e-ISSN 2549-7472
Edumatic: Jurnal Pendidikan Informatika
Volume 10, Issue 3, **559-568.**
https://e-journal.hamzanwadi.ac.id/index.php/edumatic
DOI: 10.29408/edumatic.v10i3.37027

Comparative Evaluation of Direct and Residual Self-Attention in Spatial-Frequency ResNet-18 for AI-Generated Image Detection

Habibi [1,*], Robert Marco [1]

[1] Universitas Amikom Yogyakarta, Indonesia
[*] Corresponding author: Habibi, Universitas Amikom Yogyakarta, Indonesia
✉ habibi@students.amikom.ac.id

Copyright: © 2026 by the authors

Received: 08 October 2026 | Revised: 31 October 2026 | Accepted: 8 November 2026 | Published: 28 November 2026

Abstract
The effectiveness of attention mechanisms in spatial-frequency representations for AI-generated image detection remains insufficiently understood, particularly under controlled training conditions. This study comparatively evaluates Direct and Residual Self-Attention within a spatial-frequency ResNet-18 architecture. Using 60,000 development images and a separate 6,000-image holdout set from two public datasets, the initial experiments assessed spatial features, FFT-based frequency features, feature fusion, and attention integration. The results showed that Residual Self-Attention achieved the highest initial holdout performance, with 97.28% accuracy and 97.26% F1-score. However, additional controlled experiments using three random seeds and matched training budgets revealed that continued training without attention achieved the highest mean validation F1-score (97.26%), followed by Direct Self-Attention (97.19%) and Residual Self-Attention (97.14%). The residual-gate ablation further showed that attention integration did not consistently improve classification performance. These findings indicate that the initial performance advantage of Residual Self-Attention was not reproduced under controlled training conditions. This study provides empirical evidence that architectural improvements should be evaluated against matched-training baselines to distinguish genuine performance gains from differences attributable to training procedures.

Keywords: ai-generated image detection; comparative evaluation; feature fusion; resnet-18; self-attention

To cite this article: Habibi, H., & Marco, R. (2026). Comparative Evaluation of Direct and Residual Self-Attention in Spatial-Frequency ResNet-18 for AI-Generated Image Detection. *Edumatic: Jurnal Pendidikan Informatika*, *10*(3), 559–568. https://doi.org/10.29408/edumatic.v10i3.37027

INTRODUCTION

The growing realism of AI-generated images challenges digital-content authentication and multimedia forensics. Advances in generative models have made synthetic images increasingly difficult to distinguish from authentic photographs, particularly when visual differences are subtle. These developments raise concerns about the reliability of detection methods across different image sources and generation techniques. Zhang et al. (2025) highlighted detector generalization as an important challenge in identifying generated visual content. Consequently, evaluating what classification models learn remains necessary as generative technologies continue to evolve.

Reliable AI-generated image detection requires more than achieving high accuracy on a particular dataset. Classification errors must also be considered because incorrectly identifying synthetic content as authentic may affect decisions based on digital evidence. Image

Edumatic: Jurnal Pendidikan Informatika is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

authentication additionally involves different forensic objectives, including determining whether an image is generated and identifying its possible source. Namani et al. (2025) investigated detection and attribution within a synthetic-image forensic framework, illustrating these distinct objectives. This study focuses on binary classification and examines how alternative feature-processing strategies affect detection performance.

Convolutional neural networks (CNNs) provide a foundation for extracting discriminative visual characteristics directly from image pixels. Their hierarchical representations allow classification models to learn patterns associated with authentic and synthetic content. Bird and Lotfi (2024) demonstrated CNN-based classification and Grad-CAM visualization on the CIFAKE dataset. However, spatial information alone may not represent every characteristic associated with image generation. This motivates examining additional image representations and whether integrating them can improve an established spatial classification baseline.

Frequency-domain analysis offers another perspective by describing image characteristics through their spectral representations. Such information may complement spatial features, although its effectiveness depends on the transformation and classification approach employed. Zhang et al. (2025) investigated a Discrete Cosine Transform (DCT)-based network for synthetic-image detection and attribution, demonstrating the relevance of frequency information in digital forensics. The present study instead uses Fast Fourier Transform (FFT) magnitude representations. Evaluating spatial and frequency features separately and in combination is therefore important for understanding their contributions within the selected architecture.

Combining multiple image representations is an established strategy for improving the information available to a detector. Spatial features describe visual content, while frequency representations provide another view of image characteristics. Man and Cho (2026) investigated integrating CLIP-derived spatial-semantic information and wavelet features through Transformer-based processing. Although this approach differs from the dual-branch ResNet-18 examined here, it demonstrates the relevance of multi-domain representation learning. The remaining question is whether attention processing offers an additional benefit after spatial and frequency information have already been combined.

We have investigated attention mechanisms to model relationships among extracted image features. Li et al. (2026) proposed SIDWA, integrating wavelet-based processing with deformable cross-attention for synthetic-image detection. This approach illustrates how attention can be incorporated into a frequency-aware detection architecture, although its design differs from feature-level Self-Attention applied after concatenation. Such architectural differences make direct performance comparisons difficult. Accordingly, attention effectiveness should be examined within a common backbone rather than inferred from results obtained using different feature representations and training procedures.

Attention integration may also affect the original features produced by a classification model. Gunukula et al. (2025) incorporated Squeeze-and-Excitation modules into a ResNet-based detector to recalibrate feature channels. Although channel recalibration differs from the Self-Attention mechanism evaluated here, both approaches modify learned representations before classification. Additional processing does not necessarily improve performance when the existing features are already discriminative. This raises the question of whether attention should directly process the fused representation or preserve its original information through an additional pathway.

Residual connections offer a mechanism to retain existing features while incorporating attention-derived information. Reza et al. (2026) combined attention and residual mechanisms within a CNN-based framework for deepfake facial image detection, demonstrating their application to a related classification task. However, that architecture differs from the feature-

level Residual Self-Attention considered in this study. Preserving the original spatial-frequency representation therefore remains a hypothesis whose effectiveness requires experimental evaluation rather than an assumption that residual integration will consistently improve detection performance.

Recent studies have demonstrated the potential of combining multiple image representations for synthetic-image detection. Man et al. (2026) examined interactions among spatial, frequency, and wavelet representations, highlighting the relevance of cross-domain feature processing. However, these findings do not directly establish whether attention integration provides additional benefits when applied to an existing spatial-frequency representation. In particular, the relative effectiveness of Direct Self-Attention, which processes fused features through attention, and Residual Self-Attention, which retains the original representation through a residual pathway, requires evaluation under comparable training conditions. This study addresses this research gap through two experimental stages. The initial experiments compare spatial features (S0), frequency features (S1), spatial-frequency fusion (S2), Direct Self-Attention (S3), and Residual Self-Attention (S4). Controlled experiments (C0–C2) then evaluate attention integration against continued training without attention, using matched training budgets and three random seeds. Residual-gate ablations (A0–A1) further examine the contribution of attention-derived features. The principal contribution is an empirical comparison that distinguishes initial performance differences from improvements consistently observed under controlled training conditions.

METHOD

This study employed a two-stage comparative experimental design to evaluate attention integration in spatial-frequency ResNet-18 for AI-generated image detection. The initial stage examined five configurations (S0–S4), comprising spatial features, frequency features, spatial-frequency fusion, Direct Self-Attention, and Residual Self-Attention. Because the initial S4 model was developed from a selected S2 checkpoint and followed a different training history, we conducted additional controlled experiments to evaluate the effects of attention under comparable training conditions. These experiments compared continued training without attention (C0), Direct Self-Attention (C1), and Residual Self-Attention (C2), followed by residual-gate ablations (A0–A1). This design enabled us to evaluate attention-related performance differences while accounting for the influence of continued optimization.

We used two publicly available datasets: 140k Real and Fake Faces and AI-Generated Images vs Real Images. We assembled a balanced development set of 60,000 images from 42,000 and 18,000 images, respectively, with equal class proportions. We partitioned the data into 48,000 training (80%), 6,000 validation (10%), and 6,000 testing (10%) images. We reserved an additional balanced holdout set of 6,000 images, comprising 3,000 real and 3,000 AI-generated samples. The reported evaluations included holdout results for S0–S4 and validation results for C0–A1; we did not report separate metrics for the development test partition. Because the holdout came from the same public dataset sources, we did not treat it as an independent cross-dataset benchmark.

We resized all images to 224 × 224 pixels and processed them into normalized RGB and FFT magnitude representations. Separate ResNet-18 branches extracted spatial and frequency features, which were concatenated before classification. The fusion baseline (S2) classified these features without attention, while Direct Self-Attention (S3) processed the fused representation through an attention pathway. Residual Self-Attention (S4) retained the original representation while incorporating attention-derived features. Equation (1) expresses the residual operation.

$Z_{out} = Z_{dual} + \alpha \cdot A(Z_{dual}).$	(1)

Here, Z_{dual} represents the fused spatial-frequency features, A denotes the attention transformation, and α controls the contribution of attention-derived information. Figure 1 illustrates the feature-processing pathways of the evaluated architectures.

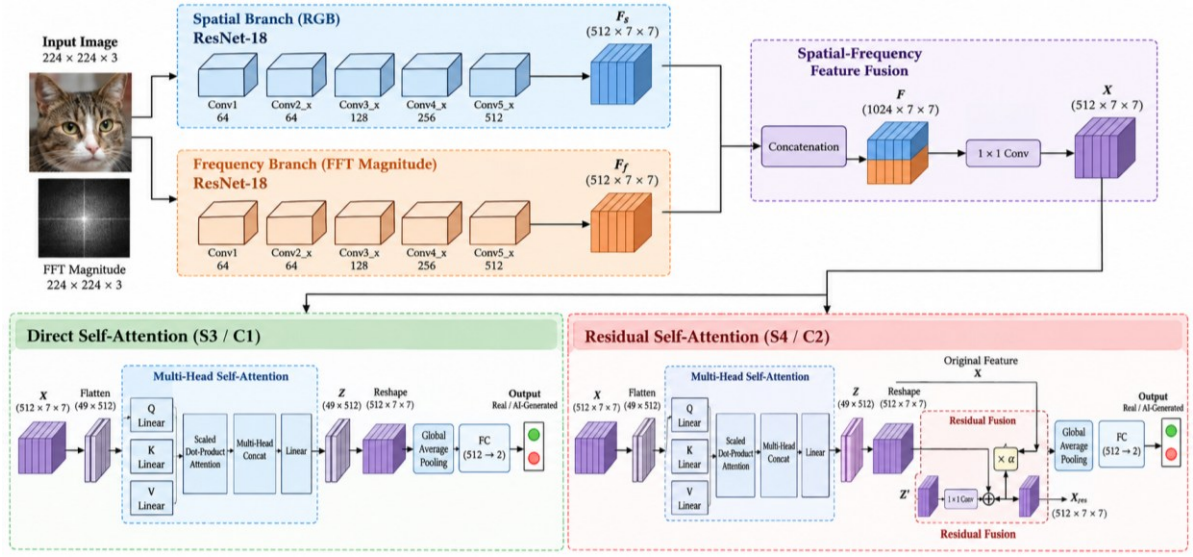

Figure 1. Spatial-frequency ResNet-18 architecture with direct and residual self-attention integration.

Figure 1 shows the shared spatial-frequency feature extraction process and the distinction between direct and residual attention integration. The initial configurations evaluate the contribution of each feature-processing stage, while the controlled configurations examine whether attention provides additional benefits beyond continued optimization. The gate ablations further assess the effect of disabling the attention contribution or fixing its coefficient at one. These experimental relationships and their objectives are summarized in Table 1.

Table 1. Experimental configurations and objectives

Configuration	Strategy	Objective
S0–S1	RGB only / FFT only	Single-domain baselines
S2	Spatial-frequency fusion	Fusion baseline
S3–S4	Direct/residual attention	Initial comparison
C0–C2	No attention / direct/trainable residual	Matched-training controls
A0–A1	Fixed gate: 0 / 1	Residual-gate ablation

Initial training used AdamW with a learning rate of 1×10^{-4}, weight decay of 1×10^{-4}, cross-entropy loss, and a batch size of 32, for up to 50 epochs with early-stopping patience of seven. The controlled experiments initialized their shared ResNet-18 backbone from a selected S2 checkpoint and evaluated C0–C2 and A0–A1 using seeds 42, 123, and 2026. Each of the 15 runs used six epochs (9,000 optimizer updates), with a physical batch size of eight and four gradient accumulation steps. Checkpoints were selected based on validation F1-score. We evaluated accuracy, precision, recall, F1-score, and confusion matrices, with AI-generated images as the positive class. We summarized controlled validation results using mean F1-score and sample standard deviation, separately from the initial holdout results.

RESULTS AND DISCUSSION

Results

The initial evaluation compared S0–S4 using a balanced holdout set of 6,000 images. As presented in Table 2, S0 achieved 0.9610 accuracy and 0.9603 F1-score, whereas the frequency-only S1 produced substantially lower values. Combining both representations in S2 increased accuracy to 0.9700 and F1-score to 0.9698, improving by 0.90 and 0.95 percentage points over S0. Direct Self-Attention in S3 did not improve the fusion baseline, while S4 achieved the highest numerical accuracy and F1-score. However, S4's advantage over S2 was limited to 0.28 percentage points in accuracy and was achieved under a different training history.

Table 2. Initial classification performance on the holdout set

Model	Accuracy	Precision	Recall	F1-score
S0	0.9610	0.9772	0.9440	0.9603
S1	0.6430	0.5996	0.8607	0.7068
S2	0.9700	0.9751	0.9647	0.9698
S3	0.9680	0.9721	0.9637	0.9679
S4	0.9728	0.9820	0.9633	0.9726

The S4 confusion matrix provides additional information about classification errors. Among 3,000 AI-generated images, 2,890 were classified correctly, and 110 were misclassified as real. Among 3,000 authentic images, 2,947 were classified correctly, and 53 were misclassified as generated. These results correspond to a false-negative rate of 3.67% and a false-positive rate of 1.77%. The difference indicates that generated images were more frequently accepted as authentic than real images were incorrectly flagged as synthetic. Thus, interpret the high aggregate accuracy alongside the error distribution rather than treating it as complete reliability in authenticity assessment.

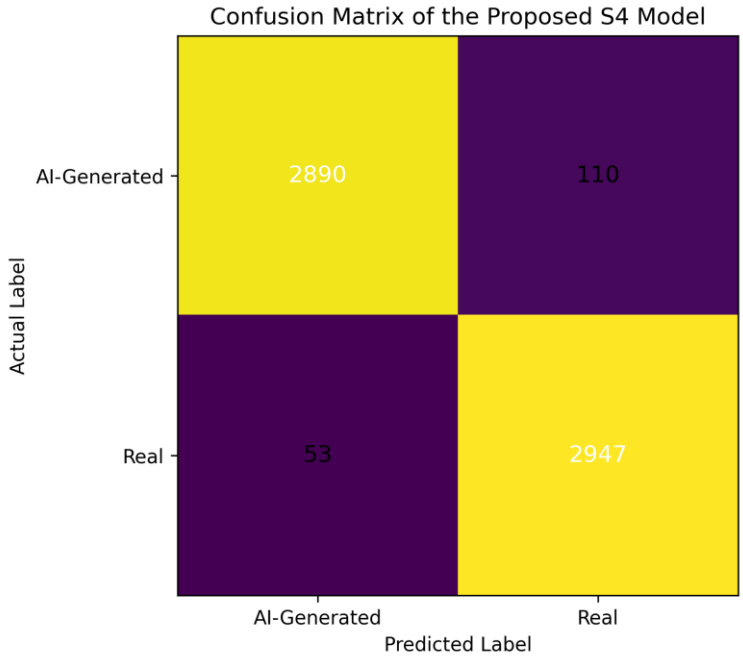

Figure 2. Confusion matrix of the initial S4 configuration on the reserved holdout set

The controlled experiments evaluated five configurations across three random seeds. As shown in Table 3, continued training without attention (C0) achieved the highest mean validation F1-score (0.972567), followed by Direct Self-Attention (C1) at 0.971879 and Residual Self-Attention (C2) at 0.971358. The difference between C0 and C2 was approximately 0.12 percentage points, indicating no improvement from residual attention under the matched training protocol. The fixed-zero configuration (A0) reproduced C0's aggregate metrics, whereas the fixed-one configuration (A1) yielded the lowest mean F1-score. Thus, S4's initial numerical advantage did not carry over to the controlled validation experiments.

Table 3. Controlled validation performance across three random seeds

Model	Mean accuracy	Mean F1-score	F1 SD
C0	0.972667	0.972567	0.000350
C1	0.971889	0.971879	0.000950
C2	0.971444	0.971358	0.000171
A0	0.972667	0.972567	0.000350
A1	0.970944	0.970837	0.000580

Performance variability differed across the controlled configurations. Although C2 exhibited the smallest F1-score standard deviation among the principal configurations, its mean performance remained below C0 and C1. Thus, lower variability did not correspond to higher classification performance. Figure 3 shows the mean F1-scores and standard deviations, illustrating the relatively small differences between configurations. Because the experiments involved only three seeds, we interpreted these observations descriptively without establishing statistical significance.

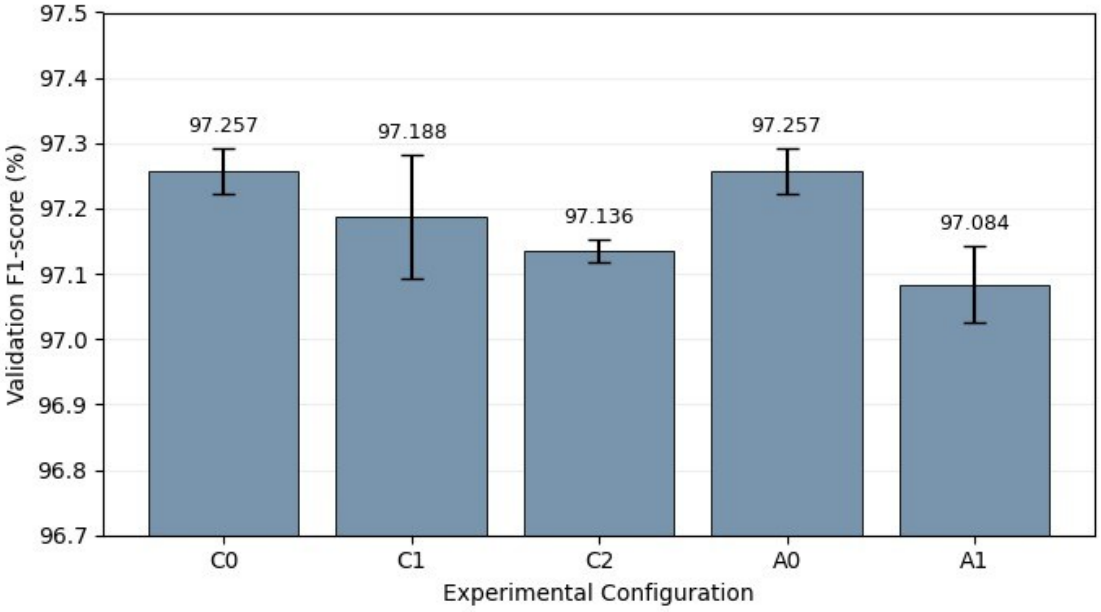

Figure 3. Mean validation F1-score with sample standard deviation across three random seeds (42, 123, and 2026)

The residual-gate ablation further demonstrated that the attention contribution depended on its configuration. A0 produced aggregate validation metrics identical to C0 across the three seeds, whereas A1 performed worse than the trainable-gate configuration C2. This indicates

that adding the attention output with a fixed coefficient of one provided no classification advantage under the evaluated conditions. Nevertheless, C2 also remained below the continued-training control. Identical aggregate metrics between A0 and C0 do not establish identical predictions or internal representations. Consequently, the initial holdout and controlled validation findings should be regarded as complementary but methodologically distinct evidence.

Discussion

The initial experiments demonstrated that spatial features provided a strong classification baseline, whereas FFT magnitude features were less effective when evaluated independently. However, combining both representations in S2 improved classification performance over the spatial-only configuration. Lokner Lađević et al. (2024) also demonstrated the relevance of CNN-based feature extraction for synthetic-image classification. Nevertheless, the improvement observed in S2 cannot be attributed exclusively to complementary frequency information because the dual-branch architecture also increased representational capacity. Thus, the results support the evaluated fusion configuration but do not establish the precise source of its improvement.

Differences among the initial configurations indicate that additional feature processing does not necessarily yield consistent improvements. While the fused representation performed better than the individual domains, introducing Direct Self-Attention in S3 reduced classification performance relative to S2. Park et al. (2024) investigated performance differences among AI-generated image detection methods and emphasized the relevance of representation and evaluation settings. The present results similarly suggest that classification effectiveness depends on how features are processed rather than simply on the complexity of the architecture. However, the observed reduction does not establish a general limitation of attention mechanisms.

Different image characteristics may provide useful evidence for authenticity assessment. Manisha et al. (2025) investigated camera fingerprint information as an alternative way to distinguish authentic and AI-synthesized images. In the present study, FFT magnitude representations supplemented spatial features, although the authors did not isolate their individual contributions within the fused representation. The evaluation setting also influences how broadly such findings can be interpreted. Islam et al. (2025) examined ensemble-based classification using multiple synthetic-image datasets. In comparison, the present experiments focused on architectural differences within two selected public dataset sources, limiting conclusions about generalization beyond those sources.

The controlled experiments provided the central finding concerning attention integration. Although S4 achieved the highest numerical holdout performance in the initial evaluation, continued training without attention produced the highest mean validation F1-score under matched training conditions. This difference suggests the initial advantage was not solely due to residual attention because the training histories were not equivalent. Li and Wang (2025) investigated detection using pretrained visual representations and proxy samples, emphasizing representation quality. The present findings likewise show that introducing additional attention processing did not consistently improve an already effective spatial-frequency model.

Residual Self-Attention was intended to preserve the fused representation while introducing information generated by attention processing. Powroznik et al. (2025) investigated residual self-attention within a medical-image classification architecture, although its implementation differed from the feature-level mechanism evaluated here. In the controlled experiments, residual attention did not outperform continued training without attention. The gate ablations also showed that the trainable configuration achieved higher mean validation performance than the fixed-one configuration, but neither exceeded the no-attention control.

These findings indicate that attention integration influenced classification outcomes without demonstrating a consistent benefit from residual preservation.

The confusion-matrix analysis further highlights the importance of considering classification errors alongside aggregate performance. The initial S4 model produced more false negatives than false positives, indicating that it incorrectly classified some generated images as authentic. Wang et al. (2025) investigated semantic inconsistencies through the RADAR framework, illustrating another form of evidence for identifying synthetic content. However, the present classifier does not provide equivalent semantic verification. Detection requirements may also differ across visual domains, as Yinghua et al. (2026) showed in their investigation of AI-generated artwork. Consequently, the reported performance should not be interpreted as definitive evidence of image authenticity across different content types.

Generalization remains an important consideration because the reserved holdout originated from the same public sources as the development data. Although this evaluation provided a separate set of images, it did not establish performance against independently sourced content or previously unseen generators. Chen et al. (2025) investigated contrastive learning for improving the generalization and robustness of computer-generated image detection, addressing an evaluation objective beyond the architectural comparison conducted here. The present findings therefore describe classification behavior under the selected experimental conditions. Independent cross-dataset and cross-generator testing would be required before drawing broader conclusions about detection robustness.

The controlled experiments also show performance variability. Residual attention produced the smallest F1-score standard deviation among the principal controlled configurations, but its mean performance remained below that of continued training without attention. Lower variation across the evaluated seeds therefore did not correspond to superior average classification performance. Yumlembam et al. (2025) investigated uncertainty measures and prediction rejection for AI-generated image detection, demonstrating another perspective on model reliability. The present study did not evaluate comparable uncertainty mechanisms, and its three repeated runs were insufficient to establish inferential significance. Accordingly, the observed differences are interpreted descriptively.

Further evaluation is needed to determine whether the observed relationships persist after image transformations. Fang et al. (2025) investigated how compression affects human judgments of AI-generated image authenticity, highlighting a relevant concern for digital-content assessment, although their study did not directly evaluate automated detector performance. The present experiments did not systematically examine compression and other post-processing conditions. Overall, the main contribution is that the initial numerical advantage of Residual Self-Attention did not reproduce under matched training conditions. This emphasizes the importance of continued-training controls and repeated evaluation when interpreting small architectural differences.

CONCLUSION

This study comparatively evaluated Direct and Residual Self-Attention in spatial-frequency ResNet-18 for AI-generated image detection. The initial experiments showed that spatial-frequency fusion improved classification performance over the spatial baseline and that residual attention achieved the highest numerical holdout results. However, the additional controlled experiments did not reproduce this advantage, with continued training without attention achieving the highest mean validation F1-score. Residual-gate ablations further showed that attention-derived contributions affected performance differently depending on configuration. The principal contribution is therefore a controlled empirical clarification of attention integration rather than evidence that residual attention is inherently superior. The findings emphasize the importance of matched training and repeated evaluation when

interpreting small architectural differences. Further research should examine independent generators, post-processing robustness, and inferential analysis to determine whether these patterns persist under broader conditions.

REFERENCES

Bird, J. J., & Lotfi, A. (2024). CIFAKE: Image Classification and Explainable Identification of AI-Generated Synthetic Images. *IEEE Access*, *12*, 15642–15650. https://doi.org/10.1109/ACCESS.2024.3356122

Chen, Y., Yin, W., Luo, A., Yang, J., & Wang, J. (2025). Improving the Generalization and Robustness of Computer-Generated Image Detection Based on Contrastive Learning. *International Journal of Intelligent Systems*, *2025*(1), Article 9939096. https://doi.org/10.1155/int/9939096

Fang, X., Chen, P., Wang, M., & Wang, S. (2025). Examining the role of compression in influencing AI-generated image authenticity. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-91545-4

Gunukula, A. R., Das Gupta, H., & Sheng, V. S. (2025). Detecting AI-Generated Images Using a Hybrid ResNet-SE Attention Model. *Applied Sciences (Switzerland)*, *15*(13). https://doi.org/10.3390/app15137421

Islam, M. T., Lee, I. H., Alzahrani, A. I., & Muhammad, K. (2025). MEXFIC: A meta ensemble eXplainable approach for AI-synthesized fake image classification. *Alexandria Engineering Journal*, *116*, 351–363. https://doi.org/10.1016/j.aej.2024.12.031

Li, J., & Wang, K. (2025). Building a Universal Detector of AI-Generated Images Without Training on Them. *International Journal of Intelligent Systems*, *2025*(1), Article 8530953. https://doi.org/10.1155/int/8530953

Li, L., Lu, T., Song, J., & Cheng, K. (2026). SIDWA: Synthetic Image Detection Based on Discrete Wavelet Transform Stem and Deformable Sliding Window Cross-Attention. *Electronics (Switzerland)*, *15*(4). https://doi.org/10.3390/electronics15040891

Lokner Lađević, A., Kramberger, T., Kramberger, R., & Vlahek, D. (2024). Detection of AI-Generated Synthetic Images with a Lightweight CNN. *AI (Switzerland)*, *5*(3), 1575–1593. https://doi.org/10.3390/ai5030076

Man, Q., & Cho, Y. I. (2026). Transformer Based on Multi-Domain Feature Fusion for AI-Generated Image Detection. *Electronics (Switzerland)*, *15*(3). https://doi.org/10.3390/electronics15030716

Man, Q., Gee, S. J., & Cho, Y. I. (2026). Multi-Domain Perception Transformer for Generalized Forgery Image Detection. *Applied Sciences (Switzerland)*, *16*(1). https://doi.org/10.3390/app16010533

Manisha, Li, C. T., & Kotegar, K. A. (2025). AI-Synthesized Image Detection: Source Camera Fingerprinting to Discern the Authenticity of Digital Images. *IEEE Access*, *13*, 29660–29672. https://doi.org/10.1109/ACCESS.2025.3537558

Namani, Y., Reghioua, I., Bendiab, G., Labiod, M. A., & Shiaeles, S. (2025). DeepGuard: Identification and Attribution of AI-Generated Synthetic Images. *Electronics (Switzerland)*, *14*(4). https://doi.org/10.3390/electronics14040665

Park, D., Na, H., & Choi, D. (2024). Performance Comparison and Visualization of AI-Generated-Image Detection Methods. *IEEE Access*, *12*, 62609–62627. https://doi.org/10.1109/ACCESS.2024.3394250

Powroznik, P., Skublewska-Paszkowska, M., Nowomiejska, K., Gajda-Deryło, B., Brinkmann, M., Concilio, M., Toro, M. D., & Rejdak, R. (2025). Residual self-attention vision transformer for detecting acquired vitelliform lesions and age-related macular drusen. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-02299-y

Reza, M. S., Elias, F., Mahmud, M. I., & Ahmed, N. (2026). Attention and residual mechanism-

based CNN architecture (ARC-Net) with enhanced fairness generalization for deepfake facial image detection. *PLOS ONE*, *21*(4), Article e0340099. https://doi.org/10.1371/journal.pone.0340099

Wang, H., Liu, X., Lu, Z., Yan, C., Jiang, X., Wang, R., & Gavves, E. (2025). RADAR: Reasoning AI-Generated Image Detection for Semantic Fakes. *Technologies*, *13*(7). https://doi.org/10.3390/technologies13070280

Yinghua, W., Linyan, L., Wenjuan, M., & Yunzhe, Z. (2026). AI-generated artwork detection using self-distilled transformers with global–local feature learning and Grad-CAM interpretability. *Scientific Reports*, *16*(1), 497. https://doi.org/10.1038/s41598-025-29229-2

Yumlembam, R., Issac, B., Aslam, N., Babu, E. K., Collyer, J., & Kennedy, F. (2025). Detection of AI generated images using combined uncertainty measures and particle swarm optimised rejection mechanism. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-28572-8

Zhang, J., Wang, Y., Tohidypour, H. R., & Nasiopoulos, P. (2025). An Efficient Frequency Domain Based Attribution and Detection Network. *IEEE Access*, *13*, 19909–19921. https://doi.org/10.1109/ACCESS.2025.3534829

Zhang, Y., Pang, Z., Huang, S., Wang, C., & Zhou, X. (2025). Unmasking AI-created visual content: a review of generated images and deepfake detection technologies. *Journal of King Saud University Computer and Information Sciences*, *37*(6), 148. https://doi.org/10.1007/s44443-025-00154-8