HIFoD: Rethinking Indicators of Compromise in Heterogeneous Image Forgery Detection, Classification and Localization

The digital era has witnessed the proliferation of image manipulation as a pervasive tool for deception and the dissemination of misleading information. While considerable progress has been made in the field of forgery detection, the emergence of heterogeneous forgery, which combines disparate image manipulation types on a single image, introduced a fresh and formidable challenge. This has led to identity theft, threat to national security, ethical breaches in journalism, and upsurge in online visual disinformation and misinformation. Deep learning techniques have been seen to perform strongly on homogeneous image forgeries, single occurrence or multiple, but limited in capacity to tackle heterogeneous forgeries [1]. This study investigates whether machine learning models, trained with a novel heterogeneous image forgery dataset (HIFoD) can deliver an optimum forgery detection and classification while still being computationally efficient.

We curated a dataset of 7,237 images comprising original and annotated-forged images. Each forged image in HIFoD contains copymove, image splicing and removal forgeries all together within an image. The images were resized to 612 x 612 x 3 and augmented with rotation, scaling, horizontal flip, and vertical flip. YOLOv4, an object detection model was employed to perform image forgery detection task. YOLOv4 [2], a CNN model, consists of four parts namely input, backbone, neck, and head, respectively. The core of any YOLO model is the backbone network, which is a CNN. This backbone is responsible for features extraction. The convolutional layers in the backbone network learn to identify and extract features from the image, in this case, features of forgeries.

In this study, three deep feature extraction methods ResNet101, DenseNet, and CSPDarknet53 were benchmarked against a modified YOLOv4 called GSDarknet model, a new model based on gravitational search algorithm (GSA). The modification was an enhancement to YOLOv4's backbone. GSA is an optimization algorithm that uses the principles of Newton's law of gravity and laws of motion to find the best solution in a search space, where better solutions are like heavier masses that attract others [3]. The performance was evaluated using Accuracy, F1-Score, mAP, IoU, while FPS and detection time were used to track the speed of detection.

GSDarknet achieved the highest balance of detection and classification performance with Accuracy of 98.1%, F1-score of 98.6%, and mAP of 86.3%. The IoU metric of 0.95 indicates a high confidence score and signifying a more precise localization of forged image areas, while 79 FPS shows that high number of frames are processed per second, requiring 20msec detection time.

This study demonstrates that the newly developed dataset strongly represents sophisticated forgeries with disparate forgery types and it is quite impressive for the model training. The performance of the GSDarknet model indicates that optimizing the feature extraction process can achieve an outstanding image forgery detection, classification and localization capabilities. Using GSA, the model dynamically balances exploration-exploitation in order to handle the selection of both highly valuable and informative features for the classifier as well as fine-tuning the output of the localization model, thereby allowing the algorithm to effectively find optimal solutions without getting stuck in local optima. Future work will investigate newer versions of YOLO models and also expand the size of the dataset.

References

[1] Pandey, Ashutosh & Parwekar, Pritee & Chakraverti, Ashish. (2022). A Review On Deep Learning Techniques for Image Forgery Detection. 890-895. 10.1109/ICCCIS56430.2022.10037746.

[2] Bochkovskiy, Alexey & Wang, Chien-Yao & Liao, Hong-yuan. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. 10.48550/arXiv.2004.10934.

[3] Rashedi, E., Nezamabadi-pour, H., & Saryazdi, S. (2009). GSA: A Gravitational Search Algorithm. *Information Sciences*, 179(13), 2232–2248.