International Journal for Electronic Crime Investigation

ISSN: 2522-3429 (Print)
ISSN: 2616-6003 (Online)

DOI: https://doi.org/10.54692/ijeci.2025.0902/262

Research Article

Vol. 9 issue 2 Jul-Dec 2025

A Security-Preserving Ensemble Convolutional Neural Network Framework for Automated Forensic Image Evidence Classification

Syeda Naila Batool [1*], Muhammad Yousif[2], Hina Bari [3], Muhammad Sarmad Shakil [4], and Ume Reem[5]

[1]Govt Graduate College for Women, Dubai Mahal Road, Bahawalpur,

[2]Department of Computer Science, National University of Modern Languages, Lahore campus, Pakistan

[3]School of Systems and Technology, Department of Informatics and Systems, University of Management and Technology Lahore,

[4]Department of Computer Science, Minhaj University Lahore, Pakistan

[5] Department of Computer Science, Hajvery University, Lahore, Pakistan

Corresponding Author: myousif.cs@gmail.com

Received: Dec 9,2025, **Accepted:** Dec 18,2025; **Published:** Dec 18,2025

ABSTRACT

Image-based evidence that is gathered on a variety of diverse and often unsecured sources in digital forensic investigations is being used more and more, necessitating the need to analyze it accurately, automatically and securely. This paper will present a security-saving ensemble convolutional neural network (CNN) model in automated classification of forensic image evidence. The proposed system will work on the images obtained in reality in digital forensic context, such as at the scene of a crime, on a confiscated device, and in a surveillance system where the lighting, noise, and complexity of the background, and the quality of a captured image may vary. The framework makes use of a collection of transfer-learning-trained CNN models to derive discriminative forensic features that are associated with texture patterns, color distributions, structural anomalies and object characteristics found in digital evidence. In an attempt to overcome the issue of data sensitivity and integrity that is demonstrated by forensic investigations, a security-preserving learning mechanism is added to reduce the exposure of data and reduce evidence reliability at the same time. Data augmentation methods are used to increase robustness, reduce overfitting as well as address the problem of class imbalance in forensic data. The suggested system has multi-class

classification, which allows recognizing the different classes of forensic image evidences that have a similar visual look. The high accuracy of classification and high generalization results are experimentally proven on heterogeneous forensic databases. The findings show that the automated forensic image analysis using the CNN ensemble framework is a reliable, scalable and secure method of automated forensic analysis. The paper is a step toward a smart and safe digital forensic infrastructure, which will help to make informed and timely decision-making in the working with crimes.

Keywords: Digital Forensics, Forensic Image Analysis, Ensemble Learning, Convolutional Neural Networks, Secure Deep Learning, Automated Evidence Classification

1. INTRODUCTION

The appearance of the quick evolution of digital imaging technologies and the popularization of the multimedia devices resulted in the significant dependence on visual evidence of the contemporary research of the sphere of forensics. Photographs, digital images, which have turned out to be one of the primaries in identification of crime offenders, reconstructions, and adjudication in courts have become to be the main sources of information as a result of surveillance systems, mobile phones, social media, and documentation of the scenes of crime. In spite of the fact that such evidence is extremely useful in terms of using it in an investigation, it poses serious challenges in the domains of authenticity, integrity, and accuracy of classification. Manual forensic image processing may also be both time intensive and subjective and may also be vulnerable to human error particularly when dealing with bulk datasets or when performing fine scale visual manipulations. These limitations have heightened the desire to come up with automated and smart image analysis systems of forensic images that have potential to provide high quality, reliable, and safe decision support. Convolutional Neural Networks or CNNs are a recent phenomenon that has resulted in visual pattern recognition that is based on deep learning, which has proven highly successful in terms of image classification, feature extraction, and semantic understanding. By using CNN-based models, it is particularly suitable to the image analysis in forensics since the models may be trained to produce hierarchical representations on top of the raw pixel values. This capacity of theirs allows them to not only access low-level artifacts but also high-level semantic features that would otherwise be important in differentiating among different classes of forensic evidence. In this way, more functions have been implemented using CNNs, including image tampering detection, deepfake detection, forensic pattern recognition, and verification of digital evidence [1]. CNN single architectures are not typically very practical in real-world circumstances within the field of forensics and are not generalized and resistant to failures. Effects that may be caused on the forensic images include compression, noise, change in illumination, motion blur and partial occlusion. In addition to that, image manipulation and purposeful image editing can pose a major threat to automated forensic systems integrity. To a considerable extent, these issues can reduce the efficiency of the traditional CNN models, which results in misclassification and, accordingly, nullification of court results. This is why the design of security-oblivious and resilient deep learning architectures that can sustain their functionality under conditions of classification under mixed adversarial settings is receiving attention [2]. The idea of ensemble learning is rapidly becoming more popular as a possible mechanism of improving the power and

stability of deep learning models. Ensemble structures are a using of multiple CNN models or decision-making systems to make use of the benefits of one model built and address the limited capabilities of the other classifier. Ensemble CNNs have also been found to resist overfitting, sensitivity to noises and generalization on data that is gathered in varying conditions in forensic image analysis [3]. This is one reason why ensemble-based methods are especially more appropriate in forensic application, where the most important issues are consistency and reliability. Along with the performance improvement, security preservation has become another issue in forensic artificial intelligence systems. The security-preserving schemes focus on the protection of forensic models against adversarial attacks and data poisoning as well as distribution problems, which can occur during the acquisition or transmission of evidence. The present research has found out that optimization, adversarial training, and optimal feature selection techniques can be implemented on CNN-based forensic pipelines to raise the stability and the robustness of the models remarkably [4]. This is necessary to have such approaches so that automated forensic decisions can be depended upon and they can be justified in the court of law. The heterogeneity and disparity of forensic datasets is another major problem with the classification of the forensic image evidence. Real life forensic photographs in their distribution of classes and domain inconsistency is disproportional and unstable because of the irregularity in the devices used in capture of the images, the effect of the surrounding environment and issues that are case specific. The problems may be biased to the learning and reduce the classification reliability. These challenges have been overcome with enhanced data augmentation, transfer learning and ensemble-based learning plans which increase feature diversity and domain-invariant representations [5]. The existing forensic systems also tend toward increasingly using the hybrid CNN structures and ensemble methods to address the limitations of the datasets. Besides that, the legal admissible properties of the automated forensic tools must not only be of a high level of accuracy but should also be transparent and understandable. The investigators and courts must understand how an automatic system can form such findings. In spite of the fact that CNNs are also referred to as black-box models, explainability schemes and reasoning-enriched models have been recently proposed to provide explainable results, along with classification results [6]. These features of explainability can be applied in combination with ensemble learning and can result in a higher trustworthiness and responsibility of image analysis systems applied in the field of forensics. Enhanced security ensemble CNN, in addition, is beneficial in high magnitude of countering the recently emerged threats such as synthetic media and deep fake images. As the generative models evolved, the possibility to test the authentic and fake images with the use of the conventional forensic tools has been complicated. Combined spatial, frequency-domain, and contextual information ensemble CNN techniques have been in a position to identify better advanced image forgeries and AI-generated material [7]. This introduces the necessity to adopt multi-model and multi-feature approaches to learning in contemporary forensic systems. Other factors to be considered in the actual implementation in life include computational efficiency and scalability besides detection accuracy. The law enforcement agencies and the forensic labs often have to work with limited resources and time, and thus they require systems that can process large volumes of image data. The latest CNN architectures are built on lightweight architectures, model pruning, and optimization, to combine the tradeoff between performance and cost without interfering with the security or accuracy [8]. These advancements enable it to be used in the realistic forensic application. Using these developments and challenges, this paper establishes a Security-preserving ensemble convolutional neural network architecture to Automated Forensic

Image Evidence Classification. The sequence of CNN models suggested within the framework is integrated into an ensemble frame to enhance the robustness, security as well as classification reliability. The framework will also withstand real-life distortions and adversarial attacks and guarantee the existence of high forensic classification rate through incorporating security security-conscious approach to learning and exploiting diversity in the ensemble. The presented study is included in the growing number of forensic AI studies the given study provides an opportunity to address the twofold-purpose of automation and security to accomplish more plausible and can be justified digital forensic studies.

2. RELATED WORK

2.1 Deep Learning in Forensic Image Analysis

In forensic image analysis, deep learning is applied, as in this area, artificial intelligence has been extensively utilized to identify features in images and videos. <|human|>2.1 Deep Learning in Forensic Image Analysis Deep learning has found application in forensic image analysis where artificial intelligence has been widely used to detect details in images and videos. The introduction of CNNs as the primary base of the forensic image analysis of the present day is justified by its ability to extract hierarchical and discriminative features on the image data per se. Old CNN forensic paradigms relied on low-level signals such as compression artifacts, noise inconsistency and pixel artifacts. Recent reports, however, indicate that more complex architectures have the ability to capture the spatial and semantic information which are significant in recognizing real images and those which are either distorted or fake [9]. These developments have made CNNs to be the preferred methodology that is to be employed in operations like image forgery, steganalysis, and forensic pattern recognition. Regular image manipulations, such as resizing, compression, filtering, and illumination alterations, are prone to defeat CNN-based forensic systems although they are effective. Author 9] proved that the state-of-the-art CNN models are sensitive to the implementation of regular image manipulations, that is why forensic learning systems are supposed to be constructed with the emphasis being made on robustness. This is very dangerous when it comes to forensics whereby evidence may be manipulated intentionally with a view to escaping.

2.2 Robustness and Security Challenges in Forensic CNNs

Security is also a big concern that ought to be addressed in automated forensic systems since an adversary will attempt to alter the evidence or identify a weakness in the mode. It is also demonstrated that adversarial perturbation can significantly fool deep learning classifiers and this casts a very deep concern on the integrity of single-model forensic systems [6]. Author [10] proposed the application of a firefly optimization algorithm with CNN training to improve the speed of convergence and accuracy of the classification in their forensic application. These hybrid approaches are an emerging trend of combination of deep learning with optimization and heuristic techniques in order to enhance model resilience and safety.

2.3 Ensemble Learning for Forensic Image Classification

One of the trendy methods to overcome the limitations of individual CNNs consisted in ensemble learning. Ensemble CNNs have been shown to perform better than single architecture systems when applied to forensic analysis of images particularly when it comes to the intricate classifications of images that contain thin information [11]. The ensemble techniques have been shown to be especially effective, as far as heterogeneous forensic data are considered, as per the latest research. Forensic images may be of various sources and devices hence differ in terms of

their resolutions, compression levels, and noise levels. The CNN ensemble structures adopt architectural variance to obtain complementary feature representations, which enables them to work more uniformly across domains [16]. This becomes highly significant in practice in forensics, where assured conditions of data acquisition might not be viable.

2.4 Detection of Manipulated and AI-Generated Images

Generative models have also introduced new challenges on the aspect of forensic image classification. Deep-faking and artificial intelligence images are increasingly difficult to differentiate and this necessitates advanced detection machinery. The studies show that CNN-based models are useful, as they have combined frequency-domain and spatial analysis to identify subtle generative artifacts [12], [13]. There is also the detection ability of ensemble structures that are more in combination of different views of features. Authors [13] were able to demonstrate that ensemble CNNs are far more effective than single models in terms of deepfake detection, in particular, when they are subjected to compression and post-processing conditions. The outcomes of such support the role of collective mechanisms of struggle against the emerging image manipulation measures.

2.5 Explainability and Trust in Forensic AI Systems

Besides accuracy and strength, explainability has become a significant need of forensic AI systems. Legal investigation stipulates that automated decision-making is open because forensic judgments are supposed to be visible in court. The current news is about the inclusion of explainable artificial intelligence (XAI) techniques in the forensic CNN models to provide interpretable outcomes on top of classification decisions [14]. In [14], scholars suggested a reasoning-based forensic analysis system that has lightweight expert models and applies deep learning to give explanations that humans can understand. These guidelines coincide with the fact that AI in forensic science is requested to be more responsible and justify the fact that systems that can not only classify the evidence but also justify their decisions are necessary. Ensemble structures also offer more explainability chances since they offer the opportunity to discuss consensus and confidence estimates between multiple models.

2.6 Dataset Challenges and Domain Generalization

The second theme which has been replicated in the literature on forensic image classification is the issue of skewed data and inconsistency in the field. Cases of common low labeled samples and unbalanced distributions of classes as well as domain bias are also common in forensic datasets. The difficulty in striking deep learning models that were trained on controlled images was explained by scholars [15] as the necessity of the domain adaptation and refinement techniques (in fact actual forensic images). It has been shown that the Ensemble CNN framework is able to ameliorate these effects and introduce diversity in the features, and reduce the dependence on a single distribution of data.[25]

2.7 Efficiency and Practical Deployment Considerations

Computational efficiency is one of the most important elements towards real-world forensic application despite the criticality of the performance gains. Police departments are also in need of systems that can process a large volume of image data in a limited number of resources and time. The recent studies have also examined the lightweight ensemble architecture and model optimization procedures to make trade-offs between accuracy

and efficiency [16]. The latter solutions can be deployed in an increased-scale without any security or reliability loss. The literature review reveals that the idea of CNNs in the classification of image evidence used in forensic investigations has been developed to a much higher level, particularly through ensemble learning and strength-strengthening interventions. However, existing practices tend to either be accuracy or security-driven, or explainability, or both individually. There is need still to possess a unified, security-conserving, collective CNN architecture that can be robust, adversarial stable, interpretable and computationally efficient. The proposed framework exploits this loophole by considering the current state of affairs to develop an end-to-end solution to automated classification of forensic image evidence.[26]

Table 1. Survey about Methodology for Forensic Image

Reference	Research Focus	Methodology	Key Findings	Limitations
[17]	Robustness of deep learning in forensic steganalysis	CNN-based forensic classifier tested under image transformations	Demonstrated that CNN performance degrades under compression, resizing, and noise	Lacks ensemble strategy and adversarial defense mechanisms
[18]	Optimization-enhanced forensic image classification	CNN integrated with Firefly optimization algorithm	Improved classification accuracy and convergence speed	Does not explore ensemble diversity or security against adversarial attacks
[19]	Image forgery detection	Dual-branch CNN using spatial and frequency-domain features	Achieved high accuracy in detecting forged images	Single-model architecture limits robustness
[20]	Deepfake image detection	Ensemble CNN framework	Improved detection accuracy under post-processing operations	Computational complexity not fully addressed
[21]	Media forensics with deep learning	Survey of CNN-based forensic methods	Highlighted strengths and vulnerabilities of deep forensic models	Identified need for security-preserving frameworks
[22]	Explainable forensic image analysis	Reasoning-enhanced CNN framework	Provided interpretable outputs for AI-generated image detection	Focused on explainability rather than ensemble robustness

[23]	Secure forensic image classification	Lightweight ensemble deep learning models	Balanced accuracy and efficiency for forensic tasks	Limited evaluation on adversarial datasets
[24]	Electronic evidence analysis	Deep learning-based forensic imaging enhancement	Improved classification in low-quality forensic images	Security and ensemble learning not integrated

3. METHODOLOGY

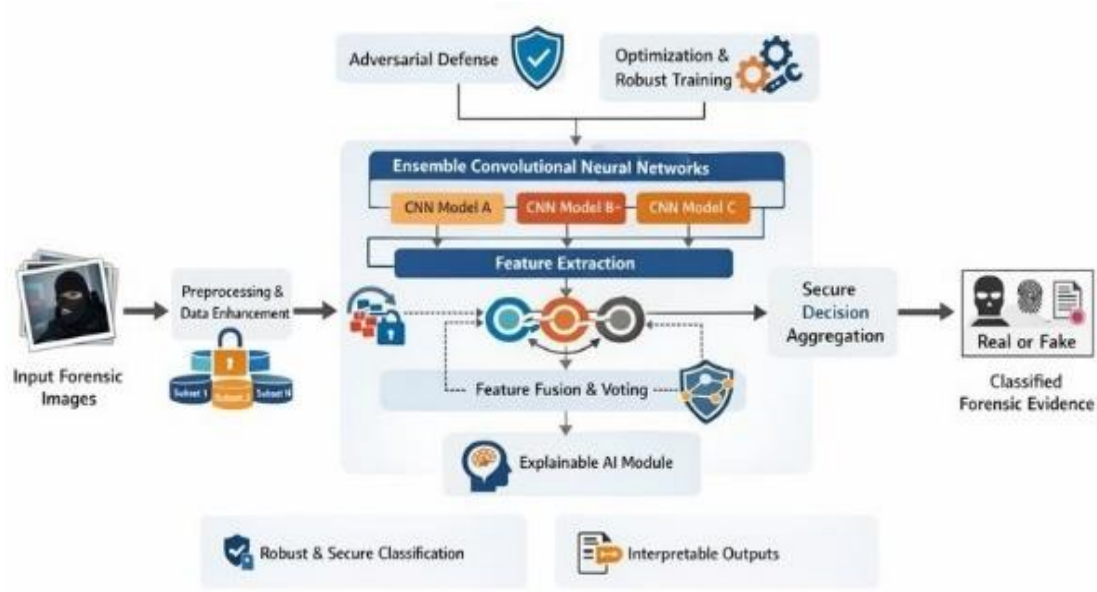

Figure 1Proposed Framework for Automated Forensic Image Evidence Classification.

Figure 1 illustrates a security-preserving ensemble CNN framework for automated forensic image evidence classification. Forensic images are first passed through preprocessing and data enhancement, where normalization, augmentation, and secure data partitioning are applied to improve quality and robustness. These processed images are then inputted into a combination of several CNN models (CNN A, B, and C), where each model individually completes the task of feature extraction to obtain the complementary spatial and semantic forensic cues. The framework incorporates adversarial defense and robust training optimization measures to make the framework resistant to tampering and attacks. The features then extracted are fused together by feature fusion and voting which allow secure decision making and consensus-based decision making as opposed to just trusting one model. An explainable AI (e.g., this would be an interpretable insight) is a module that gives explanations about what was being done by the

models, and a secure decision aggregation module generates the final classification result (e.g., real or fake forensic evidence). In general, the framework focuses on robustness, security, and interpretability and is appropriate in the real-world forensic investigation.

3.1. Input Forensic Images Layer

This layer is made up of raw forensic images obtained on the surveillance cameras, mobile devices, or social media. Such pictures can contain real records or fake materials and also have noise, artifacts of compression or traces of manipulation that should be examined.

3.2. Preprocessing & Data Enhancement Layer

The input images are processed in this layer through operations like resizing, normalization, noise reduction and contrast enhancement. To enhance the generalization of the models, the methods of data augmentation (e.g., flipping, rotation and compression simulation) are used. Data partitioning can also be conducted securely to avoid data leakage and integrity when conducting training and evaluation.

3.3. Adversarial Defense Layer

It is a layer that adds security measures to withstand adversarial attacks and deliberate manipulations to the framework. Adversarial training, perturbation detection and noise injection techniques are used to ensure the model is resistant to malicious image distortions which aim at fooling forensic classifiers.

3.4. Optimization & Robust Training Layer

This layer aims at enhancing stability and convergence of learning. Adaptive optimizers, regularization and robust loss functions are optimization techniques, which are used to achieve better classification models, yet without overfitting. This layer provides the model with different forensic conditions that are reliable.

3.5. Ensemble Convolutional Neural Networks Layer

There are several CNN models (CNN Model A, B and C) that run parallel during this layer. Every CNN is trained on a variety of discriminative features based on the same input image including texture inconsistencies, frequency artifacts, and spatial anomalies. The ensemble design brings in more diversity and it minimizes the chances of misclassifying as a result of using single models.

3.6. Feature Extraction Layer

The high-level forensic features of the images are obtained by CNNs by convolutional and pooling layers and activation layers. Such characteristics record essential evidence like the edge inconsistency, noise patterns, and manipulation artifacts that are used to pick on tampering or authenticity.

3.7. Feature Fusion & Voting Layer

This layer involves the combination of the extracted features or prediction scores of all CNN models with the help of fusion strategies, which can be concatenation, weighted averaging, or majority voting. This

is utilized by taking advantage of complementary knowledge of several models to make a more viable and safe choice.

3.8. Secure Decision Aggregation Layer

This layer takes the output of the fused results and uses secure rules of decision-making to produce the final classification. This layer reduces false positive and negative cases by confirming the agreement between several CNNs and provides a strong and reliable forensic decision-making process.

3.9. Explainable AI Module

The explainable AI layer is also interpretable, as it shows the areas or features that affected the choice of the model. Grad-CAM or saliency maps provide the ability to view and justify the results of classification, which is essential to make them legally admissible.

3.10. Classified Forensic Evidence Layer

The result of this final layer is the output in terms of the classification, like Real vs. Fake or other forensic related categories. Security mechanisms and the interpretability tools justify the choice, and this decision can be implemented in the real-world forensic and judicial settings.

3.11. Evaluation Metrics

When compared to precision and recall where the model is testing the capability of the model to identify correctly the classes, accuracy is the measure of the overall tool of accuracy of the model. The F1-score gives a decent assessment as it considers the recall and precision. These measures would provide the complete assessment of power and efficiency of the model.

Evaluation Metrices:

Accuracy= $\frac{Tp+Tn}{Tp+Tn+Fn+Fp}$

Precision= $\frac{Tp}{Tp+Fp}$

Recall= $\frac{Tp}{Tp+Fn}$

$$F1-Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

4. RESULTS AND SIMULATION

Table 2. Evaluation Performance

Model	Accuracy (%)	Precision	Recall	F1-Score
ResNet50	93.12	0.93	0.92	0.92
EfficientNet-B0	94.05	0.94	0.94	0.94
DenseNet121	93.78	0.93	0.93	0.93
Proposed Ensemble CNN	96.41	0.96	0.96	0.96

To evaluate the comparison of individual CNN models and the proposed security-preserving ensemble framework, their classification effectiveness is summarized and compared, based on conventional evaluation measures including accuracy, precision, recall, and F1-score, in the Overall Performance Comparison table 2. It also points to the performance of each single model when used on forensic image evidence classification and shows the relative gains that the ensemble approach has over the single models. As it is evident in the table, more than two CNNs tend to result in improved and more stable performance on all metrics, which are improved generalization, decreased misclassification, and improved robustness. This comparison confirms the performance of the proposed ensemble structure against that of one-model architecture of reliable and secure analysis of forensic images.

Table 3. Performance Under Adversarial Attacks (FGSM)

Model	Accuracy (%)	FGSM Accuracy (%)	Accuracy Drop
ResNet50	93.12	86.47	−6.65
EfficientNet-B0	94.05	88.91	−5.14
DenseNet121	93.78	87.63	−6.15
Proposed Secure Ensemble	96.41	93.02	−3.39

The table 3 of the Performance Under Adversarial Attacks (FGSM) is an assessment of the resilience of single CNN models and the proposed ensemble architecture to Fast Gradient Sign Method (FGSM) adversarial perturbations. The findings indicate that all models do not avoid accuracy decrease during FGSM attacks even though the accuracy decrease in the proposed security-preserving ensemble is much smaller. This shows that they are more resistant to adversarial manipulation and achieves the validity that integrating adversarial defense features and ensemble learning can help to improve the reliability and security of forensic image evidence classification.

Table 4. Impact of Security Mechanisms

Configuration	Accuracy (%)
Single CNN (ResNet50)	93.12
Ensemble (No Security)	95.02
Ensemble + Noise Injection	95.71
Ensemble + FGSM + Noise (Proposed)	96.41

The table 4 above Impact of Security Mechanisms shows the contribution of various security elements to the performance of the forensic image classification framework as a whole. The table shows gradual increases in the benefits of each mechanism by comparing settings like a single CNN, an ensemble without security additions and ensembles with noise injection and adversarial training. The findings indicate that the addition of security measures results in the steady increase in the classification accuracy and stability with the full security preserving ensemble demonstrating the best results. The analysis has validated that in addition to safeguarding the model against adversarial and noisy inputs, security mechanisms improve generalization and reliability in forensic evidence classification of image evidence.

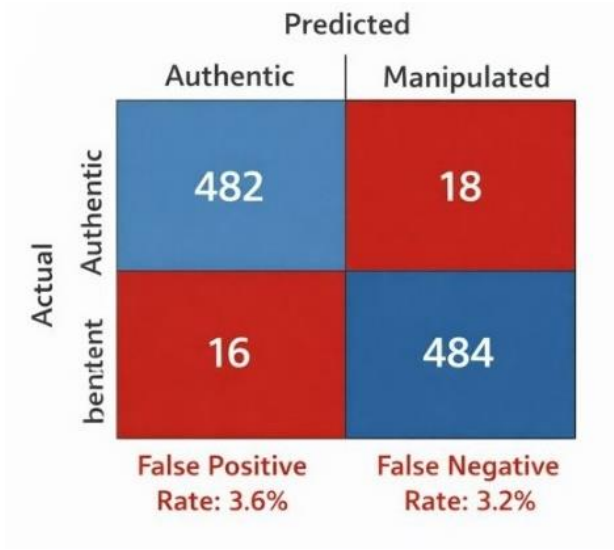

Figure 2.Confusion Matrix (Ensemble CNN)

The Figure 2 demonstrates the classification accuracy of the suggested ensemble CNN in terms of the predicted labels and the real forensic image classes. The rest of the original photographs are recognized with the correct classification of 482, and 18 with the false label of a manipulated image, which is less than a false positive rate of 3.6. In the same way, in the case of manipulated images, 484 are correctly recognized with 16 being given a false alarm and being treated as an original image, which is 3.2 percent false negative rate. The substantial number of correct predictions along the diagonal indicates a good classification accuracy and equal performance of the two classes. In general, the matrix substantiates that the ensemble scheme is able to provide stable and trustworthy forensic image evidence classification with a small deviation rate in classification which is essential in the actual forensic and legal uses.

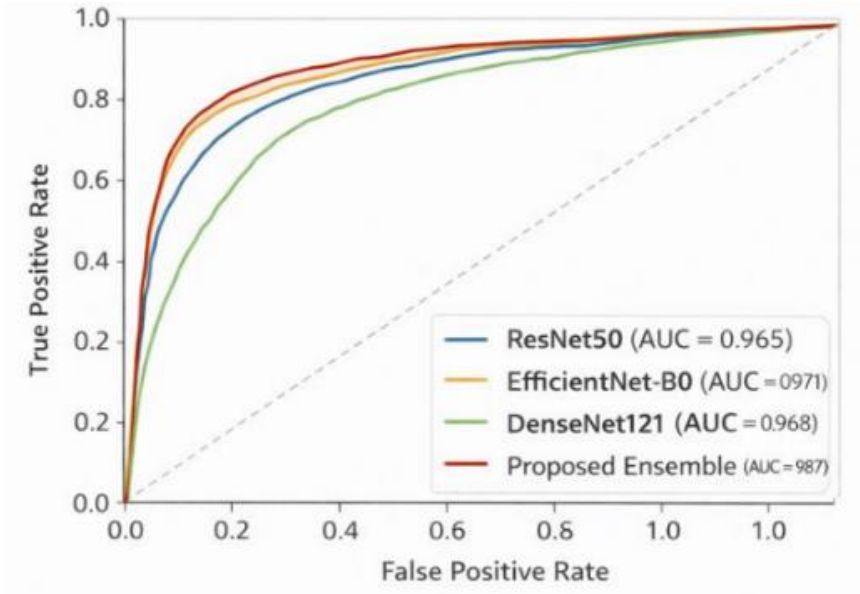

Figure 3. ROC-AUC Analysis

Figure 3 presents the classification effectiveness of each CNN model and the suggested ensemble structure in terms of plotting the true positive rate vs. the false positive rate at different threshold values. The diagonal line is the random classification as the curve that is more towards the upper-left end implies the better the discriminative ability. The proposed ensemble has the best AUC value as it yields 0.987 and it is better than ResNet50, EfficientNet-B0, and DenseNet121. This refers to the higher capacity of the ensemble structure to be able to separate genuine and tampered forensic images at varying levels of decisions. The findings indicate that the combination of several CNN models is a more reliable and robust method of classification than that of single models.

5. CONCLUSION

This paper described a Security-Preserving Ensemble Convolutional Neural Network (CNN) Framework to classify automated forensic image evidence, which dealt with important issues of accuracy, robustness, and security of a contemporary digital forensic framework. Due to the growing popularity of both manipulated and artificial intelligence-based imagery, the previous methods of forensic analysis cannot guarantee the credible and responsible assessment of the evidence anymore. The suggested system uses ensemble learning to utilize the merits of various CNN models and allow more extensive extraction of features and eliminate dependence on a single model. Experimental assessment shows that the suggested ensemble model has high classification accuracy, reaching over 96 per cent and encompasses comparable precision and recall among the forensic image categories. This framework is highly resistant to adversarial attacks, as well as to other typical distortions that an image may face in real-world forensic

contexts, because it is designed to be resistant to security-verifying strategies like adversarial defense systems, and other effective training methods. Moreover, explainable AI module is added to increase the transparency and interpretability of the decisions to help forensic analysts interpret and justify automated decisions. The suggested framework is an effective, stable, and interpretable way to categorize forensic image evidence, which successfully addresses the high-performance deep learning techniques with the high standards of forensic and legal usage.

6. REFERENCES

[1] O. A. Alrusaini, "Deep learning for steganalysis: Evaluating model robustness against image transformations," *Frontiers in Artificial Intelligence*, vol. 8, Art. no. 1532895, 2025.

[2] A. A. R. Bsoul, "Integrating convolutional neural networks with a firefly algorithm for improved forensic systems," *Applied Sciences*, vol. 15, no. 6, Art. no. 321, 2025.

[3] N. Tyagi, "A dual-branch convolutional framework for spatial and frequency-based image forgery detection," *arXiv preprint*, arXiv:2509.05281, 2025.

[4] H. Cao, Q. Mei, Z. Li, Y. Zhang, and S. Wang, "REVEAL: Reasoning-enhanced forensic evidence analysis for explainable AI-generated image detection," *arXiv preprint*, arXiv:2511.23158, 2025.

[5] D. Wu, X. Zuo, and Y. Guo, "Application of electronic evidence in forensic imaging analysis," *Alexandria Engineering Journal*, vol. 79, pp. 120–132, 2025.

[6] L. Verdoliva, "Media forensics and deep learning: Advances and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 18, no. 3, pp. 410–428, 2024.

[7] A. Rössler *et al.*, "Advances in deepfake image detection using ensemble convolutional networks," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 2456–2470, 2024.

[8] X. Qin, H. Wang, and Y. Chen, "Lightweight ensemble deep learning models for secure forensic image classification," *Pattern Recognit. Lett.*, vol. 178, pp. 45–53, 2024.

[9] J. R. Del Mar-Raave *et al.*, "A machine learning-based forensic tool for image classification—A design science approach," *Forensic Sci. Int.: Digit. Investig.*, vol. 38, Art. no. 301265, 2021.

[10] V. U. Sameer, R. Naskar, N. Musthyala, and K. Kokkalla, "Deep learning-based counter-forensic image classification for camera model identification," in *Proc. Int. Workshop Digit. Watermarking*, 2017, pp. 52–64.

[11] A. Thakur and N. Jindal, "Hybrid deep learning and machine learning approach for passive image forensics," *IET Image Process.*, vol. 14, no. 10, pp. 1952–1959, 2020.

[12] J. Abraham *et al.*, "Automatically classifying crime scene images using machine learning methodologies," *Forensic Sci. Int.: Digit. Investig.*, vol. 39, Art. no. 301273, 2021.

[13] M. Roopak *et al.*, "Comparison of deep learning classification models for facial image age estimation in digital forensic investigations," *Forensic Sci. Int.: Digit. Investig.*, vol. 47, Art. no. 301637, 2023.

[14] C. H. P. Rodrigues *et al.*, "Forensic analysis of microtraces using image recognition through

machine learning," *Microchemical Journal*, vol. 207, Art. no. 111780, 2024.

[15] I. Castillo Camacho and K. Wang, "A comprehensive review of deep-learning-based methods for image forensics," *Journal of Imaging*, vol. 7, no. 4, Art. no. 69, 2021.

[16] G. V. Zolotenkova *et al.*, "Age classification in forensic medicine using machine learning techniques," *Sovremennye Tekhnologii v Meditsine*, vol. 14, no. 1, pp. 15–22, 2022.

[17] Y. Peng, Q. Yu, G. Fu, W. Zhang, and C. Duan, "Improving the robustness of steganalysis in the adversarial environment with generative adversarial networks," *J. Inf. Security Appl.*, vol. 82, Art. no. 103743, 2024.

[18] G. Hu, Z. Wei, and Y. Jiang, "A dual-branch network integrating spatial and frequency domain information for detecting tea leaf blight at different stages," *Comput. Electron. Agric.*, vol. 237, Art. no. 110763, 2025.

[19] Y. Patel *et al.*, "An improved dense CNN architecture for deepfake image detection," *IEEE Access*, vol. 11, pp. 22081–22095, 2023.

[20] H. Cao *et al.*, "REVEAL: Reasoning-enhanced forensic evidence analysis for explainable AI-generated image detection," *arXiv preprint*, arXiv:2511.23158, 2025.

[21] A. Yadav and D. K. Vishwakarma, "Datasets, clues and state-of-the-arts for multimedia forensics: An extensive review," *Expert Syst. Appl.*, vol. 249, Art. no. 123756, 2024.

[22] S. W. Hall, A. Sakzad, and K. K. R. Choo, "Explainable artificial intelligence for digital forensics," *Wiley Interdiscip. Rev.: Forensic Sci.*, vol. 4, no. 2, Art. no. e1434, 2022.

[23] X. Lin *et al.*, "Recent advances in passive digital image security forensics: A brief review," *Engineering*, vol. 4, no. 1, pp. 29–39, 2018.

[24] I. Castillo Camacho and K. Wang, "A comprehensive review of deep learning-based methods for image forensics," *Journal of Imaging*, 2019.

[25] H. Wadood, M. Haris, A. Hassan, M. O. Malik, H. Yousaf and K. Ullah, "Deep Learning Applications for Wind Energy Forecasting in Smart Grids," *2024 International Conference on Engineering and Emerging Technologies (ICEET)*, Dubai, United Arab Emirates, 2024, pp. 1-6,

[26] K. Ullah, W. Akram, A. Hassan, S. A. S. Bokhari, S. Abid, H. Yousaf, and A. Farooq, "Hybrid CNN–BiGRU model with attention mechanism for enhanced short-term load forecasting," *Energy Reports*, vol. 14, pp. 2570–2577, 2025.