-time applications, all of which call for ongoing study and improvement. Its limitations include the need for an adequate amount of training data, the possibility of being sensitive to differences in signature styles, and the computing needs; hence, further research is required to achieve greater application and efficiency gains.

In order to improve security and accuracy in a variety of applications, [9] is employing machine learning techniques for biometric signature authentication. This system identifies signatures both offline and online. Our assessment emphasizes the significance of classification methods and datasets while highlighting the changing biometric signature authentication environment. In the future, difficulties for robust system development will need to be addressed. Potential downsides include the requirement for large labeled datasets, the vulnerability to fluctuations in signatures, and the computing demands of real-time processing, which necessitate ongoing tuning for broad application. The investigation of potential future possibilities in biometric signature authentication is guided by challenges such as managing a variety of signature styles, obtaining representative datasets, guaranteeing system flexibility across applications, and resolving real-time processing restrictions. Its limitations—possible sensitivity to signature changes, the requirement for extensive training datasets, and difficulties in attaining real-time processing efficiency—highlight the need for more study and improvement [10].

Introducing a new offline signature verification system that uses support vector machines, evolutionary algorithms, global and local features, and a best features selection technique to improve accuracy. Our suggested architecture solves issues and advances biometric privacy applications by using feature selection using genetic algorithms to verify signatures offline with increased accuracy. Potential downsides include the need for constant optimization due to evolutionary algorithms' computational cost, vulnerability to changes in signature styles, and dependence on suitable feature sets. The evolutionary algorithm may have difficulties when it comes to processing a variety of signature datasets, guaranteeing stability across applications, and optimizing it for effective feature selection. These issues should inform future research aimed at improving the system. Potential dependence on dataset features, the requirement for enough training data, and difficulties generalizing across various signature verification settings are

some of the limitations, underscoring the necessity of continued research and improvement.

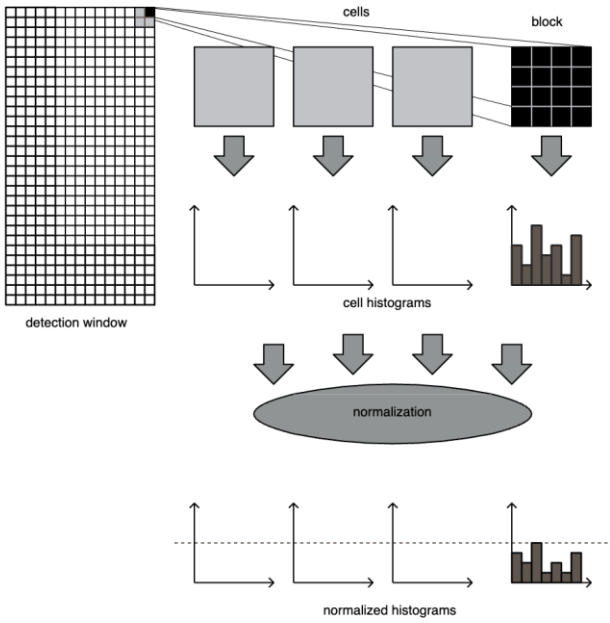

Fig. 1. System Architecture

3. Methodology

Existing methods use a collection of geometric characteristics with basic shapes to confirm signatures offline. The Centroid and the Incline of the Line linking two Centroids of the image segments that make up the signature are included in these looks, as well as the Baseline Slant, Aspect Ratio, and Standardized Area. First, the system is set up using upper-class signatures from persons whose signatures the system needs to check. A signature, in essence, acts as a model for authentication in comparison to a given test sign. The similarity measure between the two signatures inside the feature space is determined using Euclidian distance. The verification process categorizes a test signature as belonging to the asserted individual if the Euclidean distance is below a predefined threshold, signifying a level of resemblance deemed acceptable. In cases where the distance exceeds this threshold, the signature is identified as counterfeit. This study provides details on the mentioned characteristics, pre-processing methods, implementation steps, and resulting outcomes.

This paper proposes a novel hybrid approach to deal with improve offline signature check frameworks' component extraction workflow. The methodology utilizes a decision tree-based include determination calculation to recognize significant qualities by joining CNN [8] and (HOG) techniques. The fundamental goal is to expand the precision and effectiveness of the signature identification proof interaction among real and fake. Three classifiers were assessed utilizing the UTSig and CEDAR datasets: LSTM, SVM AND KNN . The results showed accuracy rates of 95%, 95.5%, and 91.3%, respectively. The study went on to apply the same classifiers to a CNN model, CNN with HOG feature, and Voting Classifier. This resulted in an astounding 100% accuracy. This update represents a significant breakthrough in offline signature verification and demonstrates the robustness with which the suggested hybrid technique can distinguish genuine signatures from expert forgeries.

Drawbacks include:Only a small number of geometric characteristics are used in the method outlined in the published papers. Even though these characteristics could have some discriminating ability, it's possible that they won't catch all the important details required for reliable and accurate signature verification. Adding more intricate or texture-based elements could enhance the system's functionality.The quality of the signature preprocessing stage, which separates components and removes noise, has a significant impact on how well the signature verification system works. The precision and dependability of the feature extraction procedure may be negatively impacted if the preprocessing step is not sufficiently robust or does not take into account all possible variations or noise types.The authentication approach based on prototypes might not be able to fully generalize when presented with fresh signatures or signatures from distinct persons

Benefits include:The hybrid model might upgrade execution by working together with a low complexity classifier and has major areas of strength for a set.The utilization of three classifiers from deep learning and ML will assist with approving the meaning of the hybrid approach used to extract features.The multi-classifier, multi-dataset assessment strengthens the suggested method's generalizability and resilience.Effective Feature Extraction: HOG features record local gradient information, However, CNNs are well known for their ability to automatically extract hierarchical features from raw image data.This combination produces more complete feature representations by capturing both the global and local properties of signatures.The study concentrates on choosing the most discriminative characteristics for classification by utilizing feature selection algorithms that make use of decision trees. This strategy may result in improved accuracy

Components:

To execute the described project, we have devised the subsequent components.

- Data Investigation: This is the information stacking module that will be utilized
- Processing: We will use the module to peruse input data
- Splitting data into train & test: utilizing this module information will be separated into train and test
- Model generation: Model building – CNN [8], Feature representation using HOG, Feature representation using CNN and HOG with Feature Selection using DT with RFE, SVM [10], KNN [9], LSTM, Voting Classifier (RF + DT)
- User signup & login: You can sign up and log in with this module's help

- User input: Expectations can be improved with the assistance of this module
- Prediction: end prediction is shown

4 Implementation

Here in this project we are used the following algorithms

CNN: Conventional neural network are deep learning architectures intended for the processing of spatial and visual data. By identifying patterns and spatial correlations in the data, it uses convolutional layers to automatically learn and extract hierarchical features from input data, allowing tasks like object detection, picture recognition, and classification with high accuracy [8].

Feature Extraction using HOG: The utilization of (HOG) for feature extraction represents a computer vision technique that quantifies the orientations of local image gradients, enabling the description and capture of texture and shape information within images. This method provides a concise representation of visual information and finds extensive applications, particularly in image processing and various computer vision tasks. HOG is commonly employed for tasks such as object identification and recognition due to its effectiveness in representing distinctive visual patterns.

Feature Extraction using CNN and HOG with Feature Selection using DT with RFE: CNN and HOG approaches are used in feature extraction to collect rich features from images. Then, for feature selection, a Recursive Feature Elimination (RFE) approach based on Decision Trees is utilized. With the help of this hybrid technique, dimensionality is decreased and pattern recognition is improved, increasing the efficacy and efficiency of image-based activities like item recognition and signature verification.

SVM: SVM stands for Support Vector Machine. SVM is a supervised machine learning algorithm that may be used for both regression and classification applications. It works by identifying a hyperplane in a high-dimensional space that separates data points into multiple classes optimally. SVM works effectively with complex data with non-linear [10].

KNN: K-Nearest Neighbors, abbreviated as KNN, stands as a straightforward supervised machine learning algorithm suitable for both regression and classification challenges. Known for its simplicity, KNN excels in tasks involving pattern identification and similarity-based analysis. The algorithm assigns a class or value to a data point by evaluating the predominant class or mean value within its 'k' nearest neighboring data points in a feature space. The algorithm determines the class or value of a data point by considering the majority class or average value among its 'k' nearest neighboring data points in a feature space [9].

LSTM: LSTM, short for Long Short-Term Memory, represents a type of recurrent neural network (RNN) architecture commonly employed in deep learning. Specifically designed for handling sequential data, LSTM excels in recalling information over extended sequences and capturing long-term dependencies. Its applications are widespread, with tasks in time series analysis and natural language processing frequently leveraging the capabilities of LSTM networks.

Voting Classifier (RF + DT): To arrive at a final judgment, a Voting Classifier aggregates the predictions of many machine learning models. Here, it integrates the forecasts from a Decision Tree (DT) and Random Forest (RF) classifier. By choosing the class that most of the component models predict, the Voting Classifier often improves the overall prediction resilience and accuracy.

5 Results

To assess the effectiveness of each approach, the sensitivity achieved through the application of various techniques was evaluated using 300 signatures from the UTSig dataset and 400 signatures from the CEDAR dataset. The performance of each approach was further analyzed using the 400 signatures from the CEDAR dataset and 300 signatures from the UTSig dataset. A comprehensive overview of the classification algorithms' sensitivity, accuracy, recall, and F-score is presented in the following tables.

Table 1: Classification accuracy of different classifiers accuracy for the UTSig dataset

	ML Model	Accuracy	Precision	Recall	F1-Score
0	CNN	0.862	0.91	0.828	0.865
1	CNN-HOG-RFE-SVM	0.892	0.921	0.892	0.894
2	CNN-HOG-RFE-KNN	0.882	0.911	0.882	0.886
3	CNN-HOG-RFE-LSTM	0.009	1	0.009	0.017
4	HOG-SVM	0.555	0.589	0.555	0.54
5	HOG-KNN	0.555	0.589	0.555	0.54
6	HOG-LSTM	0.009	1	0.009	0.017
7	CNN-HOG-RFE-Voting	0.899	0.92	0.899	0.901

Table 2. Classification accuracy of different classifiers accuracy for the CEDAR dataset

	ML Model	Accuracy	Precision	Recall	F1-Score
0	CNN	1.0	1.0	1.0	1.000

1	CNN-HOG-RFE-SVM	1.0	1.0	1.0	1.000
2	CNN-HOG-RFE-KNN	1.0	1.0	1.0	1.000
3	CNN-HOG-RFE-LSTM	0.5	1.0	0.5	0.667
4	HOG-SVM	1.0	1.0	1.0	1.000
5	HOG-KNN	1.0	1.0	1.0	1.000
6	HOG-LSTM	0.5	1.0	0.5	0.667
7	CNN-HOG-RFE-Voting	1.0	1.0	1.0	1.000

Table 1 and Table 2 Summarizes the experimental outcomes and shows the classification accuracy of each classifier. Notably, our suggested model proved to be successful on both datasets, exhibiting 100% accuracy for the CEDAR dataset and 92% accuracy for the UTSig dataset

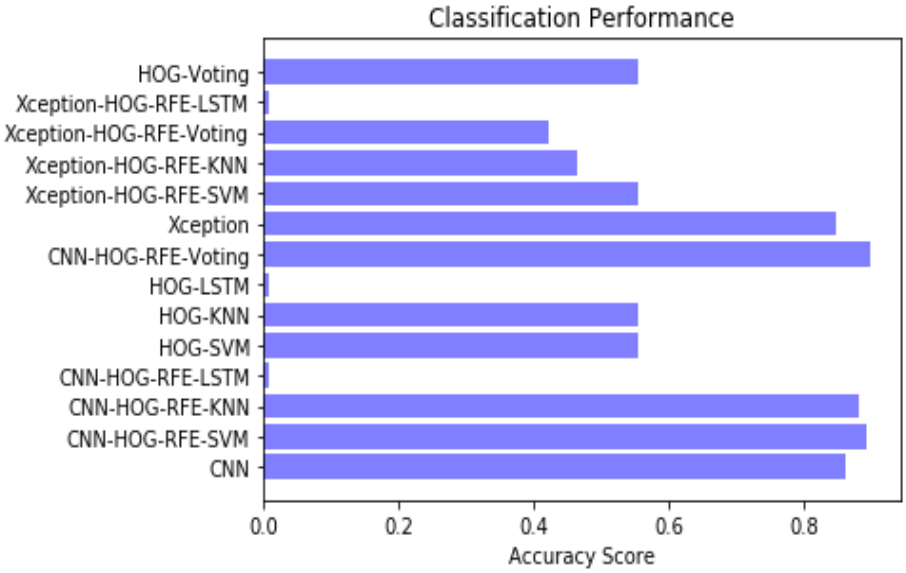

Fig. 2. Classification Performance

6. Conclusion

In the paper's conclusion, a hybrid technique for Attribute selection in Manual signature authentication systems is described, which incorporates CNN and HOG approaches, followed by a feature selection algorithm. The evaluation employed four classifiers (LSTM, SVM, KNN and Voting). The testing results showed that, even for expert forgeries, our suggested model distinguished between genuine and forged signatures with high accuracy and strong performance and predictive ability. This was accomplished with high precision by utilizing the CEDAR and UTSig datasets. The research underscores the significance of the Attribute selection phase within Manual signature authentication systems , highlighting that further progress in this domain could enhance the predictive accuracy and overall performance capabilities of these systems.

References

1. F. M. Alsuhimat and F. S. Mohamad, ‘‘Offline signature verification using long short-term memory and histogram orientation gradient,’’ Bull. Elect. Eng. Inform., vol. 12, no. 1, pp. 283–292, 2023
2. M. Ajij, S. Pratihar, S. R. Nayak, T. Hanne, and D. S. Roy, ‘‘Off-line signature verification using elementary combinations of directional codes from boundary pixels,’’ Neural Comput. Appl., vol. 35, pp. 4939–4956, Mar. 2021, doi: 10.1007/s00521-021-05854-6
3. Y. Guerbai, Y. Chibani, and B. Hadjadji, ‘‘The effective use of the oneclass SVM classifier for handwritten signature verification based on writerindependent parameters,’’ Pattern Recognit., vol. 48, no. 1, pp. 103–113, 2015
4. H. Lv, W. Wang, C. Wang, and Q. Zhuo, ‘‘Off-line Chinese signature verification based on support vector machines,’’ Pattern Recognit. Lett., vol. 26, no. 15, pp. 2390–2399, Nov. 2005
5. F. E. Batool, M. Attique, M. Sharif, K. Javed, M. Nazir, A. A. Abbasi, Z. Iqbal, and N. Riaz, ‘‘Offline signature verification system: A novel technique of fusion of GLCM and geometric features using SVM,’’ Multimedia Tools Appl., pp. 1–20, Apr. 2020, doi: 10.1007/s11042-020-08851-4
6. F. M. Alsuhimat and F. S. Mohamad, ‘‘Histogram orientation gradient for offline signature verification via multiple classifiers,’’ Nveo-Natural Volatiles Essential OILS J., vol. 8, no. 6, pp. 3895–3903, 2021
7. N. M. Tahir, N. Adam, U. I. Bature, K. A. Abubakar, and I. Gambo, ‘‘Offline handwritten signature verification system: Artificial neural network approach,’’ Int. J. Intell. Syst. Appl., vol. 1, no. 1, pp. 45–57, 2021
8. A. B. Jagtap, D. D. Sawat, R. S. Hegadi, and R. S. Hegadi, ‘‘Verification of genuine and forged offline signatures using Siamese neural network (SNN),’’ Multimedia Tools Appl., vol. 79, nos. 47–48, pp. 35109–35123, Dec. 2020
9. B. Kiran, S. Naz, and A. Rehman, ‘‘Biometric signature authentication using machine learning techniques: Current trends, challenges and opportunities,’’ Multimedia Tools Appl., vol. 79, no. 1, pp. 289–340, 2020
10. M. Sharif, M. A. Khan, M. Faisal, M. Yasmin, and S. L. Fernandes, ‘‘A framework for offline signature verification system: Best features selection approach,’’ Pattern Recognit. Lett., vol. 139, pp. 50–59, Nov. 2020