A Frequency-Domain Analysis Approach for Distinguishing GAN-Generated Images from Real Images

Key Concepts: GAN Detection, Frequency-Domain Analysis, Image Forensics

*Presented by **Nguyen Quang Huy***

Faculty of Information Technology, Ton Duc Thang University

Email: 523h0140@student.tdtu.edu.vn

QR Code

Presentation slide

GAN Implementation

Github Resources

Outline

1. Introduction
2. Related Work
3. Proposed Methodology
4. Experimental Results
5. Conclusion

Introduction

- Generative Adversarial Networks (GANs) can now produce hyper-realistic images that often bypass human distinction.
- This capability lowers the barrier for creating "deepfakes" and disinformation, posing a challenge to media authenticity.
- To develop a robust, automated methodology for image forensics that moves beyond visual inspection

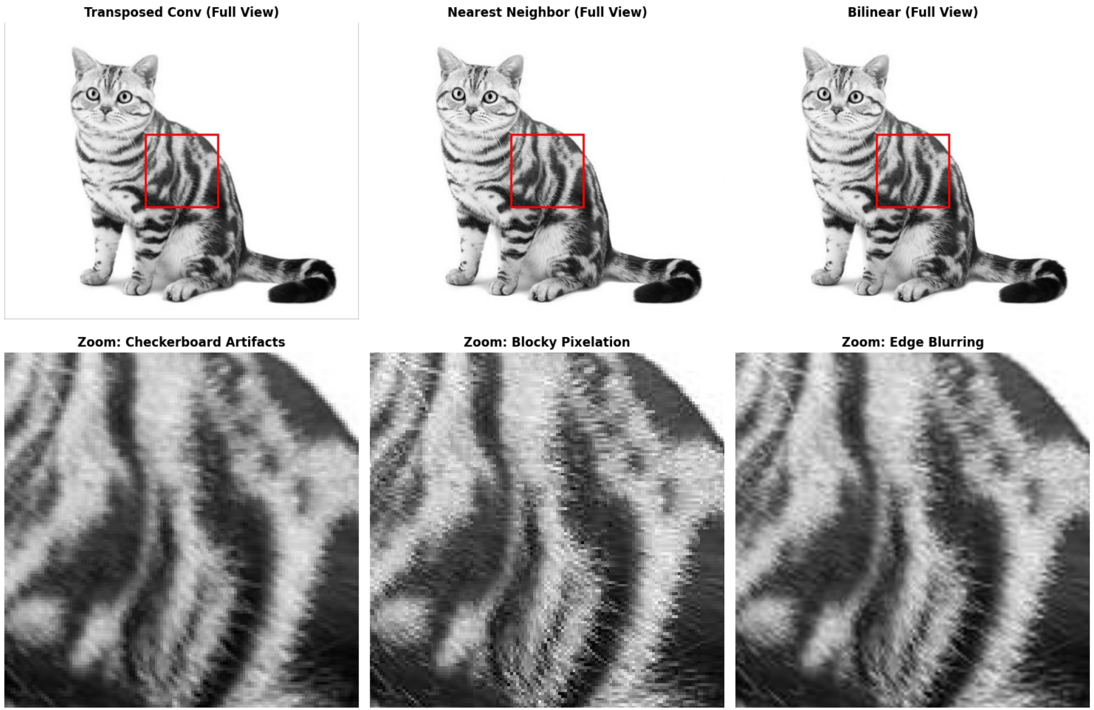
Original Image

Generated Image

Related Work

Limitations of Spatial-Domain Detection

- Invisible Artifacts
- Generalization Issues
- Sensitivity



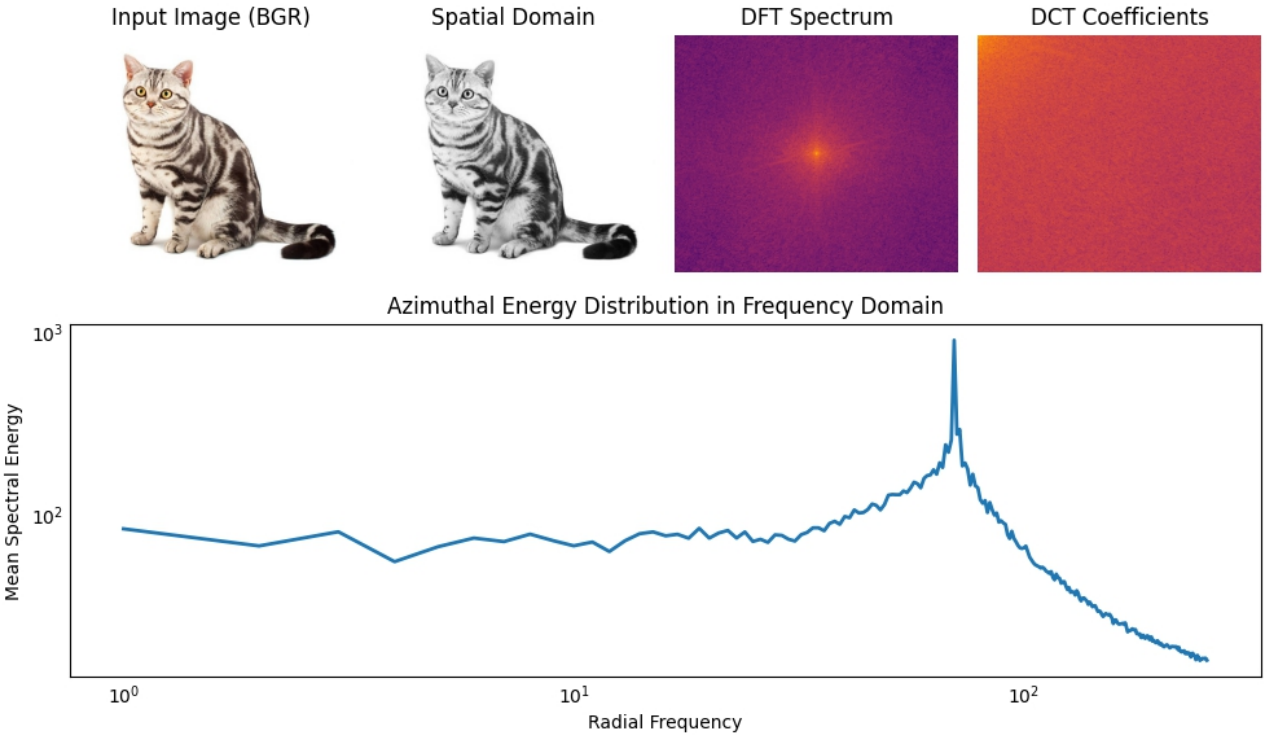
Related Work

Frequency-Domain Analysis

- Hidden Fingerprints: The mathematical process of upsampling low-dimensional codes into images leaves a distinct "fingerprint".
- Consistency: While image content (e.g., a face vs. a car) varies, the signal processing artifacts remain consistent for a given model architecture.
- DFT Advantage: The Discrete Fourier Transform (DFT) can reveal periodic grid patterns that are invisible in the pixel space

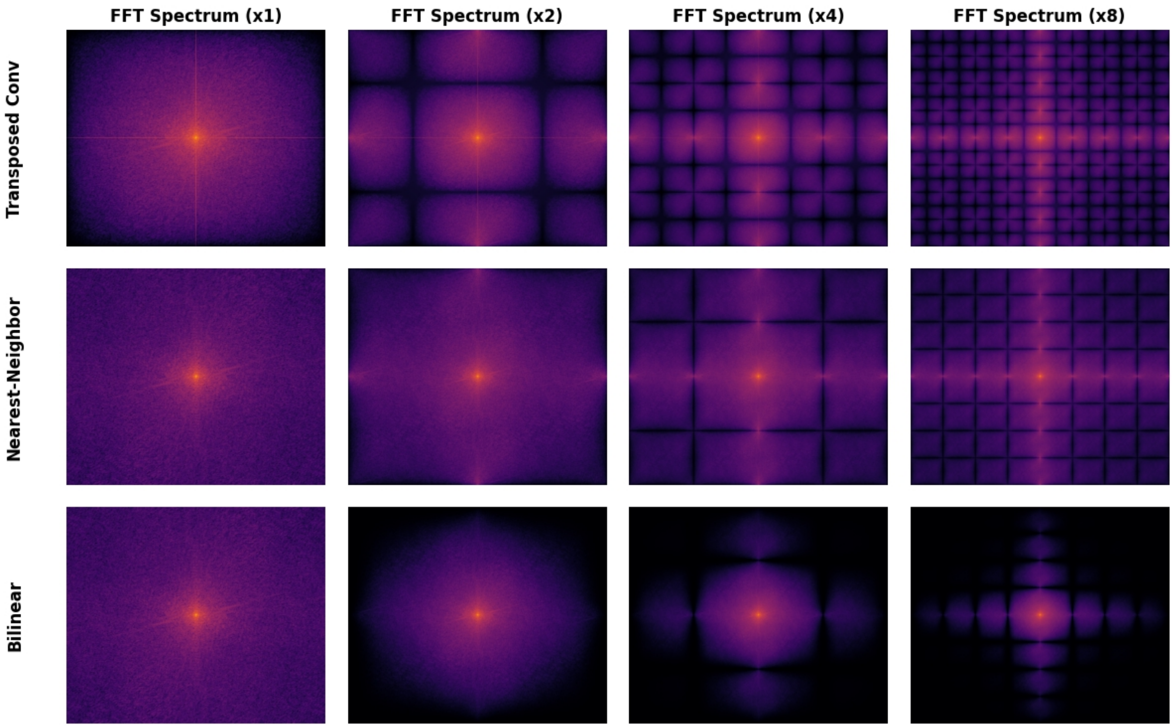
Related Work

Frequency-Domain Analysis



Related Work

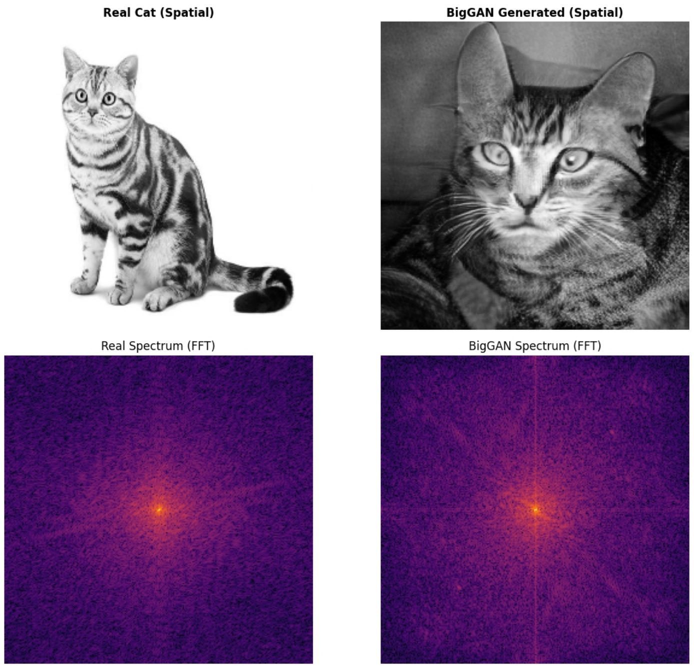
Upsampling Method



Proposed Method

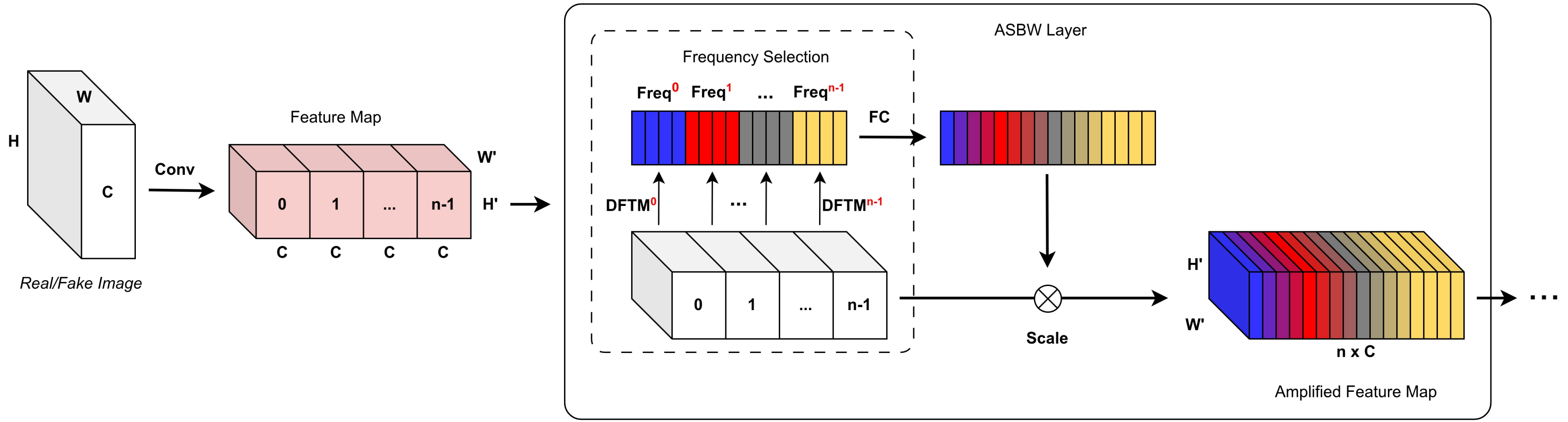
Goal

- To dynamically learn and emphasize the most discriminative frequency bands where forensic signals are concentrated.
- Design: A "plug-and-play" modular component that can be integrated into standard CNN backbones like ResNet.



Proposed Method

Architecture



Experimental Results

Dataset

"140k Real and Fake Faces" (70,000 real faces from FFHQ and 70,000 synthetic faces from StyleGAN).

Experimental Results

Backbone	Method	Accuracy (%)	F1-Score (%)
AlexNet	Baseline	50.00	33.33
	+ ASBW (Ours)	**50.00**	**33.33**
ResNet-18	Baseline	97.61	97.61
	+ ASBW (Ours)	**98.72**	**98.71**

Table 2. Performance comparison on the test set. The proposed ASBW layer demonstrates consistent improvements, particularly in the ResNet-18 architecture, validating its effectiveness in refining spectral features

Conclusion

- Key Takeaway: Frequency-domain analysis provides a robust and interpretable pathway for deepfake detection.
- Modularity: ASBW is effective as a modular enhancement for modern CNNs.
- Future Directions:
  - Testing against diffusion-based models (e.g., Midjourney).
  - Evaluating resilience against heavy image perturbations like aggressive compression