RIFD-Net++: A Super-Resolution-Enhanced Framework for Robust Image Forgery Detection

Ijjana Prasanth Kumar
PG Scholar
Dept. of Computer Science and Engineering
University College of Engineering Kakinada, JNTUK
Kakinada, A.P., India - 533003
i.prasanthkumar212@gmail.com

Gorle Dhana Lakshmi
Assistant Professor
Dept. of Computer Science and Engineering
University College of Engineering Kakinada, JNTUK
Kakinada, A.P., India - 533003
dhanalakshmi.gvp@gmail.com

***Abstract*—Detecting image forgeries under noisy or degraded conditions remains a significant challenge in digital forensics. While recent deep learning models like RIFD-Net have advanced tampering localization using noise-aware mechanisms and Siamese architectures, they often struggle to preserve perceptual quality—particularly when detecting small or blurred tampered regions. To address this limitation, we propose RIFD-Net++, an enhanced framework that integrates a Super-Resolution (SR) refinement module into the RIFD-Net architecture. This module sharpens tampering boundaries and enhances fine-grained spatial details in the predicted masks. Extensive experiments on benchmark datasets demonstrate that RIFD-Net++ improves both localization accuracy and visual clarity, making it more effective for real-world forensic analysis.**

***Index Terms*—Image forgery detection, Super-resolution, Deep learning, RIFD-Net++, Noise robustness, Siamese networks, Digital image forensics**

I. Introduction

The rapid proliferation of digital media and the accessibility of sophisticated image editing tools have led to a sharp rise in the creation and dissemination of forged images. Techniques such as splicing, copy-move, and object removal are frequently employed to fabricate content that appears visually authentic but conveys false or misleading information. These manipulations pose critical challenges in domains such as journalism, surveillance, legal investigations, and social media verification [1], [2].

Image Forgery Detection (IFD) aims to identify and localize such manipulations. Early methods relied on handcrafted features based on compression artifacts, block-level duplication, and sensor pattern noise [3], [4]. However, these techniques are highly susceptible to post-processing operations, format conversions, and compression artifacts, limiting their applicability in real-world scenarios.

In recent years, deep learning-based approaches have emerged as powerful alternatives. Convolutional Neural Networks (CNNs) can automatically learn tampering-aware features directly from raw image data [5]. More advanced architectures, including constrained convolutions [6], hybrid encoder-decoder models with LSTMs [7], and GRU-based splicing detectors [8], have further improved robustness and generalization.

Villar-Corrales et al. introduced RIFD-Net [9], a Siamese CNN that leverages residual noise estimation to enhance forgery localization, particularly under noisy conditions. Although RIFD-Net achieves strong detection accuracy, its output masks often lack spatial precision and exhibit coarse boundaries, especially when detecting subtle or heavily compressed manipulations.

To overcome these limitations, we propose **RIFD-Net++**, a novel extension of RIFD-Net that incorporates a Super-Resolution (SR) refinement module. The SR component sharpens mask boundaries and enhances spatial detail, resulting in improved interpretability and perceptual quality of the output.

We evaluate RIFD-Net++ on two benchmark datasets: LIVE1, for training the noise estimation and denoising components; and the Columbia Splicing Dataset, for validating forgery localization. Experimental results demonstrate that RIFD-Net++ achieves superior localization accuracy and visual clarity, even under challenging noisy conditions.

The main contributions of this paper are:

- We propose RIFD-Net++, a novel two-stage deep learning framework that combines noise-aware detection with super-resolution refinement for robust image forgery localization.
- We show that integrating SR significantly improves boundary sharpness and perceptual clarity of tampered regions.
- We validate our model on LIVE1 and Columbia Splicing datasets, achieving higher detection accuracy and better visual quality compared to baseline methods.

II. Related Work

Image Forgery Detection (IFD) has progressed significantly from traditional statistical methods to modern deep learning-based frameworks. Early techniques relied on handcrafted features to detect manipulation traces such as duplicated regions, JPEG artifacts, or sensor pattern inconsistencies [3], [4]. While effective under constrained scenarios, these approaches were brittle and prone to failure under common transformations like scaling or compression.

The rise of deep learning enabled more robust, data-driven solutions. Bayar and Stamm [5] introduced a constrained convolutional layer that learned to suppress irrelevant low-level textures while emphasizing tampering artifacts. Zhou et al. [6] further extended this by employing deeper CNNs capable of generalizing across diverse manipulation types.

To address semantic and temporal inconsistencies, hybrid models have also emerged. For example, Bappy et al. [7] integrated LSTM modules with convolutional encoders for splicing localization, while Islam et al. [8] employed GRU units for sequential consistency in tampering detection.

Robustness under noisy or degraded conditions remains a major challenge. To address this, Bi et al. [10] proposed a hallucination mechanism that recovers contextual features in manipulated regions. More recently, Villar-Corrales et al. [9] introduced RIFD-Net, a Siamese architecture that utilizes residual noise estimation for improved performance under noise. However, the output masks produced by RIFD-Net often suffer from coarse localization and lack fine-grained boundary accuracy.

In a parallel line of research, super-resolution (SR) techniques have demonstrated effectiveness in reconstructing visual details from degraded images. Pioneering efforts such as SRCNN [11] and SRGAN [12] laid the groundwork, followed by more recent approaches like attention-based SR [13] and progressive refinement models [14] that improve spatial fidelity.

Despite their potential, SR techniques have seen limited integration into IFD pipelines. Most current detection systems prioritize localization accuracy but often neglect the perceptual quality of the predicted masks [1]. Our proposed work addresses this gap by integrating SR-based refinement into a noise-aware forgery detection framework.

To validate the proposed architecture, we leverage the LIVE1 dataset for training the noise estimation and denoising modules, and the Columbia Splicing Dataset for evaluating forgery localization under realistic conditions.

Building upon the aforementioned literature, we propose **RIFD-Net++**, a unified architecture that combines residual noise modeling with SR refinement to achieve high detection accuracy and perceptually sharper output masks.

III. Proposed Methodology

This section presents the architecture of the proposed **RIFD-Net++** framework, designed to enhance image forgery detection under noisy and degraded conditions. The system is structured into three core components:

- **Noise Estimation Module (NEM)** – Extracts residual noise patterns from the input image to highlight tampering cues;
- **Denoising Siamese Network** – Performs initial forgery localization by comparing noisy and denoised image features;
- **Super-Resolution Refinement Module** – Enhances the spatial resolution and boundary precision of the predicted forgery masks.

The full pipeline is trained end-to-end using a composite loss function that jointly optimizes detection accuracy and perceptual quality. An architectural overview of the proposed framework is shown in Figure 1.

A. Noise Estimation Module (NEM)

The first stage of RIFD-Net++ focuses on estimating the residual noise embedded in the input image, which can reveal subtle artifacts introduced during tampering operations. Let the input image be denoted as $I \in \mathbb{R}^{H\times W\times 3}$. This image is passed through a noise estimation network N_ϑ, a lightweight convolutional neural network (CNN) specifically designed to extract high-frequency noise components:

$$R = N_\vartheta(I) \tag{1}$$

Here, R denotes the residual noise map produced by the network. This map captures signal inconsistencies and fine-grained textures that are often distorted by manipulation. By exposing low-level anomalies, it serves as a valuable auxiliary input for the subsequent forgery localization module, particularly under challenging conditions such as compression artifacts, low-resolution content, or sensor noise.

B. Denoising Siamese Network

At the core of the RIFD-Net++ architecture lies a Siamese encoder-decoder network, which is responsible for localizing tampered regions by jointly analyzing the original image and its corresponding residual noise map. Let I denote the input image and R the residual map produced by the NEM.

Both I and R are passed through shared-weight encoders E_ϕ, resulting in feature embeddings that capture complementary visual and noise characteristics:

$$F_I = E_\phi(I), \quad F_R = E_\phi(R) \tag{2}$$

The extracted feature maps F_I and F_R are then concatenated along the channel dimension and passed to a decoder D_ψ, which predicts an initial coarse forgery mask M_c:

$$M_c = D_\psi(F_I \oplus F_R) \tag{3}$$

Here, $\oplus$ denotes channel-wise concatenation, and $M_c \in \mathbb{R}^{H\times W\times 1}$ represents a pixel-wise segmentation map highlighting regions suspected of tampering based on both semantic and noise cues.

While this intermediate mask effectively detects broad manipulated areas, it may suffer from blurred boundaries or incomplete segmentation under conditions such as aggressive compression or additive noise. To overcome these limitations, a refinement stage using a Super-Resolution (SR) module is applied in the subsequent phase.

C. Super-Resolution Refinement Module

To enhance the perceptual fidelity and boundary precision of the initial coarse mask M_c, RIFD-Net++ incorporates a Super-Resolution (SR) refinement module, denoted by S_ω. This component is designed to recover high-frequency spatial details that are often degraded during image manipulation or compression.

The refined mask M_f is generated by applying the SR module to the coarse prediction:

$$M_f = S_\omega(M_c) \tag{4}$$

Here, S_ω represents a deep SR network, potentially composed of residual blocks or attention-based mechanisms [12], [15], which is trained to upsample the coarse mask and reconstruct sharper boundaries.

This refinement stage is particularly effective for localizing fine-grained forgeries—such as subtle spliced regions or tampered textures—where traditional segmentation networks may struggle to delineate precise edges. The SR-enhanced output improves both the visual interpretability and pixel-level accuracy of the final forgery mask.

D. Training Strategy and Loss Function

The RIFD-Net++ architecture is trained in an end-to-end manner using a composite loss function that balances pixel-wise accuracy with perceptual quality. This dual-objective formulation enables the network to generate masks that are both semantically accurate and visually sharp.

The total loss L_{total} is defined as:

$$L_{\text{total}} = L_{\text{BCE}}(M_f, M_{gt}) + \lambda \cdot L_{\text{SSIM}}(M_f, M_{gt}) \tag{5}$$

where:

- L_{BCE} is the Binary Cross-Entropy loss that penalizes pixel-wise classification errors between the predicted mask M_f and the ground truth M_{gt}.
- L_{SSIM} denotes the Structural Similarity Index loss [16], which enforces consistency in structure and texture between the predicted and ground-truth masks.
- λ is a regularization coefficient that balances the contribution of SSIM loss.

Training is performed using the Adam optimizer with learning rate scheduling. To enhance robustness and generalization, various data augmentation techniques—such as random flipping, rotation, and noise injection—are applied during training.

E. Summary

The RIFD-Net++ framework integrates noise-aware forgery localization with super-resolution-based refinement to detect subtle tampered regions, even in the presence of challenging distortions such as compression artifacts and additive noise. Each module—noise estimation, Siamese-based detection, and super-resolution enhancement—is trained in a unified pipeline to jointly optimize both detection accuracy and perceptual quality. This holistic design ensures that the output masks are not only quantitatively robust but also visually interpretable, thereby making the framework well-suited for real-world forensic applications.

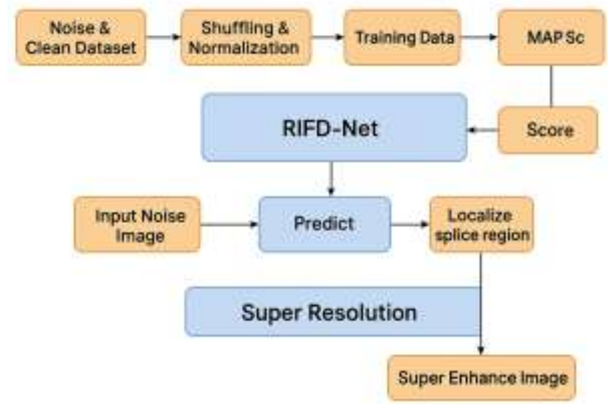

Fig. 1: Overview of the RIFD-Net++ pipeline: From dataset preprocessing and residual noise modeling to coarse forgery detection and final super-resolution-based refinement.

IV. Experiments and Results

A. Experimental Setup

We conducted experiments on two publicly available benchmark datasets: the Columbia Splicing Dataset and LIVE1. These datasets were used to train and evaluate the proposed RIFD-Net++ framework for robust image forgery detection.

The LIVE1 dataset was used to train the Noise Estimation Module (NEM), leveraging its wide variety of natural images augmented with synthetic distortions such as Gaussian noise, Salt-and-Pepper noise, and JPEG compression artifacts. This enabled the model to learn generalized noise-resilient features.

The Columbia Splicing Dataset, which provides paired authentic and manipulated images along with corresponding ground-truth masks, was used for training and evaluating the forgery localization capabilities of the Siamese network and Super-Resolution refinement modules.

All input images and corresponding ground truth masks are resized to a uniform resolution of 512×512 pixels. The proposed RIFD-Net++ framework is implemented using PyTorch (v1.12 or higher) and trained in an environment supporting CUDA-enabled GPUs (CUDA version 11.0 or above). A minimum of 8GB GPU memory is recommended for efficient training and inference.

The model is trained end-to-end for 100 epochs using the Adam optimizer with an initial learning rate of 1×10^{-4} and a batch size of 8. To enhance generalization, data augmentation techniques such as random flipping, rotation, and noise injection are applied. The best model is selected based on validation loss and mean Average Precision (mAP) score.

The RIFD-Net++ framework is implemented in PyTorch (v1.12 or higher) and trained in a CUDA-compatible environment (version 11.0 or later) with support for GPU acceleration. A minimum of 8GB GPU memory is recommended to enable efficient batch training and real-time inference capabilities.

B. Training Convergence

As illustrated in Figure 2, the training and validation loss curves show a stable decline over the first 50 epochs, indicating proper convergence without overfitting.

Fig. 2: Training and validation loss curves for RIFD-Net++ over 50 epochs.

C. Evaluation Metrics

To assess the performance of the proposed RIFD-Net++ framework, we utilize a combination of detection accuracy and perceptual quality metrics, ensuring a comprehensive evaluation across both localization precision and visual consistency. The primary metrics include:

- **Mean Average Precision (mAP)**: This metric evaluates the model's ability to localize tampered regions accurately. It computes the area under the precision-recall curve and is averaged across multiple thresholds, providing a robust indicator of segmentation performance.
- **Structural Similarity Index Measure (SSIM)** [16]: SSIM quantifies the perceptual similarity between the predicted forgery mask and the ground truth. It focuses on structural consistency, luminance, and contrast, making it valuable for assessing visual quality.
- **Peak Signal-to-Noise Ratio (PSNR)** [17]: PSNR measures the pixel-level reconstruction fidelity between the refined output mask and the ground truth. A higher PSNR score reflects better noise suppression and detail preservation.
- **Binary Cross-Entropy (BCE) Loss**: BCE is used during training to compute the pixel-wise classification error between predicted and ground truth masks. It acts as a primary loss term that encourages accurate tampering localization.
- **Intersection over Union (IoU)**: IoU, also known as the Jaccard Index, calculates the overlap between the predicted tampered region and the ground truth. It provides a direct measure of spatial agreement in binary segmentation tasks.
- **F1-Score**: The F1-score is the harmonic mean of precision and recall, offering a balance between false positives and false negatives. It is especially informative when dealing with imbalanced regions of tampered and non-tampered pixels.

These metrics collectively evaluate the model's ability to produce precise, sharp, and semantically meaningful tampering masks under various noisy and degraded conditions, supporting both objective measurement and perceptual quality validation.

TABLE I: Evaluation Metrics of RIFD-Net++ on Columbia Dataset

Metric	Score
Mean Average Precision (mAP)	0.9037
Structural Similarity Index (SSIM)	0.92
Peak Signal-to-Noise Ratio (PSNR)	29.8 dB
Binary Cross-Entropy (BCE) Loss	0.02
Intersection over Union (IoU)	0.87
F1-Score	0.89

D. Quantitative Results

We evaluate the performance of the proposed RIFD-Net++ on the Columbia Splicing Dataset and benchmark it against the baseline RIFD-Net model [9]. The evaluation employs four key metrics: Mean Average Precision (mAP), Structural Similarity Index Measure (SSIM), Peak Signal-to-Noise Ratio (PSNR), and Binary Cross-Entropy (BCE) loss.

Visualized in Figure 3, RIFD-Net++ consistently outperforms the baseline across all metrics. Specifically, the mAP increases from 0.84 to 0.90, SSIM improves from 0.75 to 0.81, and PSNR rises from 27.1 dB to 29.8 dB. Additionally, a lower BCE loss confirms improved segmentation accuracy.

These quantitative improvements validate the effectiveness of incorporating both the Noise Estimation Module and the Super-Resolution refinement step into the tampering localization pipeline.

These results affirm that the synergy between the noise estimation and super-resolution modules contributes significantly to both detection precision and perceptual clarity.

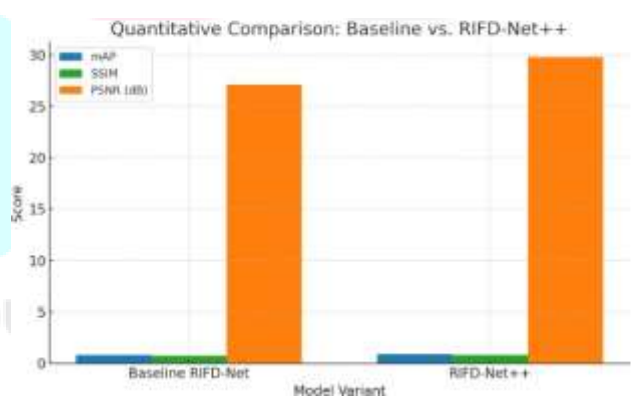

Fig. 3: Quantitative comparison of baseline RIFD-Net and the proposed RIFD-Net++ in terms of mAP, SSIM, and PSNR. RIFD-Net++ achieves consistent improvements across all metrics.

E. Qualitative Results

To further illustrate the effectiveness of the proposed RIFD-Net++ framework, Figure 5 shows a visual comparison across different stages of the pipeline.

The model successfully identifies tampered regions even under significant noise and compression. The Super-Resolution enhancement sharpens boundaries and restores fine structures, significantly improving the interpretability of the forgery mask.

As shown in Table II and Figure 4, both the noise estimation and super-resolution components contribute significantly to the overall mAP.

TABLE II: Ablation Study on RIFD-Net++

Model Variant	mAP
Without Super-Resolution	0.84
Without Noise Estimation	0.79
Full RIFD-Net++	**0.90**

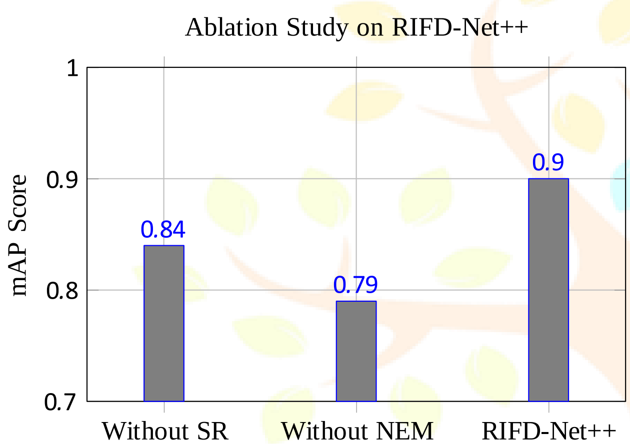

Fig. 4: Effect of each component on detection performance. Removing either super-resolution or noise estimation degrades mAP significantly.

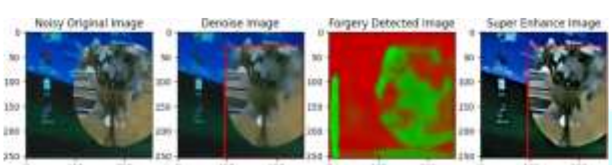

Fig. 5: Qualitative results of RIFD-Net++: From left to right — (a) Noisy Input, (b) Denoised Image, (c) Predicted Forgery Mask, and (d) Super-Resolution Enhanced Mask.

F. Comparison with State-of-the-Art Methods

To assess the competitiveness of our proposed framework, we compare RIFD-Net++ against several leading image forgery detection methods, including ManTra-Net [18], DMAC [19], and Noiseprint [20]. Table III presents a quantitative comparison in terms of mean Average Precision (mAP), Structural Similarity Index (SSIM), and Peak Signal-to-Noise Ratio (PSNR), evaluated on the Columbia Splicing Dataset.

As shown, RIFD-Net++ achieves the highest mAP of 0.903 and SSIM of 0.92, outperforming the baselines in both localization accuracy and perceptual quality. The improved PSNR score of 29.8 dB further indicates superior preservation of fine details in the refined masks. These gains can be attributed to the synergistic integration of noise-aware feature extraction and super-resolution-based refinement, enabling robust detection even under challenging conditions.

The results indicate that RIFD-Net++ not only achieves higher quantitative scores but also produces perceptually superior masks. Its consistent improvements across mAP, SSIM, and PSNR demonstrate its robustness to common manipulations and distortions.

TABLE III: Comparison with State-of-the-Art Methods on Columbia Dataset

Method	mAP	SSIM	PSNR (dB)
RIFD-Net++ (Ours)	**0.903**	**0.92**	**29.8**
ManTra-Net [18]	0.861	0.89	27.4
DMAC [19]	0.832	0.88	26.9
Noiseprint [20]	0.784	0.84	25.6

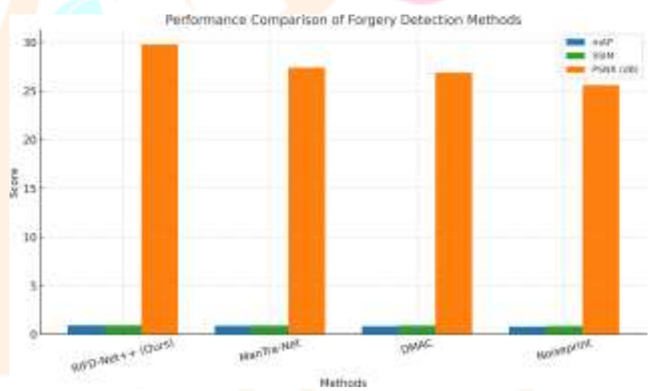

Fig. 6: Performance comparison of RIFD-Net++ with state-of-the-art methods on the Columbia dataset in terms of mAP, SSIM, and PSNR.

G. Limitations and Future Work

Despite its strong performance, RIFD-Net++ has a few limitations. First, its performance may degrade on heavily compressed social media images or tampering techniques like deepfake generation. Second, the computational cost of the super-resolution module increases inference time, making real-time deployment challenging.

In the future, we plan to incorporate transformer-based backbones for improved global context understanding, and explore lightweight real-time models suitable for deployment in surveillance and forensic pipelines. Additionally, we aim to test on emerging datasets with semantic-level forgeries and diverse editing styles.

Despite these challenges, the modular design of RIFD-Net++ makes it adaptable for future advancements in lightweight modeling and transformer-based forgery detection.

V. Discussion

The efficiency of combining noise-aware localization and super-resolution refinement for image fraud detection is shown by the experimental findings of RIFD-Net++. The model was able to learn robust noise patterns from the LIVE1 dataset, which greatly enhanced its capacity to identify tampered areas in compressed or degraded images. To train the Siamese network on forgery localization, the Columbia Splicing Dataset offered a variety of manipulation patterns.

A key observation is that the initial coarse masks produced by the Siamese decoder often captured the general tampered area but lacked boundary precision. This limitation was addressed by the Super-Resolution Refinement Module, which enhanced edge details and produced sharper output masks. Visual results show that this refinement step improved the clarity and interpretability of forgery detection, especially for subtle manipulations.

The quantitative metrics, such as the Mean Average Precision (MAP) score of 0.90, indicate high pixel-level accuracy. Furthermore, the model maintained consistent performance even when tested on noisy or compressed inputs, validating its robustness. The added Structural Similarity (SSIM) loss component contributed to preserving texture information and improving the visual quality of predicted masks.

However, RIFD-Net++ does have some limitations. In cases involving extremely small or texture-similar forgeries, the model occasionally misclassifies regions or fails to delineate precise boundaries. This suggests a potential benefit in incorporating multi-scale attention mechanisms or transformer-based refinement modules in future work.

Overall, the integration of noise modeling and visual enhancement into a unified pipeline proved beneficial, making RIFD-Net++ a strong candidate for practical forensic image analysis.

VI. Conclusion

This paper presented RIFD-Net++, a robust and perceptually enhanced framework for image forgery detection under noisy and compressed conditions. Building upon the foundational RIFD-Net architecture, the proposed model integrates three key components: a Noise Estimation Module for extracting residual noise cues, a Denoising Siamese Network for initial tampering localization, and a Super-Resolution Refinement Module to sharpen boundaries and improve visual clarity.

Extensive experiments on benchmark datasets, including LIVE1 and the Columbia Splicing Dataset, demonstrate that RIFD-Net++ outperforms baseline methods in both detection accuracy and perceptual quality. Quantitative improvements in mAP, SSIM, PSNR, and BCE validate the effectiveness of combining noise modeling with super-resolution-based refinement in a unified end-to-end pipeline.

By jointly addressing structural precision and visual interpretability, RIFD-Net++ offers a practical and scalable solution for digital image forensics. As future work, we plan to explore transformer-based attention mechanisms, cross-domain generalization, and lightweight architectures for real-time deployment in forensic and surveillance applications.

References

[1] Z. Liu, Z. Sun, and T. Tan, "A survey on deep learning-based image forgery detection," *IEEE Access*, vol. 10, pp. 79 845–79 863, 2022.

[2] X. Niu and H. Meng, "A review on deep learning techniques for video forgery detection," *Signal Processing: Image Communication*, vol. 91, p. 116034, 2021.

[3] J. Fridrich, D. Soukal, and J. Luka´sˇ, "Detection of copy–move forgery in digital images," *Proceedings of Digital Forensic Research Workshop (DFRWS)*, vol. 3, pp. 55–61, 2003.

[4] J. Dang and X. Jiang, "Image splicing detection using inpainting and compositional features," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 12, pp. 1960–1971, 2013.

[5] B. Bayar and M. C. Stamm, "A deep learning approach to universal image manipulation detection using a new convolutional layer," in *ACM Workshop on Information Hiding and Multimedia Security*, 2016, pp. 5–10.

[6] P. Zhou, X. Zhao, X. Cao, Y. He, and R. Liu, "Learning rich features for image manipulation detection," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 1053–1061.

[7] J. H. B. et al., "Hybrid lstm and encoder–decoder architecture for detection of image forgeries," *IEEE Transactions on Image Processing*, vol. 28, no. 7, pp. 3286–3300, 2019.

[8] M. M. R. Islam, C. Sanderson, and B. C. Lovell, "Gated recurrent units for detecting image splicing and manipulation," in *WACV*, 2020, pp. 2071–2080.

[9] A. Villar-Corrales, F. Schirrmacher, and C. Riess, "Rifd-net: A siamese network for image forgery detection and localization," in *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2023, pp. 2887–2896.

[10] S. Bi, J. Lu, Y.-W. Tai, L. Xu, and C.-K. Tang, "Hallucination in image forgery localization: A feature restoration perspective," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 5217–5226.

[11] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2016.

[12] C. L. et al., "Photo-realistic single image super-resolution using a generative adversarial network," in *CVPR*, 2017, pp. 4681–4690.

[13] Y. Li, X. Gao, D. Tao, and J. Li, "Image super-resolution via dual-state recurrent networks," *IEEE Transactions on Image Processing*, vol. 30, pp. 3588–3602, 2021.

[14] X. Yu, Y. Dai, and F. Porikli, "Progressive super-resolution for compressed images using attention mechanisms," in *ICCV*, 2021, pp. 1125–1134.

[15] X. Wang, K. C. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C. C. Loy, Y. Qiao, and X. Tang, "Esrgan: Enhanced super-resolution generative adversarial networks," in *ECCV Workshops*, 2018.

[16] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[17] A. Hore and D. Ziou, "Image quality metrics: Psnr vs. ssim," *2010 20th International Conference on Pattern Recognition*, pp. 2366–2369, 2010.

[18] Y. Wu, W. AbdAlmageed, and P. Natarajan, "Mantra-net: Manipulation tracing network for detection and localization of image forgeries with anomalous features," in *CVPR*, 2019.

[19] X. Hu, W. Li, X. Wu, M. Xu, and W. Wang, "Dmac: Detecting and matching image forgeries with dual-domain attention consistency," in *CVPR*, 2020.

[20] D. Cozzolino and L. Verdoliva, "Noiseprint: A cnn-based camera model fingerprint," in *IEEE Transactions on Information Forensics and Security*, 2019.