Grenze International Journal of Engineering and Technology, June Issue

A Systematic Analytical Review of Machine Learning Models for Forgery Detection and Localization in Image Document

Mayuri P. Bamnote[1] and Kishor K Bhoyar[2]
[1]Research Scholar, Department of Computer Technology, Yashwantrao Chavan College of Engineering, Nagpur, India
[2]Professor, Department of Information Technology, Yashwantrao Chavan College of Engineering, Nagpur, India
Email: mayurigetme2808@gmail.com, kkbhoyar@ycce.edu.in

***Abstract*—The rapid development of image editing tools and the rise in digital images forgery creates obvious threats, which is severe regarding the preservation of visual data integrity. Hence prompt detection and localization techniques are needed in coming time. Although numerous methodologies have been proposed, past reviews suffer from partial fragmentation of analysis, their explorations of cross-dataset generalization can be sparse and there is limited emphasis on hybridizing disentangled features and lightweight models resulting in the absence of a comprehensive taxonomy for image forgery detection approaches. In order to tackle these shortcomings, this paper presents a comprehensive overview and comparative analysis of state-of-the-art forgery detection and localization systems, by putting forward a structured taxonomy according to their architectures, applications and limitations. The review covers multiple methodologies, including transfer learning-based approaches (MobileNetV2), hybrid deep learning frameworks (FBI-Net and DIF-Net), and transformer-enhanced architectures (TransU2-Net), with assessments of accuracy, F1 score, robustness, and scalability.In addition to that, it outlines models working on domain generalization challenges, such as the CFM Framework, AUC > 0.96, along with adversarial robustness issues such as Inspector and MDCF-Net. It will helps the field by bringing everything together under one system to check how well models work, find the best ways to detect certain types of forgery, and create new hybrid and explainable AI models.**

***Index Terms*— Dataset Generalization, Transfer Learning, Transformer Models, Forgery detection, localization.**

I. INTRODUCTION

The problem of determining the authenticity of digital photos is becoming difficult due to the sophisticated editing software and the volume of digital data in place. Image editing including copy-move, splicing, and facial modifications, present an extreme risk to social credibility, legal procedures, and veritable information spread. Digital forensics is because even the human eye does not notice these forgeries most of the time crucial to deal with this problem. The methods of forgery detection and location are modified. The blocks or key points were used in older methods. More recent techniques employ deep learning and fused tools. These are new techniques that are effective on complex forgeries. Nevertheless, deep learning models require substantial computing resources and numerous examples that are labeled.

Grenze ID: 01.GIJET.12.2.56_22

This renders them difficult to apply in life. Furthermore, the majority of approaches are not able to operate with minuscule tricks trying to deceive the system. It also provides a comparison and a taxonomy to enable researchers develop meaningful and comprehensible AI solutions. The research is an extensive summary of forgery detection and localization models. It pays attention to CNNs, transfer learning and transformer architectures. It tests their functionality on benchmark data collections CASIA, Face Forensic++, and MICC. The article provides a critical analysis of the models that are employed to identify and find fake images. It divides the techniques into categories of the primary types of architecture applied e.g. CNN-based methods and transformers-based networks. The research indicates the weaknesses of previous reviews and enumerates the models that could be applied in multimedia forensics, legal authentication, and verifying content on social media. A research and development plan in the future regarding image forgery detection is also outlined in the paper.

II. Review of Existing Methods

Digital image forensics has become a significant subject today since it is more convenient than ever to make counterfeit images. There are tampering, copy-move, and AI-based approaches to fake images, which concern significant concerns regarding privacy, safety, and the veracity of information. Due to this, there have been numerous tools, approaches and designs proposed to identify and find these edits, and each of them is targeting a particular aspect of the job. The current studies comprise the primary findings, designs, applications and disadvantages of the study, which are gathered in this review. Recent advancements in deep learning have made image forgery detection a potent method of extracting features, classifying images, and locating spots of damage. Metaheuristic algorithms have been an important solver of hard optimization problems in engineering, CS, and ML. The PO method performed excellently on 27 of 33 tests on 23 normal problems and 10 of 33 extreme problems in the CEC 2019 set compared to known algorithms like PSO, DE, and SSA.[1].By using guilloche patterns—complex, computer-generated backdrop motifs that are infamously hard to counterfeit—in a unique forgery detection framework, Musab Al Ghadi and colleagues (2022) address identity document authentication. The two main stages of the suggested model are a similarity measurement step that determines whether a queried document matches a legitimate reference from the same nation CatalyzeX+6ar5iv+6arXiv+6 and a Siamese Convolutional Neural Network (SNN) that extracts highly discriminative feature embeddings from pairs of document images. This design makes the framework resistant to minor changes in backdrop designs by specifically optimizing anticollision (i.e., inter-class distinguishability) and conformance to valid guilloche structures.[2]Discrete-frequency banding signals (such as 90, 110, and 130 cycles/inch) could be embedded and then detected using spectral and matched-filter analysis techniques, as the authors demonstrated with an HP Color LaserJet 4500 as an experimental platform. They also estimated a possible embedding capacity of up to 33 bits per page. Banding signatures were added to intrinsic printer identification utilizing gray level co-occurrence texture features to increase resilience against consumable wear effects and document tampering.[3] By using Principal Component Analysis (PCA) to reduce feature dimensionality and Gaussian Mixture Models (GMM) to classify electrophotographic printers based on subtle texture and banding patterns in printed documents, Ali, Mikkilineni, Delp, Allebach, Chiang, and Chiu (2004) expanded the fundamentals of printer forensics through a machine learning-based approach. In order to isolate the major forms of sensor noise and mechanical artifacts, their data-driven pipeline first retrieved gray-level co-occurrence matrix (GLCM) characteristics from printed glyphs and letterblocks before using PCA. The work improved blind source-camera identification methods by concentrating on artifacts caused by algorithms instead of hardware noise. [5] A thorough analysis of image and document forgery detection techniques was provided by Benhamza, Djeffal, and Cheddad (2021), who categorized the main techniques by datasets, evaluation metrics, feature type (e.g., handcrafted vs. deep learning), and forgery kind (e.g., copy-move, splicing). University of Hamad Bin Khalifa +5SciSpace+5 5 bth. diva-portal. org+6diva-portal. org+6diva-portal. org+6 Hamad Bin Khalifa University. They find that deep learningbased models outperform common hand-crafted features in automatically capturing high-level artifacts for the purpose of tamper localization, which otherwise show inconsistent performance across a large variety of manipulation scenarios. They also emphasize the importance of hybrid systems that combine machine-classifiers created for structured document manipulation cases and real-life work processes as forensic signals [6].

An innovative forgery detection approach that investigates intrinsic character-level features derived directly from scanned printed documents without relying on external watermarks or metadata (Bertrand et al., 2013) arXiv+12SciSpace+12DBLP+12. They read known letters using optical character recognition (OCR) for classification and extracted alphabets, checking the discriminative feature vectors as a combination of primary inertia axes, Hu moments, bounding-box parameters and line-level horizontal alignment. To accurately identify

both near-duplicate and outlier character pieces, a two-step outlier detection process is implemented. The first tier employs shape similarity distance metrics, and the second identifies strictly identical characters (like duplicated forgeries or copy-cloned text). Results on real documents show that, while the system can correctly detect forged characters, it can produce false positives when print-scan noise looks like genuine [7]. In Expert Systems, compare text-dependent and text-independent source printer identification techniques for document forgery detection. While the text-independent mode depends on global document-level visual features or patches without character context, the text-dependent mode uses character-specific patches (such as letters or numbers) to train SVM or CNN classifiers based on local textures and printing microfeatures. Their text-dependent approach achieved approximately 98.06% identification accuracy using a dataset of documents printed by 20 different printers, while the text-independent approach achieved 95.52%. This shows that while semantic patch features provide slight improvements, text-independent techniques continue to be very successful. xs.typicalgame.com [8]. The model's concepts are applicable to document analysis when high-dimensional spectral data (such as ink or paper spectra) is available, even though it was created mainly for industrial inspection and remote sensing. On benchmark HSI datasets, the network achieves state-of-the-art performance by combining both local and global spectral characteristics. Such architectures can be modified for use in document forensics to identify manipulation, track changes, or identify ink differences in spectrally scanned official documents, especially when paired with hyperspectral capture devices. Thus, the methodology of Cheng et al. establishes a basis for upcoming deep learning-based high-precision forensic tools on high-dimensional visual data [10].

Choi et al. (2009) used Discrete Wavelet Transform (DWT) features to analyze color laser printer output and present a statistical method for printer source identification. Their study, which was presented at the IEEE ICIP 2009 conference, captures the minute textural features of printed papers that are frequently invisible to the human eye, such as halftone patterns, toner dispersion, and geometric aberrations. Multiresolution analysis is made possible by the DWT-based method, which offers reliable statistical descriptors that may differentiate amongst printers of the same model or make. This method is especially useful in forensic investigations where commercial machines are used to print modified or falsified papers [11]. Crawford Ommen and Carriquiry (2023) offer a statistical method to assist forensic handwriting examiners in reaching more objective and fact-based conclusions. The study addresses the usual subjectivity in handwriting analysis by combining Bayesian inference and likelihood ratios to quantify the strength of handwriting evidence. To address the problem of document forgery Darem et al. (2020) offer a method for spotting digital forgeries in the compressed domain concentrating on images of official documents. This physicochemical characterization data requires advanced detection techniques that do not depend on full picture data decompression. Such techniques are especially helpful in time-critical systems or limited resource environments. The approach they employ examines artifacts produced in the course of compressing (e.g., JPEG quantization tables and block boundaries) to detect tampered patches that have abnormal statistical behavior [13]. Djeddi et al. explore the impact of codebook pattern configuration on key aspect of handwriting-based document authenticationswriter recognition. (2021). Their experimental research reported in Expert Systems examines the influence of local descriptor pattern coding strategies and clustering techniques on writer identification systems performance.

It flags papers that have something unusual about them and get them forged or altered. As this is a model-free approach, it can also be applied to domains where forgery techniques are constantly evolving as it does not rely on the history of previously forged documents. [20] The system of forensic document analysis was introduced by Gupta and Kumar (2020) using ensemble learning techniques i.e., bagging and boosting (eg, AdaBoost, Random Forest) in Soft Computing. To improve in the decision process in font matching, content consistency checking or signature verification, their approach is based on a set of weak classifiers. This approach is more robust and accurate in complex, high variance forensic data sets. The majority decision mechanism in Ensemb PLL is used to integrate multiple learners, thus improving robustness of the system with respect to noise, handwriting variability, and small forgeries— which are common in practice.[21] Hamzehyan et al. (2021) propose a novel approach of printer source identification based on feature model inside a given Total Variable Printer Space (TVPS). Moreover, this method accounts both for systematic differences among devices and because it carries properties of printers (mechanical deviations, the ink spread, misalignments). According to slight differences in printed data, the method is able to classify printers, even printers from a same manufacturer and model based upon that representation. The proposed TVPS-based classifier, can offer the forensic examiner a scalable and robust architecture and enhance the granularity and the specificity in printer attribution. [22] Hazra et al. present DIGIDOC: a handwritten document analysis toolkit in forensic setups as a part of their work in Computational Vision and Bio-Inspired Computing (2023). The system is based on computer vision and machine learning techniques and is also used to assist applications like stroke analysis, forgery detection, and writer identification.

All of which are provided by DIGIDOC including preprocessing, segmentation feature extraction and classification modules. Due to the tool's emphasis on flexibility and scalability, forensic analysts can work with a wide range of scripts and handwriting styles. It is appropriate for both academic and practical legal investigations due to its integration of neural networks with domain-specific heuristics. [23] Im et al. (2022) classify historical printed books using deep learning, recognizing printing technologies as offset, digital, and letterpress. Their approach, which was published in Multimedia Tools and Applications, analyzes digitized page textures and visual signals using CNN-based structures. Even under hostile circumstances, the system may be used for forensic attribution of printed materials and outperforms baseline models on common datasets.[29] We show a non-invasive method of labeling printers through the examination of their texture. This technique is referred to in IEEE Transactions on Information Forensics and Security. It has a good working memory and is user-friendly since no prior knowledge of the text, fonts, or symbols is required. Gabor and Local Binary Patterns (LBP) can distinguish between similar printers with fair precision, even if they are rather similar.

III. Comparative Result Analysis

We will compare various techniques employed in the past studies. We inspect such items as accuracy, F1 score, AUC, and robustness, to understand how well each method could identify and find image forgery. By so doing, we are able to gain a clear picture of the strengths and weaknesses of every method. These results are summarized in the Table I, below illustrate the performance of each method and its limitations.

TABLE I: METHODOLOGICAL COMPARATIVE REVIEW ANALYSIS

Reference No.	Method Used	Findings	Strength	Limitations
[1]	Proposed a new metaheuristic optimization algorithm called Puma Optimizer (PO), inspired by the hunting behavior of pumas; tested on benchmark functions and machine learning tasks	The PO algorithm demonstrated competitive or superior performance in terms of convergence speed and solution quality when compared with other well-known metaheuristics	Simple structure, easy implementation, effective exploration and exploitation balance; applicable to a wide range of optimization and machine learning problems.	Like many metaheuristics, performance may degrade in very high-dimensional or extremely complex problem spaces; theoretical analysis is limited.
[2]	Image processing and pattern analysis techniques are applied to detect anomalies in guilloche patterns typically used in security printing.	The proposed method effectively detects forged documents by identifying inconsistencies or alterations in guilloche patterns.	Focuses on a key security feature (guilloche) commonly found in identity documents.Non-destructive and image-based, allowing for automated verification	May not generalize well to documents lacking guilloche patterns.Effectiveness could degrade with low-resolution or poor-quality scans
[3]	Utilization of both intrinsic (printer-specific artifacts) and extrinsic (externally added) signatures to embed hidden information in printed documents for security and traceability.	Demonstrated the feasibility of using signature-based methods to support secure printing and document authentication.	Combines multiple signature types for enhanced robustness. Applicable to real-world printing systems like electrophotographic devices.	Requires calibration and control over printer settings. May not be easily transferable across different printer models or technologies
[4]	Applied Principal Component Analysis (PCA) and Gaussian Mixture Models (GMM) on gray-level co-occurrence matrix (GLCM) features extracted from printed documents for printer identification	The proposed method could effectively classify printed documents according to their source printers with high accuracy.	Utilizes robust statistical and texture-based features.Demonstrated effectiveness even with limited training samples.	Performance may degrade with low-resolution scans or noisy inputs. Method depends on specific texture patterns that may vary with
[5]	Exploited artifacts introduced by Color Filter Array (CFA) interpolation patterns to identify the source camera of an image.	The method successfully linked digital images to their originating cameras based on CFA-related traces.	Effective in camera source identification without requiring access to the device. Utilizes inherent sensor-level artifacts that are difficult to forge or remove.	May struggle with heavily compressed or processed images. CFA artifacts can be weakened by resizing, filtering, or image enhancements
[6]	Surveyed various image forgery detection techniques,	Provided a detailed taxonomy and comparative	Offers a broad overview helpful for researchers	Lacks experimental implementation or

	including copy-move, splicing, and retouching, using both classical and deep learning methods.	analysis of forgery detection methods, highlighting trends and gaps in current research.	entering the field. Highlights strengths and weaknesses of different detection approaches.	benchmarking. May not reflect the most recent deep learning advancements due to publication lag.
[7]	Method Used: Leveraged intrinsic features of documents, such as printing artifacts and texture, using image analysis.	The system effectively identified fraudulent documents by analyzing their intrinsic properties	Non-intrusive, robust to superficial changes, and useful in real-world scenarios.	Performance may degrade with high compression or low resolution scans
[8]	Compared text-dependent and text-independent techniques for source printer identification.	Text-independent methods performed better in scenarios with diverse and limited textual content.	Versatile and robust across multiple document types.	Some techniques were sensitive to noise and layout variations.
[9]	Used non-destructive inspection techniques like hyperspectral imaging and microscopy.	Identified unique characteristics of ink and paper in handwritten documents	Preserves document integrity; suitable for legal evidence	Requires specialized hardware and technical expertise.
[10]	Proposed a deep feature aggregation network for anomaly detection in hyperspectral document images	The approach successfully identified forged regions and inconsistencies.	Effective for subtle and complex forgeries.	Computationally expensive and data-hungry.
[11]	Extracted statistical features from Discrete Wavelet Transform (DWT) domains for printer identification.	Achieved high accuracy in identifying color laser printers.	DWT is effective for capturing texture and frequency information.	May be sensitive to resolution or compression artifacts.
[12]	Discrete Wavelet Transform (DWT) domains were used to obtain statistical features for printer identification.	Color laser printers were identified with high accuracy.	DWT is capable of efficiently capturing texture and frequency information.	May be prone to resolution problems or compression artifacts.
[13]	Low-level features are used to detect document forgeries in the compressed JPEG domain.	Effective detection was accomplished without image decompression.	Quick and economical with resources.	Only works with compressed formats; less accurate for complicated forgeries.
[14]	Method: For offline writer recognition, codebook patterns in feature representation were examined.	Recognition performance was considerably impacted by specific codebook factors.	Assists in improving feature extraction techniques	Appropriate parameter tweaking is critical to performance.
[15]	Method Used: For global optimization challenges, Greylag geese Optimization (GGO), which was inspired by the behavior of geese swarms, was introduced.	GGO fared better than a number of conventional metaheuristic algorithms	Flexible, converges more quickly in a variety of benchmark issues.	More testing on actual document forensic activities is required.
[16]	To handle optimization difficulties, the Football Optimization Algorithm (FOA) was proposed, which simulates team sports tactics.	FOA outperformed other metaheuristic algorithms on a variety of common benchmark functions.	Strengths: Adaptability across disciplines and creative strategy modeling.	Depending on the complexity of the issue, the performance may change and require adjustment.
[17]	Using ensemble learning classifiers such as Random Forest and Gradient Boosting, a text-independent printer recognition system was created.	: High accuracy and robustness were attained without depending on particular textual content.	Adaptable to many document formats and types.	To preserve performance among printers, a sizable and varied training dataset might be required.
[18]	Assigned prints to laser printers using machine learning models and data-driven feature characterisation.	Showed that use learnt features rather than handcrafted ones increased classification	Adaptable to changing printing technology and scalable.	Data-driven approaches necessitate a large amount of labeled data and may be opaque (black-box).
[19]	For laser printer attribution, new feature extraction algorithms based on texture and spatial characteristics were proposed.	Beat conventional feature-based methods.	Sensitivity to minute differences and high discriminative capability.	Depending on scan resolution, preprocessing and calibration can be necessary.
[20]	To find anomalies in printed documents, unsupervised learning approaches (such as clustering and anomaly detection) were applied.	Suspicious documents that deviated from recognized printer signatures were successfully reported.	Operates in dynamic and unidentified situations; requires no labeled data.	Needs careful adjustment of anomaly levels and is prone to false positives.

[21]	Method: Using ensemble approaches such as bagging and boosting for classification, a forensic document examination system was constructed.	Boosting greatly increased the forensic feature classification accuracy.	Excellent multi-feature fusion performance and dependability.	Model complexity rises with ensemble size; computational cost is increasing.
[22]	Approach Used: Like speaker recognition systems, printer properties were modeled in a whole variability space.	Presented a small but efficient printer identification representation.	Preserved discriminative power while reducing dimensionality.	If training data isn't diverse, performance can deteriorate.
[23]	Method: DIGIDOC, a forensic tool for handwritten document analysis, was created utilizing image processing techniques.	DIGIDOC successfully detects the features of fake and modified handwriting.	A user-friendly forensic program with useful applications for document analysis in the real world.	Document photos that are severely damaged or of poor quality may have limits.
[24]	Method: Classified printing technologies in digitized photos of old books using deep learning networks (CNNs).	Accurate classification of printing kinds (such as offset, inkjet, and letterpress) was attained.	Useful for digital archives and historical document interpretation.	The availability of labeled historical datasets affects performance.
[25]	Method: To detect document forgeries, deep features were extracted using convolutional autoencoders.	Improved ability to differentiate between modified and authentic papers.	Automatic feature learning without manual engineering.	Requires high computational resources and large training data.
[26]	DFD-SS was introduced, utilizing spectral-spatial properties to detect hyperspectral document forgeries. Results:	By combining spectral and spatial data, a high detection accuracy was attained. Strengths	: Good at spotting minute, imperceptible changes.	Requires access to hyperspectral imaging equipment, which is a limitation.
[27]	Method Used: A multi-size block-based approach based on Benford's Law was suggested for printer identification.	By using natural statistical deviations, the approach could distinguish between printers.	Doesn't require complicated feature extraction and is lightweight.	It might not work as well on photos with a lot of compression or low resolution.
[28]	Method Used: Created a strong watermarking system for PDF documents that guarantees security and compatibility.	High watermark resilience against conversion and document tampering was attained.	Prevents common assaults and supports PDF standards.	Document complexity or size may be marginally increased by the embedding procedure.
[29]	Approach Employed: To improve source printer identification, printer-specific letter descriptor pooling was suggested.	By combining characteristics from different letters, classification accuracy was increased.	Integrates contextual data with micro-patterns.	Short or poorly printed text samples may cause performance to deteriorate
[30]	For printer classification, a single passive classifier was combined with local texture descriptors.	Provided a straightforward but efficient way to distinguish printers.	Suitable for real-time applications and low in complexity.	Not as resilient to visual distortions or noise as deep learning techniques.

IV. Conclusion and Future Scopes

Comparison of the advanced image forgery detection methodologies reveal significant improvement in addressing various categories of forgeries such as copy move, splicing and low quality edits. Majority of recent research is evidently inclined towards the use of deep learning models, hybrid models, and adaptive feature extraction to enhance the use of better accuracy, robustness, and scalability. The most popular methods in the area are CNNs, transformer-based models, and task-specific networks including dual-stream networks and feature fusion modules.

These techniques differ in the extent to which they address real-world problems such as post-processing attacks, geometric distortions and generalization between datasets. An example of such examples is the FBI-Net [29] and DIF-Net [38], which perform exceptionally well with both IoU score of approximately 70.99% and F1 score of more than 87. RSADTL-CMFD [31] and Super Point [30] are the most effective, scalable, and computationally efficient in the field of copy-move forgery detection.

On the other hand, for resource-constrained infrastructures, MobileNetV2 [27] and Inspector [25] are more appropriate that enables rapid but accurate detection of mobile as well as embedded devices. If the applications need to be adversarial-response resistant, then robust models like CFM Framework [21], MDCF-Net [24] outperform each other as they produce their AUC > 0.96 along with maintaining a high accuracy rate for security-critical face forgery detection application.

FUTURE SCOPES

Despite all the above achievements, some remain open: those relating to achieving robustness and resiliency toward more sophisticated attacks, including OSN-induced compression and adversarial perturbations. All in all, about 40% of reviewed techniques point out further generalization requirements, which sometimes come combined with domain generalization methods or synthetic data augmentations, and in a few cases with multi-feature fusions. For example, the multi-scale artifact tracking-based approach employed by MSATT and CDPNet [19] significantly enhances the detection accuracy of manipulated traces, whereas TransU2-Net [24] adopted hybrid CNN-transformer architectures that make the approach outstanding to use with object-level tampering scenarios. However, it is still considered a research field to be extensively explored in the case of dealing with diverse image characteristics and sources. Future work will be dedicated to the hybridization of techniques that combine the strengths of CNNs, transformers, and domain-specific optimization modules. For real-time and resource-constrained applications, lightweight models such as Lightweight CNN [28] and LDSFNet [21] could be further optimized to balance accuracy and computational efficiency. In forensic applications where interpretability is of importance, keypoint-based and explainable deep learning models should be prioritized. Finally, more detailed benchmark data sets that cover extensive forgery types, attack scenario types, and quality variations are necessary to allow standard comparisons and innovation about the detection approach of forgeries. In summary, while models like FBI-Net, DIF-Net, and TransU2-Net represent state-of-the-art solutions for image forgery localization, lightweight and domain-adaptive models like Inspector and MobileNetV2 are more suitable for low-resource applications to be scaled out.

REFERENCES

[1] Abdollahzadeh, B., Khodadadi, N., Barshandeh, S., Trojovský, P., Gharehchopogh, F.S., El-kenawy, E.S.M., & Mirjalili, S. (2024). Puma optimizer (PO): a novel metaheuristic optimization algorithm and its application in machine learning. Cluster Computing, 1–49.

[2] Al-Ghadi, M., Ming, Z., Gomez-Krämer, P., & Burie, J.C. (2022). Identity documents authentication based on forgery detection of guilloche pattern. arXiv preprint arXiv:2206.10989.

[3] Ali, G.N., Mikkilineni, A.K., Allebach, J.P., Delp, E.J., Chiang, P.J., & Chiu, G.T. (2003). Intrinsic and extrinsic signatures for information hiding and secure printing with electrophotographic devices. NIP & Digital Fabrication Conference, 19, 511–515. Society of Imaging Science and Technology.

[4] Ali, G.N., Mikkilineni, A.K., Delp, E.J., Allebach, J.P., Chiang, P.J., & Chiu, G.T. (2004). Application of principal components analysis and Gaussian mixture models to printer identification. NIP & Digital Fabrication Conference, 20, 301–305. Society of Imaging Science and Technology.

[5] Bayram, S., Sencar, H., Memon, N., & Avcibas, I. (2005). Source camera identification based on CFA interpolation. IEEE International Conference on Image Processing (ICIP), 3, III-69. IEEE.

[6] Benhamza, H., Djeffal, A., & Cheddad, A. (2021). Image forgery detection review. International Conference on Information Systems and Advanced Technologies (ICISAT), 1–7. IEEE

[7] Bertrand, R., Gomez-Krämer, P., Terrades, O.R., Franco, P., & Ogier, J.M. (2013). A system based on intrinsic features for fraudulent document detection. 12th International Conference on Document Analysis and Recognition (ICDAR), 106–110. IEEE.

[8] Bibi, M., Hamid, A., Moetesum, M., & Siddiqi, I. (2022). Document forgery detection using source printer identification: a comparative study of text-dependent versus text-independent analysis. Expert Systems, 39(8), e13020.

[9] Breci, E., Guarnera, L., & Battiato, S. (2024). Innovative methods for non-destructive inspection of handwritten documents. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 4825–4829. IEEE.

[10] Cheng, X., Huo, Y., Lin, S., Dong, Y., Zhao, S., Zhang, M., & Wang, H. (2024). Deep feature aggregation network for hyperspectral anomaly detection. IEEE Transactions on Instrumentation and Measurement.

[11] Choi, J.H., Im, D.H., Lee, H.Y., Oh, J.T., Ryu, J.H., & Lee, H.K. (2009). Color laser printer identification by analyzing statistical features on discrete wavelet transform. 16th IEEE International Conference on Image Processing (ICIP), 1505–1508. IEEE.Mishra, M., Jain,

[12] Crawford, A.M., Ommen, D.M., & Carriquiry, A.L. (2023). A statistical approach to aid examiners in the forensic analysis of handwriting. Journal of Forensic Sciences, 68(5), 1768–1779.

[13] Darem, A., Alhashmi, A.A., Javed, M., & AbuBaker, A.B. (2020). Digital forgery detection of official document images in compressed domain. International Journal of Computer Science and Network Security (IJCSNS), 20(12), 115.

[14] Djeddi, C., Siddiqi, I., Gattal, A., Al-Maadeed, S., & Ennaji, A. (2021). Influence of codebook patterns on writer recognition: an experimental study. Expert Systems, 38(3), e12662.

[15] El-Kenawy, E.S.M., Khodadadi, N., Mirjalili, S., Abdelhamid, A.A., Eid, M.M., & Ibrahim, A. (2024). Greylag goose optimization: nature-inspired optimization algorithm. Expert Systems with Applications, 238, 122147.

[16] El-Kenawy, E.S.M., Rizk, F.H., Zaki, A.M., Mohamed, M.E., Ibrahim, A., & Abdelhamid, A.A. (2024). Football optimization algorithm (FbOA): a novel metaheuristic inspired by team strategy dynamics. Journal of Artificial Intelligence Metaheuristics, 8(1), 21–38.
[17] El Abady, N.F., Taha, M., & Zayed, H.H. (2022). Text-independent algorithm for source printer identification based on ensemble learning. Computers, Materials & Continua, 73(1).
[18] Ferreira, A., Bondi, L., Baroffio, L., Bestagini, P., Huang, J., Dos-Santos, J.A., & Rocha, A. (2017). Data-driven feature characterization techniques for laser printer attribution. IEEE Transactions on Information Forensics and Security, 12(8), 1860–1873.
[19] Ferreira, A., Navarro, L.C., Pinheiro, G., dos Santos, J.A., & Rocha, A. (2015). Laser printer attribution: exploring new features and beyond. Forensic Science International, 247, 105–125.
[20] Gebhardt, J., Goldstein, M., Shafait, F., & Dengel, A. (2013). Document authentication using printing technique features and unsupervised anomaly detection. 12th International Conference on Document Analysis and Recognition (ICDAR), 479–483. IEEE.
[21] Gupta, S., & Kumar, M. (2020). Forensic document examination system using boosting and bagging methodologies. Soft Computing, 24(7), 5409–5426
[22] Hamzehyan, R., Razzazi, F., & Behrad, A. (2021). Printer source identification by feature modeling in the total variable printer space. Journal of Forensic Sciences, 66(6), 2261–2273.
[23] Hazra, A., Pal, D., Pal, B., & Bandyopadhyay, A. (2023). DIGIDOC: a handwritten document analysis tool for forensic application. In Computational Vision and Bio-Inspired Computing: Proceedings of ICCVBIC 2022 (pp. 379–397). Springer, Singapore.
[24] Im, C., Kim, Y., & Mandl, T. (2022). Deep learning for historical books: classification of printing technology for digitized images. Multimedia Tools and Applications, 81(4), 5867–5888.
[25] Jaiswal, G., Sharma, A., & Yadav, S.K. (2022). Deep feature extraction for document forgery detection with convolutional autoencoders. Computers & Electrical Engineering, 99, 107770.
[26] Jaiswal, G., Sharma, A., & Yadav, S.K. (2022). DFD-SS: document forgery detection using spectral–spatial features for hyperspectral images. Journal of Visual Communication and Image Representation, 89, 103690.
[27] Jiang, W., Ho, A.T., Treharne, H., & Shi, Y.Q. (2010). A novel multi-size block Benford's law scheme for printer identification. In Advances in Multimedia Information Processing – PCM 2010 (pp. 643–652). Springer, Berlin.
[28] Jiang, Z., Wang, H., & Han, S. (2024). A robust PDF watermarking scheme with versatility and compatibility. Multimedia Tools and Applications, 1–27.
[29] Joshi, S., Gupta, Y.K., & Khanna, N. (2022). Source printer identification using printer specific pooling of letter descriptors. Expert Systems with Applications, 192, 116344.
[30] Joshi, S., & Khanna, N. (2017). Single classifier-based passive system for source printer classification using local texture features. IEEE Transactions on Information Forensics and Security, 13(7), 1603–1614.