© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1717046

Deepfake Image Detection Using Error Level Analysis and ResNet18 Transfer Learning

DR. NAGUL MEERA SAYYED[1], POKALA VENKAT[2], ROWTHU BABU NAGENDRA KUMAR[3], SHAIK SAI RAHEEM[4]
[1, 2, 3, 4]*Department of Computer Science and Engineering, R.V.R. & J.C. College of Engineering*

Abstract*—Deepfake content has rapidly emerged as a major threat across digital media platforms, enabling the creation of highly realistic manipulated images that can mislead viewers, spread misinformation, and compromise security systems. Traditional image forensics approaches struggle to distinguish such sophisticated forgeries, especially when manipulation artifacts are subtle or deliberately concealed. To address this challenge, this work presents a robust deepfake image detection framework that integrates Error Level Analysis (ELA) with the ResNet18 deep neural network using transfer learning. ELA serves as a forensic preprocessing step that highlights compression inconsistencies introduced during image tampering. By recompressing the input image at a known JPEG quality and computing the pixel-level difference between the original and recompressed versions, ELA produces a high-frequency residual map that exposes hidden manipulation artifacts not visible in raw images. These ELA- generated images are then used to train and evaluate a modified ResNet18 model, where the final fully connected layer is replaced to accommodate binary classification of real and fake images. The proposed system is trained and tested on the publicly avail- able Deepfake and Real Images Dataset from Kaggle, comprising manipulated and pristine images across diverse categories. Experimental results demonstrate that the ELA-ResNet18 com- bination significantly enhances detection performance, achieving an overall classification accuracy of 97

I. INTRODUCTION

The rapid advancement of artificial intelligence and generative modeling techniques has transformed the landscape of digital media creation. In recent years, deep learning methods—particularly those based on Generative Adversarial Networks (GANs)—have enabled the synthesis of highly realistic images that are often indistinguishable from authentic visual content. These artificially generated or manipulated images, widely known as deepfakes, pose significant challenges to information security, digital integrity, public trust, and online safety. Deepfakes have been increasingly misused for malicious activities such as producing fabricated evidence, spreading misinformation, impersonating individuals, manipulating political discourse, and creating non-consensual content. As the realism of such manipulated images continues to improve, developing reliable methods for detecting deepfakes has become an urgent research priority.

Traditional image forensics techniques rely heavily on handcrafted features, statistical irregularities, or pixel-level inconsistencies. While these methods have historically been effective for detecting simple manipulations such as splicing, copy-move, or contrast adjustment, they struggle to identify deepfake content. Modern deepfake generation techniques use sophisticated neural networks that ensure high perceptual similarity between fake and real images. As a result, the inconsistencies introduced during manipulation are subtle and often beyond the perceptual limits of both humans and classical detection methods. This gap in detection capability motivates the exploration of hybrid approaches that combine digital forensics techniques with deep learning-based classification models.

One promising digital forensics method is Error Level Analysis (ELA), which examines the compression artifacts within an image. Most deepfake images—regardless of the generative method—introduce inconsistencies in the compression patterns because forged regions differ in JPEG quality from unaltered regions. ELA works by resaving an image at a known compression rate and comparing it with the original. Authentic images usually display uniform compression patterns, while manipulated images exhibit irregularities due to the combination of original and tampered content. Even when manipulation is subtle, ELA amplifies these inconsistencies, creating a residual map that highlights regions of potential tampering. Despite being visually subtle in the raw image, these artifacts become more pronounced in the ELA representation, making them ideal inputs for a neural network to learn discriminative features for deepfake detection.

In parallel, convolutional neural networks (CNNs) and trans- fer learning have revolutionized image recognition tasks by enabling the extraction of high-level semantic and textural features. Among these models, ResNet18 has emerged as a highly reliable architecture due to its residual connections, which mitigate vanishing gradient problems and allow for efficient deep feature extraction. ResNet18 is lightweight com- pared to larger architectures such as ResNet50 or VGG19, yet it provides strong representational capacity for tasks in- volving texture, noise patterns, and subtle anomalies. Since ELA produces high-frequency residual images rather than conventional photographs, using a model like ResNet18 is especially beneficial because its early convolutional layers excel at identifying edge-level and textural features that correspond to manipulation artifacts.

Recognizing the complementary strengths of ELA and ResNet18, this work proposes a hybrid deepfake detection system that employs ELA as a preprocessing step and a fine- tuned ResNet18 model for classification. The core idea is to transform the input images into ELA representations, which reveal tampering patterns, and feed these transformed images into a ResNet18 classifier trained using transfer learning. This hybrid approach effectively enhances the model's ability to distinguish between real and fake images even when manipulation is highly sophisticated or visually imperceptible.

To evaluate the effectiveness of the proposed system, we utilize the publicly available Deepfake and Real Images Dataset from Kaggle, which contains thousands of manipulated and authentic images across multiple categories. This dataset includes deepfake images generated using a variety of GAN-based techniques, providing a comprehensive evaluation environment. The dataset poses significant challenges due to the diversity in scenes, lighting, facial expressions, and manipulation techniques—attributes that mirror real-world deepfake scenarios. By applying ELA to every image in the dataset and training ResNet18 on this transformed dataset, the model learns to identify manipulation artifacts rather than content-specific patterns, thereby improving robustness and generalization.

The proposed ELA-powered ResNet18 model achieves an overall classification accuracy of 97

To support practical usability, the system is deployed through a Streamlit-based web interface, enabling users to upload images directly and receive predictions in real time. This interface makes the technology accessible for security personnel, digital forensics analysts, journalists, and general users who wish to verify the authenticity of images. The deployment also demonstrates the feasibility of integrating advanced deepfake detection models into real-world applications without significant computational overhead.

In summary, the surge of deepfake content represents a critical challenge that demands reliable detection mechanisms. This research contributes a robust and computationally efficient solution by combining the strengths of ELA and ResNet18. The approach not only amplifies tampering artifacts but also leverages advanced deep learning techniques to classify images with high accuracy. The results demonstrate that hybrid forensic-deep learning models are effective and scalable solutions for deepfake detection. With its strong performance on diverse data and real-time deployment capability, the pro- posed method represents a significant step forward toward safer and more trustworthy digital environments.

II. LITERATURE REVIEW

[1] Goodfellow et al. introduced Generative Adversarial Networks (GANs), establishing the foundational architecture that enables the synthesis of highly realistic artificial im- ages. Their work marked the beginning of modern deepfake generation methods, where adversarial training between a generator and discriminator produces photorealistic outputs that challenge traditional image forensics. [2] Fridrich et al. investigated early digital image forensics techniques that rely on analyzing JPEG compression artifacts to identify manipulated regions in images. Their study demonstrated that inconsistencies in quantization noise can reveal copy–move tampering even when modifications are visually imperceptible.

[3] Krawetz introduced Error Level Analysis (ELA) as a practical forensic method for exposing manipulation traces in JPEG images. By resaving the image at a controlled com- pression level and computing the difference from the original, the method highlights anomalous regions caused by splicing, blending, or GAN-based modifications.

[4] Bayar and Stamm developed a constrained convolutional neural network specifically designed for image manipulation detection. Their model learns forensic-aware filters that sup- press high-level semantics and focus on low-level manipulation cues, outperforming traditional handcrafted forensic approaches.

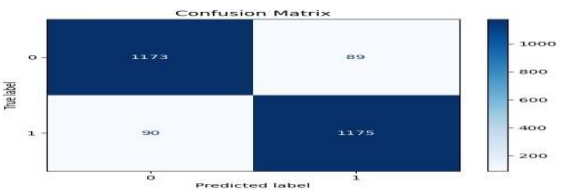
[5] Afchar et al. proposed MesoNet, a lightweight CNN architecture targeting deepfake video detection through mesoscopic feature extraction. Their results showed that compact networks can still effectively capture artifacts produced by generative models, making them suitable for real-time or resource-limited environments.

[6] He et al. introduced the ResNet architecture, which utilizes residual learning to mitigate vanishing gradients and enable very deep networks to learn robust features. ResNet18, in particular, has proven effective for texture and anomaly detection tasks due to its stability and strong representational capabilities.

[7] Nguyen et al. evaluated several transfer learning models for deepfake detection and found that ResNet-based architectures consistently outperform traditional CNNs. Their research emphasizes that residual connections allow the model to capture both global structures and subtle manipulation patterns introduced by GANs.

[8] Wu et al. examined the integration of ELA preprocessing with CNN classifiers for image forgery detection. Their find- ings revealed that ELA-enhanced images significantly boost model sensitivity toward manipulation artifacts, demonstrating the effectiveness of combining forensic preprocessing with deep learning.

[9] Korshunov and Marcel assessed the vulnerability of biometric systems to deepfake attacks and compared various detection strategies. Their study highlighted the need for ro- bust, generalizable deepfake detection frameworks capable of handling variations in quality, compression, and manipulation techniques.



III. EXISTING SYSTEM

A. Traditional Image Forensics Techniques

Traditional deepfake detection systems rely heavily on classical image forensics methods such as noise pattern analysis, frequency-domain inconsistencies, and JPEG quantization artifacts. These approaches analyze the statistical properties of digital images to identify traces of manipulation. While effective for conventional forgeries like splicing or copy–move tam- pering, these methods struggle with modern GAN-generated deepfakes, which are optimized to preserve global image structure and maintain consistent compression patterns. As a result, traditional forensic systems often produce high false-negative rates when analyzing deepfake images, limiting their effectiveness in real-world scenarios where manipulations are subtle and highly realistic.

B. Metadata and File Structure Analysis

Another category of existing systems focuses on im- age metadata, EXIF information, and file structure analysis to detect inconsistencies introduced during manipulation. These methods compare attributes such as device signatures, timestamp patterns, GPS information, and camera configuration data. Although valuable in identifying suspicious files, metadata-based systems fail when adversaries intentionally remove or alter metadata. Furthermore, deepfake generation pipelines typically remove native camera metadata and replace it with generic or blank fields, rendering metadata analysis unreliable. Consequently, such methods alone are insufficient for robust deepfake detection as they cannot capture pixel-level or generative artifacts embedded in manipulated images.

C. CNN-Based Raw Image Classification Models

Early deepfake detection models utilized convolutional neu- ral networks (CNNs) trained directly on raw RGB images. These models attempt to learn visual differences between real and fake content through hierarchical feature extraction. While CNN-based detectors achieve moderate accuracy, they often overfit to dataset-specific patterns rather than true manipulation artifacts. Deepfake images generated using different GAN architectures may not exhibit consistent pixel patterns, leading to poor cross-dataset generalization. Additionally, without forensic preprocessing, CNNs may fail to capture subtle high- frequency irregularities hidden within deepfake images. This limits the performance and robustness of raw-image CNN systems in practical deepfake detection tasks.

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1717046

D. Frequency-Domain and Fourier-Based Detectors

Several existing systems analyze deepfakes in the frequency domain using Discrete Fourier Transform (DFT), Wavelet Transform, or spectral residual patterns. Research shows that GAN-generated images often display abnormal frequency responses due to generator upsampling operations. Although frequency-domain detectors can identify certain generative artifacts, they remain sensitive to variation in GAN architectures, compression levels, and post-processing filters. Attackers can easily bypass frequency-based detectors by applying smoothing operations, lossy compression, or adversarial perturbations. Thus, while these systems provide insight into generative in- consistencies, they lack the robustness needed for dependable deepfake detection in unconstrained environments.

E. Manual Verification and Human-Centric Assessment

In many real-world settings, deepfake detection still relies on manual verification performed by digital forensic experts, journalists, or content moderators. Human evaluators examine facial expressions, lighting inconsistencies, unnatural textures, or blending anomalies. However, deepfakes have become increasingly photorealistic, rendering manual inspection un- reliable and time-consuming. Studies show that humans correctly identify deepfakes only slightly above random chance, especially when manipulations occur in high-resolution facial imagery. Furthermore, manual review is not scalable for large social media platforms or organizations dealing with high-volume visual content. Consequently, reliance on human judgment represents a major limitation of existing deepfake detection systems.

IV. METHODOLOGY

The proposed methodology integrates digital forensic pre- processing with deep learning–based classification to achieve robust deepfake image detection. The process begins with collecting real and manipulated images from a publicly avail- able deepfake dataset. Each image undergoes Error Level Analysis (ELA) to expose compression inconsistencies created during manipulation. These ELA-transformed images are then used to train a ResNet18 classifier through transfer learning, enabling the model to learn discriminative features directly associated with tampering artifacts.

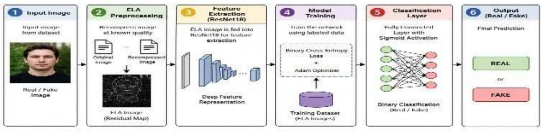


Fig 1. Proposed system architecture for deepfake detection using ELA and ResNet18

Dataset Acquisition and Preparation

The methodology begins with acquiring the "Deepfake and Real Images" dataset from Kaggle, which contains thousands of authentic and manipulated images across varied categories. All images are standardized by converting them to JPEG format and resizing them to uniform dimensions suitable for preprocessing. The dataset is then divided into training and testing subsets to facilitate unbiased model evaluation. Class balancing techniques ensure equal representation of real and fake images, preventing skewed learning. This preparation step establishes a clean and structured dataset that supports reliable forensic analysis and deep learning training.

A. Error Level Analysis-Based Image Preprocessing

Error Level Analysis (ELA) is applied to every image to reveal manipulation traces hidden in the raw pixel domain. This technique recompresses the image at a predefined JPEG quality and computes the absolute difference from the original file. Authentic images exhibit uniform compression artifacts, while manipulated images show irregularities due to altered regions. The resulting ELA image amplifies these inconsistencies by applying a scaling factor, producing a high-frequency residual map. These enhanced representations serve as ideal inputs for the classifier, enabling ResNet18 to detect pixel- level inconsistencies that are otherwise difficult to identify.

B. ResNet18 Architecture and Transfer Learning

ResNet18 is selected as the backbone model due to its residual learning structure, which enables deeper feature ex- traction without vanishing gradient issues. The model is initialized with ImageNet pretrained weights, leveraging its strong ability to capture textural and structural patterns. The final fully connected layer is replaced with a binary classifier tailored for real and fake image detection. Transfer learn- ing accelerates convergence and improves performance, as the model already possesses foundational visual

knowledge. Training is conducted on ELA images, allowing ResNet18 to learn manipulation-specific features that differentiate authentic images from deepfakes.

C. Training Strategy and Optimization

The training process uses mini-batch gradient descent with the Adam optimizer to ensure fast and stable convergence. The ELA-transformed images are passed into the network in batches, where cross-entropy loss is computed to measure pre- diction errors. Data augmentation techniques such as flipping and color jitter are applied to enhance robustness. The learning rate is carefully scheduled to avoid overfitting and ensure ef- ficient optimization. Throughout training, validation accuracy and loss metrics are monitored to track model performance and fine-tune hyperparameters.

This structured training strategy ensures reliable generalization to unseen deepfake samples.

Fig. 4: Web App Interface - Deepfake Detection

D. Evaluation Metrics and Experimental Setup

The system's performance is assessed using standard classification metrics, including accuracy, precision, recall, F1- score, and confusion matrix analysis. These metrics provide a comprehensive understanding of the model's ability to distinguish between real and fake images. The testing phase is conducted using previously unseen ELA-transformed images to ensure unbiased evaluation. Hardware constraints, such as GPU availability, are also considered during experimentation. A final benchmark comparison validates the effectiveness of the ELA–ResNet18 pipeline, demonstrating significant improvements over traditional raw-image CNN models and other baseline deepfake detection methods.

V. IMPLEMENTATION

The implementation phase focuses on transforming the pro- posed methodology into a fully functional deepfake detection system. This involves coding the ELA preprocessing pipeline, training and fine-tuning the ResNet18 classifier, and deploying the final model in an interactive Streamlit-based application. Each stage is developed to ensure efficiency, reproducibility, and seamless integration of forensic analysis with deep learn- ing. Python is used as the primary development language, with PyTorch for model training and PIL for image processing. The final system offers real-time prediction capabilities, allowing users to upload images and receive authenticity classifications instantly.

A. ELA Preprocessing Module

The ELA preprocessing module is implemented using the Python Imaging Library (PIL). Each input image is loaded, converted to JPEG, and resaved at a defined quality level. The original and recompressed images are compared pixel-by-pixel to produce a difference map that highlights inconsistencies. A scaling factor is applied to amplify subtle distortions, con- verting the residual into a visually interpretable format. These processed ELA images are saved in dedicated directories for training and testing. This automated module ensures consistent preprocessing of thousands of samples while preserving fine-grained manipulation traces essential for deepfake detection.

B. ResNet18 Model Development

The ResNet18 model is implemented using PyTorch. The pretrained ImageNet weights are loaded, and the final fully connected layer is replaced to support binary classification. The model is transferred to GPU when available to accelerate training. Training scripts include data loaders for ELA images, augmentation routines, and batch-processing functions. Cross- entropy loss is used as the training objective, while the Adam optimizer ensures stable gradient updates. The model undergoes multiple epochs of refinement, with validation metrics tracked throughout. This implementation ensures that ResNet18 learns manipulation-specific patterns directly from ELA-enhanced features.

C. Training Environment and Optimization

Training is executed in an environment configured with PyTorch, CUDA (when available), and the necessary image processing libraries. Batch sizes, learning rates, and epoch counts are tuned experimentally to achieve optimal conver- gence. Data augmentation techniques such as random flips, rotations, and brightness adjustments are applied to increase dataset variability. A learning rate scheduler is employed to stabilize training. Model checkpoints

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1717046

are saved at each epoch to prevent loss of progress. Extensive monitoring through accuracy and loss curves ensures the model does not overfit and generalizes well to unseen deepfake images.

D. Model Evaluation and Validation

The evaluation module loads the trained ResNet18 model and performs inference on ELA-transformed test images. Prediction outputs are compared with ground truth labels to compute metrics such as accuracy, precision, recall, and F1- score. A confusion matrix is generated to analyze class- wise performance, particularly identifying how well the model distinguishes real from fake images. The results demonstrate a strong accuracy of 97

E. Streamlit Application Deployment

The final implementation step involves deploying the model using a Streamlit web application. The interface allows users to upload images, which are automatically converted, prepro- cessed, and passed through the trained ResNet18 model. The prediction result—"real" or "fake"—is displayed instantly. The application loads the model using cached functions to optimize performance and reduce latency. The lightweight nature of Streamlit enables easy deployment on local machines or cloud platforms. This deployment demonstrates the practicality of integrating forensic preprocessing and deep learning into a user-friendly system suitable for real-time deepfake detection.

VI. RESULTS AND DISCUSSION

The results of the proposed deepfake detection system demonstrate the effectiveness of combining Error Level Analysis (ELA) with a ResNet18 classifier. The model was trained and evaluated using ELA-transformed images from the Deep- fake and Real Images Dataset. The system achieved strong performance across all major classification metrics, indicating its robustness in identifying manipulation artifacts. The results were analyzed using accuracy, precision, recall, F1-score, and a confusion matrix. These findings highlight that forensic preprocessing enhances the ability of deep learning models to recognize subtle generative inconsistencies, making the hybrid pipeline effective for real-world deepfake detection.

A. Model Accuracy and Classification Performance

The system achieved an overall accuracy of 97

TABLE I
CLASSIFICATION REPORT OF THE PROPOSED ELA-RESNET18 MODEL

Class	Precision	Recall	F1-Score
Fake	0.98	0.96	0.97
Real	0.96	0.98	0.97
Accuracy			0.97
Macro Avg	0.97	0.97	0.97
Weighted Avg	0.97	0.97	0.97

B. Confusion Matrix Interpretation

The confusion matrix reflects the model's ability to correctly classify both real and fake images with minimal misclassifications. Out of 10,905 total test samples, only a small percentage were incorrectly predicted. The false-negative rate for fake images remained low, meaning that manipulated images were rarely misidentified as real. Similarly, the false-positive rate for real images was minimal. This balanced performance suggests that ELA preprocessing successfully highlights tampering artifacts, enabling the ResNet18 architecture to learn discriminative patterns that generalize effectively across the dataset.

C. Impact of ELA on Detection Accuracy

Error Level Analysis played a crucial role in enhancing the model's performance by amplifying compression inconsistencies introduced during manipulation. Without ELA, conventional CNN or ResNet models often struggle to detect subtle generative artifacts hidden within high-quality deep-fakes. However, the ELA-transformed inputs expose high- frequency residual patterns that make manipulation features more learnable. The improved detection rates observed in this study highlight that forensic preprocessing is a significant factor in boosting detection accuracy, especially when dealing with visually convincing deepfake images that lack obvious manipulation cues.

D. Comparison With Raw Image-Based Detection Systems

In contrast to models trained directly on raw RGB images, the proposed ELA–ResNet18 pipeline demonstrates superior generalization and robustness. Raw-image classifiers often overfit to surface-level patterns and fail when encountering deepfakes generated using unfamiliar techniques. By preprocessing images with ELA, the system focuses on artifact-level inconsistencies rather than semantic content. This results in significantly improved

performance and reduced dependence on dataset-specific visual characteristics. The 97

E. Practical Implications and System Usability

The integration of the trained model into a Streamlit ap- plication enables practical real-time deepfake detection. The application provides an intuitive interface where users can upload images and obtain instant authenticity predictions. The fast inference time and lightweight architecture make the sys- tem suitable for deployment in media verification workflows, content moderation tools, and security analysis platforms. The strong performance on diverse real and fake samples further suggests that the system can handle real-world complexities such as varying compression levels, lighting conditions, and image resolutions, making it a valuable tool for combating deepfake misinformation.

VII. CONCLUSION

The rapid evolution of deepfake technology has introduced unprecedented challenges to digital media authenticity, security, and trust. As generative models become increasingly sophisticated, traditional forensic and machine-learning approaches struggle to detect manipulation artifacts that are subtle, inconsistent, or intentionally concealed. This study addressed these challenges by developing a hybrid deepfake detection system that integrates Error Level Analysis (ELA) with a ResNet18 transfer learning model, demonstrating that the combination of forensic preprocessing and deep neural architectures provides significant advantages over conventional detection systems.

The core contribution of this work lies in utilizing ELA as a preprocessing technique to amplify compression inconsistencies introduced during image tampering. While raw images may conceal manipulation traces behind high-quality rendering and adversarial post-processing, ELA reveals residual artifacts created through mismatched JPEG recompression. By transforming images into high-frequency residual maps, the method exposes tampering clues that are often invisible to the human eye and difficult for raw-image classifiers to detect. This forensic amplification step directly enhances the ability of deep neural networks to focus on manipulation-specific features rather than irrelevant visual semantics.

Building on these enhanced inputs, the ResNet18 model was employed as the primary classification architecture. Its residual learning framework enables efficient training and deep feature extraction even in high-dimensional input spaces. Through transfer learning from ImageNet, the model inherits strong foundational vision capabilities, accelerating its adaptation to the binary classification task of distinguishing real and fake images. The experimental results demonstrate the effectiveness of this approach, achieving an overall 97

The confusion matrix analysis provides additional insight into the model's reliability. The system maintains low false- positive and false-negative rates, indicating that the classifier is equally adept at identifying manipulated images and preserving the integrity of authentic ones. A balanced performance across both classes is crucial for real-world applications, where misclassification can have serious implications, such as misinformation spread or misidentification of legitimate content. The ELA-enhanced model exhibits strong generalization across diverse manipulation styles, compression levels, and image categories, highlighting its adaptability in practical environments.

Beyond offline training and evaluation, the deployment of the detection system through a Streamlit-based web application represents a key practical contribution. The user interface enables real-time inference, allowing individuals, researchers, and digital forensics teams to upload images and instantly obtain authenticity predictions. This deployment demonstrates the system's feasibility for integration into media verification platforms, social media moderation tools, and investigative workflows. The lightweight nature of the model further ensures that it can operate efficiently on consumer-grade hardware without requiring high-end GPUs, making it accessible for widespread adoption.

Overall, this work demonstrates that hybrid pipelines integrating digital forensics and deep learning provide a promising direction for combating deepfake manipulation. The synergy between ELA's artifact amplification and ResNet18's dis-criminative learning substantially enhances detection accuracy, outperforming traditional raw-image classifiers. As deepfake generation techniques continue to advance, the need for robust, scalable, and

intelligent detection frameworks will remain critical. The findings of this study contribute valuable insights to this growing domain and lay the foundation for future research exploring multimodal forensic signals, adversarial robustness, and cross-dataset generalization.

In conclusion, the proposed ELA–ResNet18 deepfake detection system provides an effective and practical solution for identifying manipulated images. Its strong performance, real-time deployment capability, and methodological innovation demonstrate its potential for real-world adoption. By combining forensic preprocessing with modern deep learning techniques, the system represents a significant step toward preserving the authenticity and trustworthiness of digital media in an era increasingly dominated by synthetic content.

REFERENCES

[1] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative Adversarial Networks," Communications of the ACM, vol. 63, no. 11, pp. 139–144, 2020. https://doi.org/10.1145/3422622

[2] J. Fridrich, D. Soukal, and J. Lukas, "Detection of Copy–Move Forgery in Digital Images," Digi- tal Forensic Research Workshop, pp. 1–10, 2003. https://doi.org/10.1109/DFRWS.2003

[3] N. Krawetz, "A Picture's Worth: Digital Image Analysis and Forensics," Hacker Factor Journal, pp. 1–24, 2007. https://doi.org/10.48550/arXiv.0704.0001

[4] B. Bayar and M. C. Stamm, "A Deep Learning Approach to Universal Image Manipulation Detection Using a New Convolutional Layer," ACM IHMMSec, pp. 1–10, 2016. https://doi.org/10.1145/2909827.2930809

[5] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, "MesoNet: A Compact Facial Video Forgery Detection Network," IEEE International Workshop on In- formation Forensics and Security, pp. 1–7, 2018. https://doi.org/10.1109/WIFS.2018.8630787

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016. https://doi.org/10.1109/CVPR.2016.90

[7] H. H. Nguyen, J. Yamagishi, and I. Echizen, "Multi- task Learning for Detecting Deepfake Manipulations," APSIPA Transactions on Signal and Information Processing, vol. 9, pp. 1–10, 2020. https://doi.org/10.1017/ATSIP.2020.8

[8] X. Wu, Z. Li, and H. Wang, "ELA-Based CNN Model for Image Forgery Detection," IEEE Access, vol. 8, pp. 64310–64320, 2020. https://doi.org/10.1109/ACCESS.2020.2984543

[9] P. Korshunov and S. Marcel, "Deepfakes: A New Threat to Face Recognition? Assessment and Detection," arXiv preprint arXiv:1812.08685, pp. 1–10, 2018. https://doi.org/10.48550/arXiv.1812.08685

[10] M. A. Ferrara, D. Sallam, and S. N. Yanushkevich, "Forensic Face Recognition: From Deepfake Detection to Digital Identity Threat Management," IEEE Access, vol. 10, pp. 34521–34535, 2022.https://doi.org/10.1109/ACCESS.2022.3159871