International Journal of Safety and Security Engineering
Vol. 15, No. 7, July, 2025, pp. 1481-1492
Journal homepage: http://iieta.org/journals/ijsse

A Deep Learning-Based Approach for Gender Prediction in Digital Forensics

Abeer D. Salman[1*], Aymen Jalil Abdulelah[1], Ali Al-Kubaisi[2]

[1] Electronic Computer Center, University of Anbar, Ramadi 31001, Iraq
[2] Department of Artificial Intelligence, College of Computer Science and Information Technology, University of Anbar, Ramadi 31001, Iraq

Corresponding Author Email: abeer.dawood@uoanbar.edu.iq

https://doi.org/10.18280/ijsse.150715

Received: 11 June 2025
Revised: 15 July 2025
Accepted: 26 July 2025
Available online: 31 July 2025

Keywords:
digital forensics, digital investigation, evidence integrity, SHA-256, ResNet18

ABSTRACT

In real life, everyone has unique characteristics and abilities that distinguish them from others. Handwriting is one of these characteristics. This feature has an important benefit in many applications, including digital forensics, cybersecurity, and many other fields. The evolution of handwriting systems becomes a necessary and urgent matter due to the development of technology that has increased the use of digital handwriting captured using scanners or electronic pens. In this research, a novel method was applied to enhance the investigation of digital forensics by using handwriting as evidence to determine the gender. The current methods suffer from many limitations related to accuracy and a lack of a uniform technique for evidence verification. This research addresses these limitations by ensuring evidence integrity through SHA-256 cryptographic hashing and performing gender classification using ResNet18-based convolutional neural networks, and thus it can address the limitations of the existing methods by proposing the first dual-purpose forensic framework. Earlier studies focus on either gender prediction or evidence verification, while this study integrates both critical forensic requirements in a unified pipeline. This work is enhancing digital forensics by providing investigators with a reliable, secure, and automated tool for handwriting-based gender identification. The QUWI database was used to evaluate the system, which holds 1,017 handwritten samples in Arabic and English. From the results obtained, we can say that the proposed method has achieved high efficiency in helping criminal investigators by analyzing evidence with high efficiency. The time taken to generate the hash of each document is 0.23 seconds, which means that it exceeds the standard requirement of 0.5 seconds, and this will make the system suitable for real-time forensic applications. Results prove superior performance with 97% gender classification accuracy, significantly outperforming existing methods.

1. INTRODUCTION

If a person does something that breaks the law without motive or justification, this act is considered a crime recognized by the state as a felony or misdemeanor [1]. Cybercrime is type of crime that includes hacking, data breaches, and online fraud [2]. Many pieces of evidence help the police find suspects in cyber and non-cybercrimes, including surveillance cameras and mobile phone data (photos, videos, text messages, and contacts, in addition to social media) [3]. In our world today, the rate of cybercrime has increased, and with it, the need for digital forensics (DF) has increased [4]. DF is the use of technological techniques to identify, collect, preserve, document, and analyze electronically stored data to ultimately provide evidence that facilitates criminal investigations in uncovering crimes committed using electronic devices and identifying the circle of suspects [2, 5, 6]. DF holds immense importance in modern-day crime investigation and detection. It is a way of finding evidence in digital media like hard drives, floppy disks, and pen drives [7]. DF is normally applied in computer-related crimes such as intellectual property theft, unauthorized access to computer systems, unauthorized access to confidential data, e-terrorism activities, money laundering, misuse of data, etc. [6]. The evidence comprises several forms of data, including audio, video, and text, stored on electronic devices.

Biometrics is defined as the utilization of physiological and behavioral characteristics to determine a person's criminal identity, like face, fingerprint, and voice recognition, iris scanning, and handwriting examination. The efficacy of biometric systems is contingent upon many recognition processes, including feature extraction, feature robustness, and feature matching. Biometric data encompasses detailed information about individuals, which plays a significant role in evaluating cases within DF. Forensic science has traditionally used manual methods for identification, but modern science has developed novel approaches by combining biometrics with computer intelligence [1, 8]. The application of biometric characteristics, including voice, facial recognition, handwriting, and gait, for the profiling of computer users (e.g.,

gender identification) has been a notable study interest in the last decades. Current technologies have several shortcomings that may affect their efficiency in dealing with digital crimes. (1) They may be noisy; (2) they require extra tools; (3) they can be quickly avoided (e.g., vocal attributes) or eliminated (e.g., fingerprints) [5].

When investigating a crime, it is usually required to record some human characteristics of the suspects, such as height, weight, age, gender, etc. [3]. When a crime occurs, the criminal investigation team must examine the entire site and keep the found data on storage devices; consequently, the amount of data stored is enormous. The value of the data lies in the amount of information obtained, which will later aid in decision-making [4, 6]. This needs the evolution of more efficient techniques for analyzing evidence in an investigation. If manual analysis of this data is used, we cannot guarantee the validity or accuracy of the results obtained; therefore, there is a necessity to automate the forensic procedure for precise and rapid outcomes [4, 9]. Criminal investigators implement various machine learning algorithms to analyze images or collect data to determine soft biometric features such as age, gender, and more [3, 9].

A lot of academics are attentive to using machine learning techniques for varied forms of data fusion. However, there is no framework to aid them in their research. Choosing a suitable machine learning algorithm, along with applying preprocessing on the dataset, affects the results of any machine learning algorithm. Thus, datasets must be sufficiently robust to produce the best outcomes [9]. The skill to name the gender from their handwritten text carries weighty potential in DF. At the crime scene, handwritten text can be found in different forms of digital evidence, including scanned documents, notes, and signatures. This research intends to survey the methodologies to distinguish the gender of a human that wrote the evidence based on their handwriting and take part in forensic investigations by providing more profiling tools.

There are three limitations faced by the applications of DF that used handwriting-based gender prediction: (1) Existing methods focus exclusively on prediction accuracy while neglecting evidence integrity requirements essential for legal proceedings, (2) no unified framework addresses both cryptographic verification and gender classification simultaneously, and (3) current approaches lack the processing speed necessary for real-time forensic investigations. To handle these gaps, this research introduced the first integrated forensic framework that merges SHA-256 cryptographic hashing to ensure evidence integrity with ResNet18-based deep learning for gender prediction. The proposed framework differs from existing approaches that treat evidence security and gender classification as separate concerns; it ensures both forensic compliance and high-accuracy prediction within a single, efficient pipeline.

The combination of handwriting analysis with DF faces several stringent challenges, such as keeping a chain of custody through integrity evidence verification, achieving classification accuracy suitable for legal standards, and ensuring processing efficiency for time-sensitive investigations. Furthermore, existing gender prediction systems operate independently of evidence security protocols, creating a fundamental disconnect between technical capability and forensic requirements. The main contributions of this research paper are:

- Enhancing the field of digital forensics by developing a reliable method for gender identification from handwritten text. Evidence authenticity and gender prediction are the two crucial aspects of digital forensics. The combination of these aspects is a challenging task. So, this is the novelty of this paper.
- A comprehensive framework is introduced in this paper, first using SHA-256 hashing to ensure evidence integrity, followed by utilizing the training model of a convolutional neural network (CNN) that is called the ResNet18 architecture for gender prediction. The integration we will implement in this paper not only aids gender classification capabilities but also preserves the required chain of custody for forensic evidence.
- Demonstrate how to integrate advanced deep learning techniques into forensic workflows while maintaining strict constraints on evidence managing and verification.
- Enhancing criminal investigations via automatically analyzing the handwritten notes while adhering to stringent digital forensics standards. A robust gender classification model evolved, benefitting from the Qatar University Writer Identification dataset (QUWI), which holds 1,017 handwritten samples in various languages and scripts.

The paper is structured as follows: A survey of the recent studies in this direction is presented in Section 2. Section 3 introduces gender prediction techniques that are used in DF using handwritten features. An outline of the DF process is illustrated in Section 4. Section 5 shows the details of the proposed dual-stage investigation methodology combining machine learning with cryptographic verification. Experimental results and comparative analysis are presented in Sections 6 and 7, respectively. The last section presents the conclusion of this paper.

2. LITERATURE REVIEW

Recent studies have focused on extracting specific patterns from handwriting and using them to find certain human characteristics, including gender, as follows: Dargan et al. [10] in 2024, introduced gender prediction from offline handwriting in Gurumukhi script. They accomplished classification accuracy that surpasses benchmark performance, showing potential for future deep learning applications. They used dataset consisting of 35,000 characters each from male and female writers. Dataset preprocessed to remove noise, normalization, and bitmap image creation. The authors used more than one gender classification models, such as K-nearest neighbor (KNN), decision trees (DT), and random forests (RF).

In 2024, Hasan et al. [11] examine the efficacy of various methods for arabic handwritten numbers recognition (AHNR), focusing on the recognition of handwritten arabic numbers using CNNs, local binary patterns (LBP), and histogram of oriented gradients (HOG). The authors used CNNs for feature extraction and classification as well as LBP and HOG to extract a variety of features with KNN as a classifier. The CNN model achieved nearly 99% accuracy, better than the other models. Additionally, the CNN model has proven superior computational efficiency, being 0.61 seconds faster than the HOG method.

Agduk and Aydemic [12] in 2022 predicted the gender of a person based on handwritten text signatures utilizing transfer learning methods. It evaluates the effectiveness of 32 transfer learning algorithms and 28 classification methods for this task. DenseNet169 (most effective), DenseNet201, ResNet50,

ResNet101, EfficientNet variations, AlexNet, VGGNet, among others are the algorithms used. They used classification algorithms: Random Forest, Linear Discriminant Analysis (LDA), Logistic Regression (LR), Support Vector Machines (SVM), Gradient Boosting Classifier, Neural Networks, and others. The highest accuracy of 92.46% was conducted by DenseNet169 for feature extraction and Linear Discriminant Analysis for classification. An accuracy of 92.77% was achieved using DenseNet169 for feature extraction and the Hist Gradient Boosting Classifier for classification. Table 1 surveys related works that aim to recognize gender from handwritten.

Table 1. Survey of handwriting-based gender prediction studies

Study	Features Extracted	Name of Classifier	Datasets Used	Accuracy Obtained	Security Used
Dargan et al. [10]	The features extracted involve: Document Examiner Features, Computational Features, Macro and Micro Features, Geometrical Features, and Transformed Features.	KNN, DT. RF, Adaptive Boosting	Digital images of handwritten Gurumukhi script	94.6% for using the hybridization of features and Adaptive Boosting classifiers.	Non
Hasan et al. [11]	It used CNNs for feature extraction and classification. LBP is employed as a feature extraction technique, and HOG is used for extracting shape and structure-based features.	KNN, CNN	A dataset contains 70,000 images	98.95% for CNN, 95.7% for KNN with LBP.	Non
Agduk and Aydemir [12]	using transfer learning methods and classification algorithms to define gender attributes.	HIST Gradient Boosting	The dataset includes handwriting samples	92.77% for Hist Gradient Boosting Classifier.	Non
Rabaev et al. [13]	A features are extracted based on data-driven, mutual information, PCA-based feature extraction, and textural features are also extracted. Additionally, features such as handwriting slant angle, stroke order, and pen pressure are also extracted. The study employed a native feature extraction technique based on Bilinear (B-CNNs) combined with ResNet blocks for gender classification tasks.	Adaptive Multi-Gradient (AMG) and Multi-Gradient Directional (MGD) features Bilinear CNN	KHATT, QUWI, and HHD	99.29% of using the B-ResNet on the HHD Dataset. 76.17% of using the KHATT dataset. 88.33% for the QUWI dataset (English handwriting), 85.23% for (Arabic handwriting) and 78.04% for (English and Arabic handwriting). 84% for the HHD Dataset.	Non
Al-Harbi [14]	The features used in the study include: N-grams: Specifically, unigrams and bi-grams.	SVM	Dataset for Jordanian dialect	92.42% using the SVM classifier.	Non
Morera et al. [15]	The features extracted from images are: Shape Features, Statistical Features, with applying geometric transformations for data augmentation.	CNN	IAM and KHATT	83.19% for IAM dataset. 68.90% for KHATT dataset.	Non
Mirza et al. [16]	The features include the mean and variance of images filtered by a Gabor filter bank, and the Fourier transforms of these matrices are then utilized.	Feedforward ANN	QUWI	70% for the Arabic samples only. 67% for English samples only.	Non

3. HANDWRITTEN-BASED GENDER PREDICTION FOR DIGITAL FORENSICS

Biometric systems use one or more of a person's vital characteristics for the authentication process. These systems are divided into physiological and behavioral systems. Physiological characteristics are linked to the characteristics of the individual, such as fingerprints, irises, retinas, DNA, hands, faces, and voices. while the behavioral biometric systems encompass attributes such as keystrokes and voice.

Biometric data, encompassing physiological and behavioral information, is often employed to recognize and find criminal suspects. The utilization of biometric characteristics for identification verification in forensic investigation belongs to the early 20th century [5]. A handwritten signature is a biometric measure called soft biometrics, which has a long history and is used for predicting many traits, such as the gender of the writer. Predictive capabilities instituted on handwriting open vast horizons for predicting many traits [9].

Gender prediction can be an important aspect of forensic investigations, especially when the crime scene contains samples of handwritten papers that can be used to identify the writer, as it can provide additional information for identifying a suspect or victim [10, 12, 15]. Gender can be identified from handwriting samples, which is essential in situations such as personal and gender identity that may be needed in deciding the author of a document found at a crime scene [12, 17]. For many years, handwriting has been used by document

examiners, forensic analysts, and in psychoanalysis. Handwriting-based gender prediction analysis can serve as a valuable measure in forensic inquiries because of its ability to recognize the gender of a handwritten document's author, which can assist in refining the suspect pool in criminal investigations and may serve as evidence in legal activities [16].

Offline and online recognition are types of recognition processes. The first one uses the text that is found on paper or another physical medium and is then digitized, while the second deals with text generated by an electronic digitizer device [15]. Offline handwriting text analysis entails the classification of handwriting into distinct categories, including age, gender, and nationality. This process involves the extraction of features from handwriting, such as textural characteristics, letter shape, size, curvature, pen pressure, and text inclination, and the application of algorithms through various techniques to recognize the characters and words present in the text. Offline handwritten text analysis entails the transformation of handwritten characters into machine-encoded text, employing techniques such as pattern recognition, machine learning, and artificial intelligence algorithms [15]. In this context, machine learning (ML) is considered a solution. The main attractive features of it are that it streamlines the analytical transaction and provide accurate results [4, 9].

ML and artificial intelligence (AI) are the prominent technologies of the contemporary era. All processes are progressing towards automation. DF must advance concurrently with the digital world to enhance data processing capabilities. Deep learning and deep neural network methods are employed for analysis of text and digital files and are also used with multiple types of forensics applications that include linguistic, image, video, and email. It is utilized in timeline reconstruction, pattern identification, and object categorization [18, 19].

In recent years, ML has gained major interest. It is a method acquired via training and experience rather than through programming. The dataset must be trained before the application of these algorithms. The selection of algorithms is another crucial part in machine learning. Depending on the need for supervision, ML algorithms can be classified into two types: supervised learning and unsupervised learning. In the first type, it is necessary to train the algorithm on a labeled dataset for examples, SVM, DT, RF, CNN, LR, and Naïve Bayes (NB). In unsupervised learning, it is necessary to train the classifier on an unlabeled dataset. For instance, clustering and Principal Component Analysis (PCA). Another category is deep learning, which relies on unsupervised learning. It employs hierarchical structures of artificial neural networks (ANNs) [4, 6, 20]. Data analysis is conducted primarily in two phases. The initial phase involves preprocessing, during which we must input our dataset into the machine learning algorithm. If the data is gathered personally, it must first be trained, while it is filtered based on requirements if it was pre-trained. After training the dataset, the ML algorithm will be used for classification purposes. The results of this step will be considered as input for the final stage of digital forensics [9].

All stages of identification of the person and gender are illustrated in Figure 1, starting with the data collection step and ending with the classification step, passing through preprocessing, feature extraction steps. The quality of the collected data is poor, so it needs to be improved, which is done in the preprocessing step that prepares the data for the feature extraction stage. This step includes various operations such as segmentation, binarization, normalization, and noise removal (particularly applied for image or signal) [21, 22].

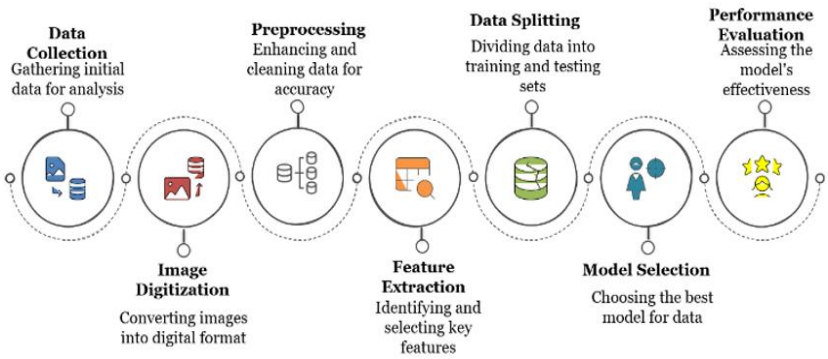

Figure 1. Gender identification system process

To ease extracting the features from handwritten document images, the image can be split into parts of writing (blocks, paragraphs, lines, words, and characters).

In the feature extraction stage, some techniques are applied to the processed data to generate valuable features that are used in the next stage to train or learn the classification algorithm.

The classification stage included implementing algorithms that take the feature set as input and produce the identification of gender/person as a result. Examples of classification algorithms are DT, LR, KNN, SVM, RF, and deep neural networks. The first phases (preprocessing and feature extraction) that are applied to the training data will also be used for the testing data, which includes new handwriting samples. Finally, the classifier finds the gender depending on the extracted features [23].

Traditional person/gender recognition methods usually depend on feature engineering and directly extracted features from handwriting images and then classify using a classical classifier. All existing works focused on offline handwriting-based individual/gender recognition with various feature extraction methods—codebook, structural features, texture features, and structural and textural features. For handwriting characterization, the RootSIFT descriptor and gaussian mixture models (GMMs) were applied on handwritten images, while cloud of line distribution (COLD) and hinge features were extracted from handwritten images for gender identification [23].

Deep learning employs multiple hidden layers in the architecture of an ANN to form a deep neural network (DNN) capable of autonomously learning data and patterns, hence making informed decisions. It has bought significant popularity among researchers and practitioners compared to earlier approaches due to its generalization capabilities and ability to yield more accurate findings when trained on extensive datasets [24].

In the case of small data size, transfer learning and data augmentation approaches are used to aid deep learning work. To ingest the information from the global context, a residual recurrent neural network (GR-RNN) was used. To compute feature vectors and obtain writer identification results, various techniques can be employed, such as baseline, fragment-only, fragment-RNN, fragment-global-context-RNN, and fragment-global-context-residual-RNN [23].

4. DIGITAL FORENSICS PROCESS

DF was called computer forensics at its rise and dates to the late 1990s and early 2000s. The versatility of digital forensics is demonstrated by its application in politics, law enforcement, the corporate sector, the government, and education. Locating and gathering digital evidence from available sources is the first step in every digital crime investigation procedure. Digital evidence is any notable data requisite to set up criminal behavior's evidence. This data from digital artifacts can be analyzed and admitted in court [25]. The digital forensics investigation technique differs based on device kind and environment, leading to many branches, including memory forensics and mobile forensics. DF pertains to the identification, documentation, and response to security breaches. This regards the acquisition, analysis, and reporting of digital evidence, utilizing different technical skills to expose signs of cybercrime. Performing a comprehensive and precise data analysis can be extremely daunting and challenging because of the multitude of tools, each tailored for a unique function, scattered across several platforms.

To ensure that the right tool is utilized for each stage of the investigation, a detailed understanding of their application in the distinct DF branches and phases of the forensic lifecycle is required [25]. The domain of digital forensics is in significant demand owing to the persistent concerns of data breaches and information hacking. Computer forensics covers many areas: identification, preservation, analysis, documentation, and presentation of digital evidence. To avoid compromising the evidence, specialized methodologies and techniques are executed for extracting the information from digital media with precision. The final information obtained can indicate the offender's purpose, the methodologies employed in the commission of the crime, and the resultant detrimental outcome [1]. DF investigation (DFI) encompasses several stages, starting with search, identification, collection, and finally, producing a report related to the digital evidence. The investigators are doing several procedures to enhance DFI. For instance, Thakar et al. introduced a framework known as the Next Generation Digital Forensic Investigation Model (NGDFIM) to aid investigators during the examination process. The framework makes up three phases [18, 19]:

1. **On-site processing phase:** In this stage, the crime teams collect any evidence at the crime site, such as notes, photographs, sketches, and videos. The physical and biological evidence also must be collected and preserved. Examples of physical evidence are fingerprints, footprints, and tire, or shoe impressions. While DNA, other bodily fluids, hair, skin, and bone material are examples of biological evidence [1].
2. **Analysis:** with activities concerning the preparation of tools and equipment that might be useful in dealing with incident evidence. The analysis phase helps in turning data found on a suspect's device into evidence that can be used in court. Tools like The Sleuth Kit and Volatility provide multiple features that help an investigator during the analysis phase of the investigation.
3. **Presentation:** creation of a report with findings where the evidence will be presented as proof that a crime was committed and will prove the identification of the criminal [1].

Every step has its importance, but the crux of forensics lies in the analysis phase. With the increase in data volumes, analysis becomes difficult, and it is almost impossible to draw error-free results [9]. The above framework is more efficient, reducing time for imaging and analysis, saving the privacy of the suspect, and improving the overall process of digital investigation, so we depend on it through this research.

5. METHODOLOGY

In this research, we propose a two-stage DF system that combines evidence verification and gender prediction based on handwritten documents. SHA-256 hashing is used in the first stage to verify evidence. Using a hash provides a cryptographic seal, which is useful for verifying the authenticity of handwritten documents during the investigation process. The second stage of the proposed system encompasses using a model of CNNs named ResNet 18 for gender prediction. This model has attractive features that make us use it, one of which is overseeing complex images during classification tasks, and its ability to mitigate vanishing gradient problems through residual connections. The ResNet18 structure consists of 18 layers with skip connections that allow for gradient flow, making DNN training more efficient with stable performance. Through the results that will be explained later, we demonstrate the effectiveness of the proposed model for the task of handwriting analysis for gender prediction. The structure of our comprehensive system is illustrated in Figure 2, which details the system's stages, from initial document processing through validation to gender prediction. The introduced system guarantees the security and accuracy of DF while supporting the standards required for forensic evidence processing.

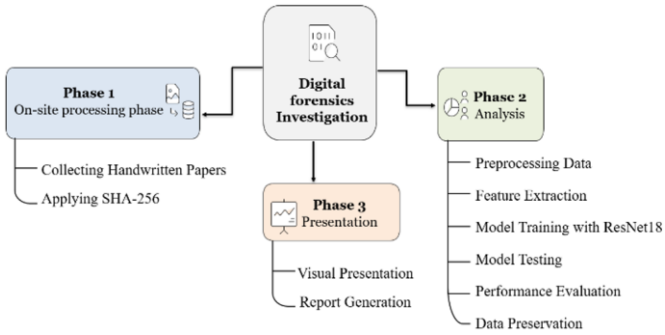

Figure 2. Proposed system architecture

5.1 Securing evidence

In recent years, with the advancement in technology, the need for the authentication of digital images has increased, because many tools are arising that enable unauthorized people to modify the images and republish them [26]. The integrity of collected evidence must be ensured during all investigation phases, from collection until investigation completion. The way to do this is by using hashing [27]. A mathematical hash function is a cryptographic function that accepts input messages of arbitrary size and outputs a fixed-size result known as the hash value or digest. The output is usually a combination of numbers and letters that function as a distinct digital "fingerprint" for the input information. The output of the hash function must be deterministic, which means that for a given input, the output will always be the same [27]. The main powerful feature of the hash function is that the generated hash value cannot be used to retrieve the original message, so privacy and security are preserved while sharing the message. For this feature, the hash is effective for password storage and digital signature [28, 29]. There are various hash functions, like message digest (MD2, MD4, and MD5), hashed message authentication code (HMAC), and secure hash algorithm (SHA-0, SHA1, SHA2, and SHA-3). Each one has its features, but SHA-256 is highly secure and widely used in DF. This paper suggests SHA256 to legitimate DF images. Hashing can be used at distinct stages of DF. Hashing can be used at the evidence collection stage, where it is used at the time of acquiring digital evidence, such as it is used for (original handwriting image files, extracted features (e.g., HOG features or CNN embeddings), or prediction results (gender labels) [27].

The SHA-256 processing is as follows:

Padding, in this step, we arrange the final input message into the following, as Figure 3 below explains. This step is called the padding step. Then the padded message is broken into 512-bit blocks [30].

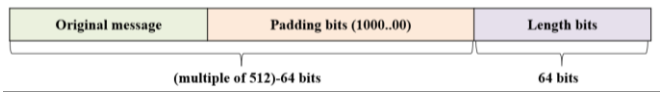

Figure 3. Padding

In SHA-256, the message is split into blocks, each one have 512 bits, $B^{(0)}, B^{(1)}, \ldots, B^{(N-1)}$, which are then processed in 64 rounds on each block of data $B^{(i)}$ utilizing a variety of bitwise operations and nonlinear functions, such as logical functions (AND, OR, XOR) and arithmetic functions (ADD, SHIFT). A hash value H_N with a partial 256-bit is gained as the total processing of the current $B^{(i)}$ block is done, and the last hash H_N is computed and delivered after the last data block $B^{(N-1)}$ is computed. During each round, the block is partitioned into sixteen words, each one has 32 bits, and these words are used to update a set of eight 32-bit variables, known as the working variables [26, 27, 29]. Figure 4 shows the steps of the SHA-256 hash function algorithm.

where,

$$Ch(x,y,z) = (x \wedge y) \oplus (\neg x \wedge z) \tag{1}$$

$$Maj(x,y,z) = (x \wedge y) \oplus (x \wedge z) \oplus (y \wedge z) \tag{2}$$

$$\sum_{0} (x) = ROTR^{2}(x) \oplus ROTR^{13}(x) \oplus ROTR^{22}(x) \tag{3}$$

$$\sum_{1} (x) = ROTR^{6}(x) \oplus ROTR^{11}(x) \oplus ROTR^{25}(x) \tag{4}$$

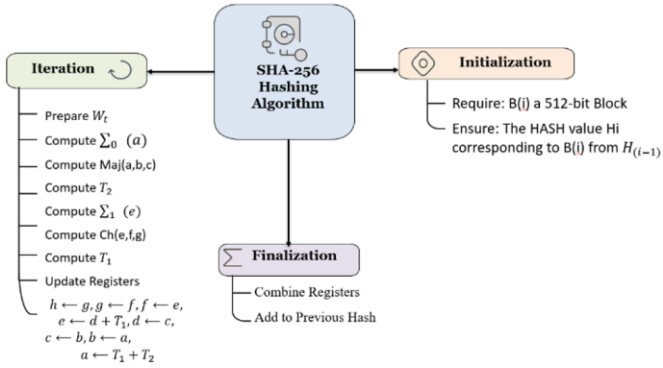

Figure 4. SHA-256 hashing algorithm

To process the $B^{(i)}$ data block, the main function uses eight working variables: a, b, c, d, e, f, g, and h, each of which is 32 bits. After the 64 loops, these variables comprise a buffer state, which is the partial hash value H_{temp}. The result H_{temp} is summed with the hash of the previous data block that was previously computed to obtain the hash value corresponding to the message until the data block $B^{(i)}$. When the first data block is processed, a first hash value H_0 is used. At the beginning of the computation $H_{(i+1)}$, the H_i value needs to be loaded into the buffer state. At every iteration in the hash algorithm, a constant K_t is desired. This is constantly well known in the SHA-256 specification. The equations below compute the temporal variables T_1 and T_2.

$$T_1 = h + \sum_{1}^{256} (e) + ch(e,f,g) + K_t^{256} + W_t \tag{5}$$

$$T_2 = \sum_{0}^{256} (a) + Maj(a,b,c) \tag{6}$$

A 32-bit value W_t is calculated for each iteration in the algorithm above as explained during the first 16 iterations, the first 16 values W_t are taken directly from the 512-bit message, and the remaining iterations σ_0 and σ_1 are used to compute the values of W_t.

$$W_t = \{M_t^{(i)}\ 0 \le t \le 15\ \sigma_1^{\{256\}}(W_{t-2}) + W_{t-7} + \sigma_1^{\{256\}}(W_{t-15}) + W_{t-16}\} \quad 16 \le t \le 63 \tag{7}$$

$$\sigma_0(x) = ROTR^{7}(x) \oplus ROTR^{18}(x) \oplus SHR^{3}(x) \tag{8}$$

$$\sigma_1(x) = ROTR^{17}(x) \oplus ROTR^{19}(x) \oplus SHR^{10}(x) \tag{9}$$

In our paper, after we get the evidence (images, notes, etc.), we compute the hash value using SHA-256. We can ensure

that the evidence is not altered if we get the same hash value for the processed evidence, so its integrity is preserved.

5.2 Prediction with deep learning

This section describes the details of deep learning models that are used for handwriting evidence-based gender prediction as follows:

5.2.1 Data acquisition

To assess the performance of writer identification systems, databases available online can be used. One such public database used for this purpose is QUWI, which consists of handwritten documents in both Arabic and English. 1,017 individuals of varying ages, genders, and educational levels were collected in this database. These volunteers were asked to write a distinct text and a random text of their choice in both Arabic and English. This makes the data suitable for use in text-based writer identification tasks. The most noteworthy features of these databases, which led us to use them in our proposed system, are the balanced distribution of males and females (48% and 52%, respectively) and the holding of both Arabic and English texts, enabling them to be used in various systems.

5.2.2 Preprocessing

To ensure good model results, it is important to focus on the quality of the inputs. To achieve high quality, the preprocessing of handwritten documents is required. This preprocessing plays a significant role in the gender recognition process. In this paper, we applied several preprocessing procedures to the documents to gain a high-quality input for the ResNet18 model.

Image Standardization: Dataset images are high-resolution images (600 dots Per Inch), so they will first be converted to a standard format. To keep colors, all images are converted to the RGB color space, which indicates pen pressure and writing style. ResNet18 accepts images with (224 x 224) pixels, so the images must be resized to the desired size while keeping the aspect ratio. To normalize the pixel values, we divide them by 255, and thus all values will be in the range [0,1]. Finally, the normalization was followed by standardization using the ImageNet mean and standard deviation.

Evidence quality enhancements: We'll be making several improvements to the document to gain high-quality features, such as using the histogram equalization technique for promoting contrast and using a 3×3 kernel filter to eliminate the Gaussian noise. As well as employ bilateral filtering to keep the edges that support the text's font characteristics.

Text Region Extraction: This step involves separating background from text using Otsu's threshold for initial separation. Contiguous text regions can be identified by analyzing connected components. This step also applies to morphological operations to connect nearby characters while preserving the integrity of the fonts.

Evidence Augmentation: One of the common problems is overfitting, which can be solved in this step, thus increasing the robustness of ResNet-18 by implementing various augmentation mechanisms to the evidence, such as random rotation and random gradients. The first one deals with the variance in the writing angle, whereas the second deals with the inconsistency of font size. Finally, the brightness and contrast must be adjusted because it is affected by writing tools and scanning circumstances.

Data Systemizing: This step involves creating small batches of equally distributed male and female samples. To process large-sized documents, a sliding window approach is utilized. Training, validation, and selection groups are prepared with stratified sampling while maintaining gender distribution.

These steps will be applied to process Arabic and English texts to ensure proper processing while preserving the text properties.

5.2.3 Feature extraction and model training

This stage of the proposed model is extracting the features from the preprocessed evidence to use later in training the model. Transfer learning is used for this demonstration in conjunction with the pre-trained ResNet18 model. We can obtain leveraging weights after training the model on a large-scale dataset. Low-level visual features like edges and textures, and high-level features like shapes and patterns, can be obtained from this stage. These features can be used for gender classification based on handwriting. The ResNet18 architecture processes handwritten documents through five components, as shown in Figure 5. This architecture was designed for use in handwriting analysis yet maintains its computational efficiency for other forensic applications.

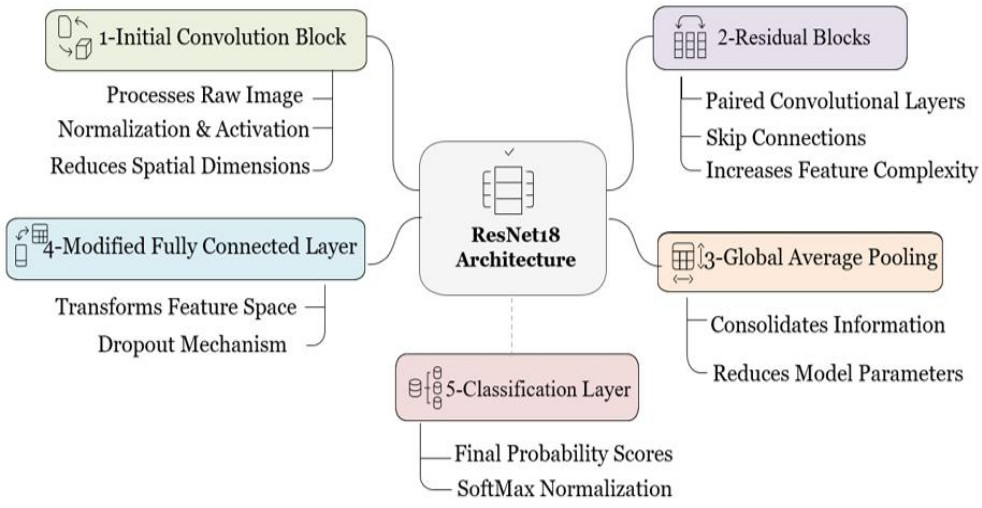

Figure 5. ResNet18 architecture for handwritten document processing

The transfer learning process follows three critical stages: Model Initialization (Algorithm 1), Training Protocol (Algorithm 2), and Feature Adaptation Process (Algorithm 3).

Algorithm 1: Model Elaboration

```
Input: Pre-trained ResNet18 model
Output: improved model for gender prediction
Start
   Load pre-trained ResNet18 weights.
   keep all convolutional layers.
   Swap the final classification layer with:
     - Dense layer consists of 512 neurons
     - Activation function (such as ReLU)
     - Dropout layer (0.5 probability)
     - Output layer (2 neurons)
   set class weights [1.2 for female, 1.0 for male]
End
```

Algorithm 2: Model Training

```
Input: Training data, Validation data, Model
Output: Trained model
Start
   FOR epoch = 1 to MAX_EPOCHS:
     For each batch in the training data:
       Forward pass:
         Offprint features over frozen layers.
         Produce predictions over new layers.
         Compute weight loss.
       Backward pass:
         Update unfrozen layer weights.
         Stratify gradient clipping.
       Adjust the learning rate based on validation performance.
     Assess the model on the validation set.
     keep the best-performing model.
   End
End
```

Algorithm 3: Feature Adaptation

```
Input: Trained model, Fine-tuning data
Output: Fine-tuned model
Start
   Unfreeze deeper layers progressively
   FOR each unfrozen layer:
     Predetermined a smaller learning rate.
     Fine-tune on training data.
     Observe validation performance.
     IF performance plateaus:
       Adjust the learning rate.
   Validate on the test set.
End
```

5.3 Constraints affecting methodology and system performance

Despite an integrated framework being proposed to enhance forensic handwriting analysis, built on combining two significant approaches (SHA-256 hashing and deep learning via ResNet18), some computational and methodological limitations have appeared that affect scalability and practical deployment. For resource-limited forensic labs, these approaches pose big challenges due to the SHA-256 algorithm requiring memory buffering that accumulates during high-volume processing, and ResNet18 demands substantial computational power and memory. Large-scale investigations require a large volume of memory and more time because the system requirements are scaled with document volume. The constraints of the ResNet18 algorithm include fixed input sizes, which require an input image with a fixed size that may cause the loss of some information from the handwritten image. While the constraints of SHA-256 belong to its deterministic feature, where it gives the same result for the same input, the system can not know the degree of the change in the documents, where changing one character gives a different result. These constraints limit the flexibility and explainability, which is critical in legal contexts. Using one database may add less generalization to the system because of limited demographic and script variety. For legal admissibility, it is difficult to accept the model in criminal investigations or as part of evidence, so applying this system in the criminal field requires a high level of transparency and accuracy, which are still beyond the capabilities of the current system. Overcoming obstacles includes GPU acceleration, memory optimization, and using fallback verification mechanisms with suggesting the use of lightweight models, incremental hashing, expanded datasets, and explainable AI to enhance performance and applicability.

Despite these limitations, the framework indicates strong potential for secure and automated handwriting-based gender identification in modern forensic settings, which is explained in the next section.

6. EXPERIMENT RESULTS AND DISCUSSION

Our proposed system exhibits high security strength in terms of evidence verification, in addition to its outstanding performance in gender discrimination, as demonstrated by the experimental evaluation of the system. To show this, we will present a comprehensive analysis of the security and gender discrimination aspects in this section.

6.1 Evidence security analysis

6.1.1 Hash function performance evaluation

After applying the 256-bit hash to the QUWL database, we found that the average time required to compute a hash per document was 0.23 seconds, which transcends the standard requirement of 0.50 seconds, making the proposed model suitable for real-time forensic applications. Table 2 illustrates the high performance of SHA-256. Further demonstrating the cryptographic strength of the algorithm is the achievement of zero collisions across all 4,068 documents. The system maintains memory efficiency during the reproduction process, where it generates 256 bits per hash, regardless of document size.

Table 2. SHA-256 performance analysis on QUWI dataset documents

Performance Metric	Result	Benchmark Requirement
Average Hash Generation Time	0.23 seconds	< 0.5 seconds
Hash Collision Rate	0% (0/4,068 docs)	0%
Bit Change Detection	100% (256-bit differences)	> 50%-bit difference
Reproducibility Rate	100%	100%
Memory Usage	256 bits per hash	256 bits
Document Size Range	100KB - 15MB	-
Verification Success Rate	100%	100%

The system can be scaled to be applied across a variety of documents used in criminal investigations. where it processed documents ranging in size from 100 KB to 15 MB. This feature makes the system valuable for forensic applications, due to the speed and reliable efficiency being two critical requirements.

6.1.2 Security properties validation

Our system effectively prevents both attempted reverse engineering and malicious document substitution, as demonstrated by the successful verification of the resistance of the pre-image and second pre-image. Changing a single bit in the document generates a completely different hash value, providing strong protection against document alteration. Table 3 displays the SHA-256 security features implemented in our system.

Table 3. SHA-256 security property validation

Metrics	Evaluation Result	Description
Pre-image Resistance	Validated	No successful reverse engineering of the original document from the hash
Second Pre-image Resistance	Validated	No collisions were discovered in the updated documents.
Collision Resistance	Validated	No hash collisions across the whole dataset
Avalanche Effect	> 50%	Single-bit changes result in significantly different hashes
Processing Speed	Linear	Hash time proportional to document size

6.2 Gender prediction performance

Accuracy, precision, recall, and F1-score are metrics utilized to assess the performance of the proposed system. Eqs. (10)-(13) indicates the equations of each metric:

$$Accuracy = \frac{TP + TN}{(TP + FN + TN + FP)} \tag{10}$$

$$Precision = TP/(TP + FP) \tag{11}$$

$$Recall = TP/(TP + FN) \tag{12}$$

$$F - score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{13}$$

where, TP is True Positive, TN is True Negative, FN is False Negative, FP is False Positive.

Table 4. Detailed model performance metrics

Metric	Female	Male	Overall
Precision	0.93	1.00	0.97
Recall	1.00	0.96	0.97
F1-Score	0.96	0.98	0.97
Accuracy	0.97	0.97	0.97

Our proposed model was able to predict gender with excellent ability, as shown beyond doubt in Table 4. The metrics were highly balanced across categories, referring to strong generalization capabilities. As the notation of the calculated precision values that are 0.93 for females and 1.00 for males, we can say the system accomplishes exceptionally reliable positive predictions. The recall score was mostly good. The model was capable to recognize all female writings correctly with a score of 1.00, compared to the high recall score of 0.96 for the male samples.

Precision related to positive predictions. It indicates how many of the samples that are were actually positive and it was also identified as positive by the model. As noticable the precision values for for females and males were 0.93 and 1.00, respectively that is show remarkably dependable positive predictions in our gender prediction system. always the system was accurate when predicts the male authorship, it removing false male classifications completely, according to the perfect male precision (1.00). In forensic applications, high precision is fundemental for investigative credibility . The investigative truthfulness is related to the precision value. As shown results, a precision value was 0.93, which means the system identified 93% of documents as female writer, while only 7% were false positives. Increasing false positive rates has an effect on the accuracy of indicating the suspects in criminal investigations, as it may misallocate investigative resources or jeopardise the integrity of the case.

Recall measure is related to the relevant instances. It indicates how many of the actual positives were correctly identified.The recall values were 1.00 for females and 0.96 for males. Our system demonstrates remarkable detection capabilities in both male and female groups. The perfect female recall (1.00) ensures the system recognises all the documents written by woman in the dataset (i.e., no female samples are misclassified as male). For a male recall of 0.96, 96% of male served documents are correctly classified as male, while 4% are misclassified as female. High recall is important in applications such as question-answering, text categorization and retrieval systems when it is vital to retrieve all or most of the relevant documents (e.g., archived emails, legal documents, etc.).

In information retrieval, relevance is a measure of how good it is that an item meets the information need of the user. High recall is only one measure of performance. Low recall is often associated with false negatives. While this often occurs in academic information retrieval in different test databases, there is growing concern that this may happen in real-life retrieval situations. In forensic identification, missing a suspect also known as a false negative or a miss can lead to conducting a further investigation or it could cause misidentification, in which an innocent subject may be labeled as a suspect. The model may be especially good at identifying female handwriting traits because of the higher female recall than male recall. This could be because female writing patterns have unique characteristics that the ResNet18 architecture can recognise more readily.

F1-score metric providing a balanced performance where it takes into consideration both false positives and false negatives, so it offers a harmonic mean between precision and recall. With a total F1-score of 0.97, our system obtains F1-scores of 0.96 for females and 0.98 for males. These values show a superb balance between recall and precision as mention before, for example the precision value of males is perfect making up for slightly lower recall value, as appered by the slightly higher F1-score for males (0.98) compared to females (0.96).

Because both false positives and false negatives have serious effect in criminal investigations, balanced

performance is crucial for trustworthy forensic analysis. The main classification performance trends are shown by the confusion matrix. Although the system generates 32 false positive female classifications, it exhibits perfect sensitivity for female detection (no false negatives). This asymmetric error pattern implies that some samples of male handwriting have traits that the model identifies with the writing styles of women. Forensic practitioners who must interpret model predictions in investigative contexts must comprehend these misclassification patterns. Table 5 provides deteils of confusion matrix. Table 5 shows the detiels of confusion matrix.

Table 5. Comprehensive analysis of the confusion matrix

Actual Gender	Predicted Female	Predicted Male	Total	Class Metrics
Female	468 (TP)	0 (FN)	468	Recall: 1.00
Male	32 (FP)	487 (TN)	519	Recall: 0.96
Total	500	487	987	-
Class Metrics	Precision: 0.93	Precision: 1.00	-	Accuracy: 0.97

The integrated proposed model depends on a coordinated and stable learning model, as shown in Table 6. The model could learn efficiently without overfitting, where the training loss reduced from 1.1127 to 0.2967 over 20 epochs, along with improvements in validation metrics. The final accuracy that the system conducted is 0.9689 which close to the training accuracy of 0.9730. This makes the model highly generalizable.

Table 6. Training progression metrics

Epoch	Training Loss	Validation Loss	Training Accuracy	Validation Accuracy
1	1.1127	1.0233	0.6523	0.6789
5	0.6814	0.6532	0.8234	0.8456
10	0.5433	0.5234	0.8967	0.9123
15	0.4646	0.4234	0.9345	0.9456
20	0.2967	0.2789	0.9730	0.9689

The classification performance is evaluated at various decision thresholds by Analysis of the Receiver Operating Characteristic (ROC). Our system is capable of predicating male and female handwriting samples by Area Under the Curve (AUC) of 0.987. The system keep high true positive rates while decreasing false positive rates through a range of threshold settings, as shown in Table 7.

Table 7. Metrics of performance at various confidence thresholds

Thresholds	Accuracy	Precision	Recall	F1-Score	Coverage
0.50	97.30%	0.965	0.980	0.972	100%
0.70	98.15%	0.978	0.971	0.974	94.2%
0.80	98.67%	0.983	0.965	0.974	87.3%
0.90	99.12%	0.991	0.958	0.974	76.8%
0.95	99.45%	0.997	0.943	0.969	62.4%

Important trade-offs for forensic applications are revealed by this threshold analysis. While the higher confidence thresholds enhance the accuracy and precision, they cause a decrease in coverage, so a small number of documents are predicted automatically. Important trade-offs for forensic applications are revealed by this threshold analysis. While the higher confidence thresholds enhance the accuracy and precision, they cause a decrease in coverage, so a small number of documents are predicted automatically. Based on the case requirements, forensic practitioners can adjust system parameters, where, in dangerous investigations, higher thresholds may be used to ensure maximum reliability at the expense of requiring manual analysis for uncertain cases.

6.3 Forensic properties validation

After comapre our findings with accepted forensic science standards benchmarks, we found, the proposed system achieve what is usually needed for forensic applications. Our system achieves an accuracy of 97.30%, thus surpassing the minimum accuracy threshold that is suggested by the FBI's Technical Working Group for Digital Evidence is 95%. Additionally the balanced precision and recall metrics guarantee dependable performance across both gender categories.

After the overall performance analysis for forensic, some ramifications must be considered: researchers can be completely confident in male predictions while being appropriately cautious with female classifications that have a 7% false positive rate (ideal male precision vs. ideal female recall).

As shown by the strong confidence and high overall performance metrics, the system is suitable for operational deployment in forensic labs. However, the threshold analysis shows that forensic professionals must weigh between coverage against accuracy depending on case-specific specifications. While preliminary investigations may use lower thresholds to maximise the number of automatically processed documents, high-stakes investigations may benefit from higher confidence thresholds that sacrifice some coverage for increased reliability.

Our integrated framework provides forensically appropriate, dependable gender prediction capabilities that meet or surpass established standards for automated classification systems in criminal investigations, according to this multidimensional performance analysis.

7. COMPARATIVE ASSESSMENT

In this section, we compare our proposed system with some previous works that also distinguish gender based on handwriting but using different methods. In terms of accuracy, our system was able to distinguish gender with an accuracy of 97.30%, while the system applied by Dargan et al. [10] accomplished an accuracy of 94.60%, Agduk and Aydemir [12] system achieved an accuracy of 92.77%, and the lowest accuracy was achieved by Morera et al. [15] system of 83.19%. If we compare the systems in terms of processing speed, we find that our system was able to process data in a competitive performance with a speed of up to 0.31 seconds per document, making it appropriate for employees in real-time operations. Table 8 compares the proposed system with prior systems in terms of the method used, accuracy, and processing time.

As mentioned earlier, our approach is based on ResNet18. We compared this model with other models, such as DenseNet169, VGG16, and AlexNet as shown in Table 9. Despite the fewer parameters used in ResNet18 model compared to others, it outperformed them in accuracy, training time, and memory usage. We also know that computing

resources in forensic applications are limited, so our system is well-suited to collaborating with them efficiently.

Table 8. Performance comparison

Study	Method	Accuracy	Processing Time
Proposed approach	ResNet18 + SHA-256	97.30%	0.31s/doc
Dargan et al. [10]	Hybrid Features + AdaBoost	94.60%	0.45s/doc
Agduk and Aydemir [12]	DenseNet169 + HGBC	92.77%	0.52s/doc
Morera et al. [15]	CNN	83.19%	0.38s/doc

Summarizing the meaningful results, we were able to obtain after implementing the system, we find that the system ensures the integrity of evidence from tampering. This is achieved with high efficiency in terms of processing time and the absence of collisions between large numbers of documents. Compared to previously implemented systems, our system has achieved considerable progress by achieving high accuracy, reaching 97%, while consuming minimal computing resources. the system achieved high accuracy

Table 9. Model architecture performance comparison

Architecture	Parameters	Training Time	Accuracy	Memory Usage
ResNet18 (used)	11.7M	4.5 hours	97.30%	45MB
DenseNet169	14.3M	6.2 hours	92.77%	68MB
VGG16	138M	8.3 hours	91.45%	528MB
AlexNet	60M	3.8 hours	88.92%	227MB

8. CONCLUSION

In this research paper, we present a novel approach that improves digital forensics investigations by identifying the gender of handwritten documents and, before that, securing them to prevent tampering. SHA-256 was used to generate hashes for each document to keep its integrity. For gender identification, CNN-based ResNet-18 was used. Our system, based on a combination of these techniques, achieved remarkable results that could develop forensic document analysis. The system was applied to the QUWI database. SHA 256 was able to ensure the integrity of evidence and avoid collisions by generating a unique hash for each document (evidence) in QWLU. This process was performed at a high speed of 0.23 seconds per document. The gender identification accuracy reached 97.30%, with 11.7M parameters surpassing previously implemented systems while consuming minimal computing resources. This efficiency was coupled with a processing speed of 0.31 seconds per document. All these system features make it applicable to real-time digital forensics applications. The future directions of this field of study include expanding the proposed system to a variety of languages, such as Chinese, Cyrillic script, and Japanese languages, which requires collecting 500+ samples per gender per script, as well as investigating transfer learning effectiveness across different writing systems while maintaining cryptographic integrity verification. In the future, the researchers can evolve the multi-modal forensic system by combining handwriting prediction with pressure-sensitive stylus data, voice, and facial recognition systems. Additionally, explainable AI can be applied.

REFERENCE

[1] Saini, M., Kapoor, A.K. (2016). Biometrics in forensic identification: Applications and challenges. Journal of Forensic Medicine, 1(108): 2. https://doi.org/10.4172/2472-1026 .1000108

[2] Salman, A.D., Hasan, E.H. (2023). Survey study of digital forensics: Challenges, applications and tools. In 2023 16th International Conference on Developments in eSystems Engineering (DeSE), Istanbul, Turkiye, pp. 788-793. https://doi.org/10.1109/DeSE60595.2023.10469020

[3] Anda, F., Lillis, D., Le-Khac, N.A., Scanlon, M. (2018). Evaluating automated facial age estimation techniques for digital forensics. In 2018 IEEE Security and Privacy Workshops (SPW), San Francisco, CA, USA, pp. 129-139. https://doi.org/10.1109/SPW.2018.00028

[4] Qadir, S., Noor, B. (2021). Applications of machine learning in digital forensics. In 2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2), Islamabad, Pakistan, pp. 1-8. https://doi.org/10.1109/ICoDT252288.2021.9441543

[5] Shen, C., Xu, H., Wang, H., Guan, X. (2016). Handedness recognition through keystroke-typing behavior in computer forensics analysis. In 2016 IEEE Trustcom/BigDataSE/ISPA, Tianjin, China, pp. 1054-1060. https://doi.org/10.1109/TrustCom.2016.0175

[6] Mohammad, R.M. (2018). A neural network based digital forensics classification. In 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA), Aqaba, Jordan, pp. 1-7. https://doi.org/10.1109/AICCSA.2018.8612868

[7] Rout, A.N.B., Pandit, S., Patil, H., Patil, V. (2020). Detection of gender using digital forensic. International Journal of Computer Trends and Technology, 68(3): 7-13. https://doi.org/10.14445/22312803/IJCTT-V68I3P102

[8] Ulupinar, S., Dogan, S., Akbal, E., Tuncer, T. (2017). The importance of standardization in biometric data for digital forensics. In 2017 International Conference on Computer Science and Engineering (UBMK), Antalya, Turkey, pp. 781-785. https://doi.org/10.1109/UBMK.2017.8093529

[9] Qadir, A.M., Varol, A. (2020). The role of machine learning in digital forensics. In 2020 8th International Symposium on Digital Forensics and Security (ISDFS), Beirut, Lebanon, pp. 1-5.

[10] Dargan, S., Kumar, M., Mittal, A., Kumar, K. (2024). Handwriting-based gender classification using machine learning techniques. Multimedia Tools and Applications, 83(7): 19871-19895. https://doi.org/10.1007/s11042-023-16354-1

[11] Hasan, B.M., Jaber, Z.J., Habeeb, A.A. (2024). Digits recognition for Arabic handwritten through convolutional neural networks, local binary patterns, and histogram of oriented gradients. Baghdad Science Journal, 21(10): 14. https://doi.org/10.21123/bsj.2024.9173

[12] Agduk, S., Aydemir, E. (2022). Classification of handwritten text signatures by person and gender: A

comparative study of transfer learning methods. Acta Informatica Pragensia, 11(3): 324-347. https://doi.org/10.18267/j.aip.197

[13] Rabaev, I., Alkoran, I., Wattad, O., Litvak, M. (2022). Automatic gender and age classification from offline handwriting with bilinear ResNet. Sensors, 22(24): 9650. https://doi.org/10.3390/s22249650

[14] Al-Harbi, O. (2019). A comparative study of feature selection methods for dialectal Arabic sentiment classification using support vector machine. arXiv preprint arXiv:1902.06242. https://doi.org/10.48550/arXiv.1902.06242

[15] Morera, Á., Sánchez, Á., Vélez, J.F., Moreno, A.B. (2018). Gender and handedness prediction from offline handwriting using convolutional neural networks. Complexity, 2018(1): 3891624. https://doi.org/10.1155/2018/3891624

[16] Mirza, A., Moetesum, M., Siddiqi, I., Djeddi, C. (2016). Gender classification from offline handwriting images using textural features. In 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), Shenzhen, China, pp. 395-398. https://doi.org/10.1109/ICFHR.2016.75

[17] Al Maadeed, S., Hassaine, A. (2014). Automatic prediction of age, gender, and nationality in offline handwriting. EURASIP Journal on Image and Video Processing, 2014(1): 1-10. https://doi.org/10.1186/1687-5281-2014-10

[18] Ivanova, M., Stefanov, S. (2023). Digital forensics investigation models: Current state and analysis. In 2023 8th International Conference on Smart and Sustainable Technologies (SpliTech), Split/Bol, Croatia, pp. 1-4. https://doi.org/10.23919/SpliTech58164.2023.10193176

[19] Kaushik, M.S., Kandali, A.B. (2023). Convolutional neural network based digital image forensics using random forest and SVM classifier. In 2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Bengaluru, India, pp. 860-865. https://doi.org/10.1109/IDCIoT56793.2023.10053434

[20] Abdulrazzaq, M.M., Ramaha, N.T., Hameed, A.A., Salman, M., et al. (2024). Consequential advancements of self-supervised learning (SSL) in deep learning contexts. Mathematics, 12(5): 758. https://doi.org/10.3390/math12050758

[21] Alaei, F., Alaei, A. (2022). Handwriting analysis: Applications in person identification and forensic. In Breakthroughs in Digital Biometrics and Forensics, Cham: Springer International Publishing, pp. 147-165. https://doi.org/10.1007/978-3-031-10706-1_7

[22] Hashim, Z., Mohsin, H., Alkhayyat, A. (2024). Offline handwritten signature identification based on hybrid features and proposed deep model. Iraqi Journal for Computer Science and Mathematics, 5(1): 220-236. https://doi.org/10.52866/ijcsm.2420.05.01.016

[23] Tiwari, A., Mehrotra, V., Goel, S., Naman, K., Maurya, S., Agarwal, R. (2021). Developing trends and challenges of digital forensics. In 2021 5th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, pp. 1-5. https://doi.org/10.1109/ISCON52037.2021.9702301

[24] Salman, A.D., Al-Dahhan, R.R. (2025). Ensure privacy-preserving using deep learning. Mesopotamian Journal of CyberSecurity, 5(2): 703-720. https://doi.org/10.58496/MJCS/2025/042

[25] Suvarna, D., Mahesh, K.M., Gupta, M., Gabburi, S., Honnavalli, P., Sapna, V.M. (2024). The development of a digital forensic framework for ease of forensic analysis. In 2024 12th International Symposium on Digital Forensics and Security (ISDFS), San Antonio, TX, USA, pp. 1-5. https://doi.org/10.1109/ISDFS60797.2024.10527263

[26] Breitinger, F., Liu, H., Winter, C., Baier, H., Rybalchenko, A., Steinebach, M. (2013). Towards a process model for hash functions in digital forensics. In International Conference on Digital Forensics and Cyber Crime, Cham: Springer International Publishing, pp. 170-186. https://doi.org/10.1007/978-3-319-14289-0_12

[27] Gilbert, H., Handschuh, H. (2003). Security analysis of SHA-256 and Sisters. International Workshop on Selected Areas in Cryptography, Berlin, Heidelberg: Springer. pp. 175-193. https://doi.org/10.1007/978-3-540-24654-1_13

[28] Salih, R.K., Kashmar, A.H. (2024). Enhancing blockchain security by developing the SHA256 algorithm. Iraqi Journal of Science, 65(10): 5678-5693. https://doi.org/10.24996/ijs.2024.65.10.30

[29] Al Maadeed, S., Ayouby, W., Hassaine, A., Aljaam, J.M. (2012). QUWI: An Arabic and English handwriting dataset for offline writer identification. In 2012 International Conference on Frontiers in Handwriting Recognition, Bari, Italy, pp. 746-751. https://doi.org/10.1109/ICFHR.2012.256

[30] Quist-Aphetsi, K., Senkyire, I.B. (2019). Validating of digital forensic images using SHA-256. In 2019 International Conference on Cyber Security and Internet of Things (ICSIoT), Accra, Ghana, pp. 118-121. https://doi.org/10.1109/ICSIoT47925.2019.00028