



Artificial Intelligence Approaches for Digital Image Forgery Detection

Prof. Vijaya Jadhav¹, Shreyas Shinde², Sanika Lomate³, Shreya Matkar⁴

¹ Assistant Professor, ² MCA Student, ³ MCA Student, ⁴ MCA Student

Department of Master of Computer Applications

ASM's Institute of Business Management & Research, Pune, Maharashtra, India

VijayaJ1979@gmail.com, shreyasshinde365@gmail.com, sanikalomate9370@gmail.com,
shreyamatkar60@gmail.com

How to Cite this Article:

Shinde, S., Lomate, S. & Matkar, S. (2026). Artificial Intelligence Approaches for Digital Image Forgery Detection. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(10), 1-9.
<https://doi.org/10.55041/ijcope.v2i10.070>

License:

This article is published under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

© The Author(s). Published by International Journal of Creative and Open Research in Engineering and Management.



<https://doi.org/10.55041/ijcope.v2i10.070>

Abstract—With the growing use of digital images in online communication and social media, verifying whether an image is authentic has become increasingly important. Image manipulation techniques such as copy-move, splicing, inpainting, and object removal can change the meaning of visual information. At the same time, JPEG compression and repeated recompression can weaken the forensic traces that are useful for detection. This review looks at two AI-based approaches for detecting digital image forgery. The first study combines Error Level Analysis (ELA) with the VGG-16 convolutional neural network to classify images as original or tampered. The second study focuses on compressed images shared through social networks and uses a modified ShuffleNet V2 architecture for lightweight forgery classification. The two studies are reviewed based on their objectives, datasets, preprocessing or feature representation, deep learning architecture, compression conditions, findings, and limitations. The comparison shows that ELA with VGG-16 can help expose compression-related manipulation artifacts, while the modified ShuffleNet V2 approach focuses on the practical challenge of detecting forgery after JPEG and social-network recompression. The review highlights dataset diversity, compression robustness, computational efficiency, and generalization across different manipulation and sharing conditions as important research gaps.

Keywords— Artificial Intelligence, Digital Image Forgery, Deep Learning, Error Level Analysis, VGG-16, ShuffleNet V2, Image Compression.

I. INTRODUCTION

Digital images are now an important source of information in social media, online communication, education, news, documentation, and digital evidence. At the same time, easily available image-editing software and manipulation tools allow users to change visual content by copying regions, inserting content, removing objects, or combining information from different images. These modifications can reduce

confidence in an image's authenticity and make digital investigation more difficult.

Traditional image-forensics methods often rely on manually designed features and traces produced during image acquisition or editing. These methods can work well when manipulation leaves clear artifacts, but detection becomes more difficult when an image is resized, recompressed, or repeatedly shared through an online platform. Social-network services commonly apply JPEG compression, which can modify or reduce



the traces that a forensic detector would otherwise rely on.

Artificial Intelligence (AI), particularly Deep Learning, provides a data-driven way to approach image forgery detection. Convolutional Neural Networks (CNNs) can learn visual patterns directly from training data and use them to distinguish authentic images from manipulated ones. However, a detector that performs well on a controlled dataset may not perform in the same way when compression, manipulation type, image source, or dataset distribution changes.

This review considers two related studies that focus on different practical aspects of image forgery detection. Sudiatmika combine Error Level Analysis with VGG-16 for detecting original and tampered images. Studiawan investigate forgery classification after social-network compression using a modified ShuffleNet V2 architecture. The purpose of this review is to understand how these approaches address image forgery, how their methodologies differ, and what research gaps remain for reliable real-world image authentication.

II. REVIEW METHODOLOGY

This review follows a focused literature-review approach using two selected studies related to AI-based digital image forgery detection. The studies were selected because they address the common problem of identifying manipulated images while considering different forensic conditions. The first study focuses on ELA-based preprocessing and deep CNN classification, while the second focuses on lightweight classification under JPEG and social-network compression.

For each study, the review considers the research problem, objective, dataset, manipulation conditions, preprocessing or feature representation, deep learning architecture, evaluation setting, major findings, and limitations. The extracted information is then compared across common dimensions instead of treating the reported results as directly equivalent. This distinction is important because the datasets, image conditions, architectures, and experimental settings are different.

The review process can be summarized as follows: Literature Survey → Selection of Two Related Studies → Method and Dataset Extraction → Individual Study Analysis → Comparative Analysis → Findings and Limitations → Research Gap Identification → Future Directions.

III. REVIEW OF AI-BASED IMAGE FORGERY DETECTION APPROACHES

A. Image Forgery Detection Using Error Level Analysis and Deep Learning

Sudiatmika, Rahman, Trisno, and Suyoto proposed an approach that combines Error Level Analysis (ELA) with Deep Learning to identify original and tampered images. The study uses the CASIA v2.0 dataset, which contains authentic and manipulated images. ELA is used as a forensic preprocessing technique to show differences in compression levels within an image. These differences can make suspicious regions more visible and provide a transformed representation for classification.

After ELA preprocessing, the images are processed using the VGG-16 Convolutional Neural Network. The model learns image features from the ELA-based representation and classifies the input as original or forged. The study reports a training accuracy of 92.2% and a validation accuracy of 88.46% after 100 epochs. These results show the potential of combining forensic preprocessing with a deep CNN, although they should be understood in the context of the dataset and experimental conditions used in the original study.

The main contribution of this approach is the combination of a visually interpretable preprocessing step with a well-established CNN architecture. ELA can highlight compression inconsistencies that may be difficult to notice through normal visual inspection. However, the approach also depends on how manipulation artifacts appear in the selected dataset. Changes in image source, compression history, or manipulation characteristics may affect the visibility of ELA patterns and, in turn, classification performance.

B. Forgery Classification on Compressed Images from Social Networks

Studiawan, Islami, and Shiddiqi focus on a practical problem in digital image forensics: Forged images shared through social networks may undergo JPEG compression or repeated recompression, which can change or hide forensic traces. The study combines 5,007 images from the MICC-F2000, MICC-F220, IMD2020, and Columbia datasets. The images include authentic and forged examples involving copy-move, splicing, inpainting, and object-removal manipulations.

To address the computational and compression-related challenges, the study proposes a modified ShuffleNet



V2 architecture. The modified architecture includes a Ghost module, a Squeeze-and-Excitation (SE) module, and a FReLU activation function. The model is evaluated under different image conditions, including uncompressed images, JPEG compression, repeated compression, and Facebook recompression. The design therefore focuses not only on forgery classification but also on the conditions under which images are commonly shared through social platforms.

The importance of this study comes from its focus on image conditions after sharing. Rather than assuming that forensic evidence remains unchanged, the study considers how compression affects the detection problem. Using a lightweight architecture also addresses the practical need to reduce computational cost. At the same time, achieving robustness across platforms, compression levels, and manipulation types

remains a challenge for deployment beyond the conditions studied.

combine historical prices, technical indicators, and FinGPT-derived Twitter sentiment, testing CNN, LSTM, attention-based LSTM, and a hybrid attention-CNN model against each other. It is a fairly direct example of where the field seems to be heading: architectures that try to capture long-range dependency, local pattern, and textual context all in one system.

One caveat worth flagging: attention weights are often read as a form of explanation — which time steps “mattered” — but that is not the same as a causal account of why the model predicted what it did. Treating predictive performance and interpretability as separate questions, rather than assuming one implies the other, seems like the more careful reading of this literature.

V. COMPARATIVE ANALYSIS

The two reviewed studies use Deep Learning for automated forgery classification, but they differ in their problem settings and methodological focus. The first study uses ELA to transform the image before VGG-16 classification, while the second focuses on efficient forgery classification under compression conditions using a modified lightweight architecture. Therefore, their reported performance should not be treated as a direct ranking of one approach over the other

Aspect	Study 1	Study 2	Dataset	Manipulations	Main Focus	Limitation
Approach	ELA + VGG-16	Modified ShuffleNet V2	CASIA v2.0	Multiple tampering types	Artifact-based classification	Dataset/generalization
Compression	Used in ELA representation	JPEG, repeated, Facebook	Multiple datasets	Copy-move, splicing, inpainting, object removal	Social-media robustness	Compression may weaken traces
Architecture	VGG-16 CNN	Lightweight modified ShuffleNet V2	5,007 images	Authentic + forged	Detection with lower computational burden	Generalization across conditions
Reported results	92.2% training; 88.46% validation	Evaluated across compression settings	Different experimental settings	Different manipulation distributions	Different practical priorities	Results are not directly comparable

The comparison shows that image forgery detection depends on more than the classifier architecture alone. The representation of forensic evidence, dataset source and diversity, manipulation type, and compression history all affect the detection problem. The ELA-based approach makes compression-level differences an explicit part of the input, while the ShuffleNet V2 approach treats compression as an operating condition that the classifier needs to handle. Together, the studies point to two complementary directions: forensic preprocessing and robustness at the architecture level.



VI. RESEARCH CHALLENGES AND FUTURE DIRECTIONS

A. Dataset Diversity and Generalization

Both studies use datasets containing labeled authentic and manipulated images. A model trained on one particular dataset may learn patterns specific to that dataset along with patterns related to forgery. Future studies should evaluate models across multiple datasets, image sources, resolutions, and manipulation pipelines to better understand generalization.

B. Compression and Recompression Robustness

Compression is a major challenge in practical image forensics. A forged image may be compressed once, repeatedly compressed, or processed differently by different social platforms. Future research can evaluate a wider range of quality factors, recompression histories, resizing operations, and platform-specific transformations. Robust detectors should continue to retain useful forensic information even when the original manipulation traces have been weakened.

C. Computational Efficiency

Deep CNN models can require significant computational resources during training and inference. The second study highlights the importance of lightweight architectures for practical deployment. Future work can explore efficient architectures while maintaining reliable detection across different image conditions. This is especially relevant for systems that need to process large numbers of shared images or run on resource-constrained devices.

D. Multiple Manipulation Types

Real-world images may contain multiple types of manipulation. Copy-move, splicing, inpainting, object removal, resizing, and other edits may occur together. Future datasets and evaluation protocols should therefore include mixed and sequential manipulations instead of relying mainly on isolated editing operations.

E. Explainability and Reliable Evaluation

A classification result alone does not fully explain why an image was identified as forged. Future AI-based forensic systems can benefit from localization or explanation mechanisms that identify suspicious regions and provide supporting evidence for the prediction. Evaluation should also clearly separate training, validation, and unseen test data, and report the

effects of compression and manipulation conditions so that improvements can be interpreted correctly.

VII. CONCLUSION

This review examined two AI-based approaches for digital image forgery detection. The first study combines Error Level Analysis with VGG-16 to identify manipulation-related compression differences and classify images as original or forged. The second study uses a modified ShuffleNet V2 architecture to address forgery classification under JPEG, repeated-compression, and social-network conditions.

The reviewed studies show that Deep Learning can support automated image forgery detection, but reliable performance depends on the data and operating conditions used to develop and test the model. ELA provides a useful forensic representation for exposing possible manipulation artifacts, while the lightweight ShuffleNet V2 approach emphasizes practical classification under compressed-image conditions. The review therefore identifies dataset diversity, compression robustness, computational efficiency, manipulation diversity, and explainable evaluation as important areas for future research. A robust real-world system will need to consider these factors together instead of relying on only one architecture or one dataset.

VIII. REFERENCES

- [1] I. B. K. Sudiatmika, F. Rahman, T. Trisno, and S. Suyoto, "Image forgery detection using error level analysis and deep learning," TELKOMNIKA (Telecommunication, Computing, Electronics and Control), vol. 17, no. 2, pp. 653–659, Apr. 2019, doi: 10.12928/TELKOMNIKA.v17i2.8976.
- [2] H. Studiawan, A. M. Islami, and A. M. Shiddiqi, "Forgery classification on compressed images from social networks to assist forensic analysis," Discover Computing, vol. 28, no. 1, Art. no. 331, 2025, doi: 10.1007/s10791-025-09825-6.
- [3] M. Zanardelli, F. Guerrini, R. Leonardi, and N. Adami, "Image Forgery Detection: A Survey of Recent Deep-Learning Approaches," Multimedia Tools and Applications, vol. 82, pp. 17521–17566, 2023, doi: 10.1007/s11042-022-13797-w.
- [4] N. T. Pham and C. S. Park, "Toward Deep-Learning-Based Methods in Image Forgery Detection: A Survey," IEEE Access, vol. 11, pp. 11224–11237, 2023, doi: 10.1109/ACCESS.2023.3241837.



- [5] F. Zare Mehrjardi, A. M. Latif, M. Sardari Zarchi, and R. Sheikhpour, “A Survey on Deep Learning-Based Image Forgery Detection,” *Pattern Recognition*, vol. 144, Art. no. 109778, 2023, doi: 10.1016/j.patcog.2023.109778.
- [6] G. K. Birajdar and V. H. Mankar, “Digital Image Forgery Detection Using Passive Techniques: A Survey,” *Digital Investigation*, vol. 10, no. 3, pp. 226–245, 2013, doi: 10.1016/j.diin.2013.04.007.