Ensemble-Based Fake Image Detection in Sentinel-2 RGB Data

Hilal Yildiz

M.Sc. Computational Science and Engineering
Technical University of Munich (TUM)
Garching, Germany

Luca Chiarabini, Daniela Espinoza-Molina, and Andrés Camero

German Aerospace Center (DLR)
Remote Sensing Technology Institute (IMF)
Wessling-Oberpfaffenhofen, Germany

ABSTRACT

The rapid advancement of image manipulation techniques has significantly increased the prevalence of highly realistic fake imagery, posing critical challenges for applications relying on the integrity of remote sensing data. This study investigates the detection of manipulated Sentinel-2 RGB imagery from a representation-driven perspective. Challenging the assumption that all forgeries share common patterns, we demonstrate that GAN-generated images and copy-move forgery (CMF) exhibit distinct characteristics requiring fundamentally different feature representations. Through systematic analysis using frequency-domain (FFT), wavelet-based, and subspace learning methods (Geo-DefakeHop), we show that individual model performance is strictly dependent on the alignment between representation and manipulation characteristics. Experimental results reveal that specialized models, including representation-specific ResNet architectures and Geo-DefakeHop variants, exhibit limited generalization due to representation mismatches. To address this, a representation-aware hierarchical ensemble framework is proposed. The system integrates these specialized models through a conditional decision mechanism that adaptively selects representations based on input characteristics. The proposed ensemble achieves a detection accuracy of 0.715, significantly outperforming all individual specialized configurations, including the best standalone Geo-DefakeHop (0.645) and various ResNet-based models using FFT and wavelet decompositions (0.513–0.535). By explicitly accounting for representation compatibility, the framework leverages complementary strengths without requiring prior knowledge of the forgery type. Overall, this research establishes that fake image detection is inherently representation-dependent and provides a principled, adaptive framework for robust and generalized remote sensing image forensics.

Index Terms— Remote Sensing, Generative Adversarial Networks, Copy-Move Forgery, Hierarchical Ensemble

1. REFERENCES

[1] H.-S. Chen, K. Zhang, S. Hu, S. You, and C.-C. J. Kuo, "GeoDefakeHop: High-performance geographic fake image detection," *APSIPA Transactions on Signal and Information Processing*, vol. 13, 2024.

[2] S.-Y. Wang, O. Wang, R. Zhang, A. Owens, and A. A. Efros, "CNN-generated images are surprisingly easy to spot... for now," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 8695–8704.

[3] T. Dzanic, K. Shah, and F. Witherden, "Fourier spectrum discrepancies in deep network generated images," *Advances in Neural Information Processing Systems*, vol. 33, pp. 3022–3032, 2020.

[4] X. Zhang, S. Karaman, and S.-F. Chang, "Detecting and simulating artifacts in GAN fake images," in *Proc. IEEE Int. Workshop on Information Forensics and Security (WIFS)*, 2019, pp. 1–6.

[5] B. Zhao, S. Zhang, C. Xu, Y. Sun, and C. Deng, "Deep fake geography? when geospatial data encounter artificial intelligence," *Cartography and Geographic Information Science*, vol. 48, no. 4, pp. 338–352, 2021.

[6] J. Sui, D. Ma, C.-C. J. Kuo, and M.-O. Pun, "FLDCF: A collaborative framework for forgery localization and detection in satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–13, 2024.

[7] U. A. Ciftci and I. Demir, "Deepfake satellite imagery detection with multi-attention and super resolution," in *Proc. IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS)*, 2023, pp. 4871–4874.

[8] L. Verdoliva, "Media forensics and deepfakes: An overview," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 5, pp. 910–932, 2020.