M2 Internship Subject:

Detecting Manipulated 3D Gaussian Splatting: Evaluation and Improvement of Generalization Performance

Keywords: Multimedia forensics, 3D Gaussian splatting, 3D manipulation detection, cross-dataset generalization, data efficiency.

Context, State of the Art, and Objective

Detecting manipulated multimedia content has been an active area of research since the proliferation of easy-to-use manipulation tools, primarily those based on AI (artificial intelligence). Numerous methods have been proposed to detect image or video manipulations, including those developed by our research group at GIPSA-lab, which focus mainly on digital image forensics [1,2,3]. Recently, a new type of multimedia content known as 3D Gaussian Splatting (3DGS) [4] has gained significant popularity. 3DGS is a revolutionary 3D reconstruction and rendering technique that represents three-dimensional scenes using a large number of 3D Gaussian ellipsoids rather than traditional polygonal meshes. 3DGS models reconstructed from the real world require protection, and it must be possible to detect any illegal manipulation of authentic 3DGS models. This forensic challenge associated with 3DGS has recently attracted researcher interest, giving rise to some pioneering studies [5,6] (please refer to these two references for examples of authentic and manipulated 3DGS models). In these two key references for this internship, two datasets comprising authentic and manipulated 3DGS models were created, and two detection methods, each trained on one of these datasets, were proposed.

The objective of this internship is to carry out, for our research group, a first experience in forensic analysis applied to 3DGS, with an emphasis on very concrete aspects, i.e., the *generalization capacity* and the *data efficiency* of 3DGS manipulation detectors. The issue of generalization arises when a detector is trained on known source data and then tested on unknown target data. This is a highly realistic scenario, as new 3DGS manipulation methods can emerge daily, making it impossible to account for every possible manipulation technique during the detector's training phase. Furthermore, data efficiency is of paramount importance; it enables training with fewer data and reduced computational resources, while potentially leading to better generalization with reduced overfitting.

Methodologies and Tasks

The internship student is expected to accomplish the following tasks.

1. We begin with a brief literature review on the relevant topics, covering approximately 5 to 10 scientific articles, including the two key references mentioned above [5,6].
2. The first step involves evaluating the detection methods proposed in [5,6] in terms of generalization capability. Specifically, the two existing methods will be tested on the dataset that was not used during the detector's training phase; for instance, the detector proposed in [5] will be tested on the dataset constructed in [6], and vice versa. This cross-dataset evaluation will potentially reveal the limitations regarding the generalization capabilities of the two existing methods. Through this evaluation, we will also gain a better understanding of the factors contributing to any potential drop in generalization performance, as well as the specific behaviors of the two detection methods. Generalization performance is expected to be suboptimal, a common finding in the field of forensics for other types of multimedia content, such as image forensics [1,2,3]. Even if generalization proved rather satisfactory, this would not diminish the value of our work, as we will focus on data-efficient detection methods, whereas existing methods rely on supervised learning that requires large quantities of training samples.

3. The second step involves improving the generalization capability of methods for detecting manipulated 3DGS models, while prioritizing data and computational efficiency. To this end, we aim to extend the generalizable and data-efficient methods we previously proposed for synthetic image detection which rely on proxy samples [2] or anomaly detection [3]. The core idea behind these two research lines is to replace data-intensive supervised learning with either self-supervised learning using easily generated proxy samples, or unsupervised learning that leverages only authentic samples combined with mathematical frameworks of anomaly detection.
4. The internship concludes with the writing of the report and the preparation of the oral defense.

Implementation will be based on PyTorch. Computations may be performed either on a dedicated workstation within our research group or on the GRICAD platform (https://gricad.univ-grenoble-alpes.fr/).

Expected Results and Impact

Where appropriate, we will prepare and submit a paper to an international conference, or even an article to an international journal, based on the research conducted during this internship. We will adopt an open source approach for this project to enhance its quality and impact. Beyond the expected scientific outcomes, we hope that this internship project will also have a societal impact by making available to the public useful 3DGS forensic analysis tools to detect manipulated 3D models that people might encounter in their professional or personal lives. We believe that our internship proposal aligns very well with the objectives of MIAI and EFELIA MIAI by focusing on the development of data-efficient AI-based detection methods, addressing human and societal well-being in the face of AI-induced multimedia manipulation, and training an AI Master student through a relevant, cutting-edge research topic.

Work Environment and Application

The proposed internship will take place within the PICSA team (*Physique Image Cerveau Signal Apprentissage*), part of the Data Science Department at GIPSA-lab (*Grenoble Images Parole Signal Automatique*, http://www.gipsa-lab.grenoble-inp.fr/). The internship student will work with Dr. Kai Wang (CNRS researcher and member of the PICSA team) and one of his Ph.D. students. A monthly stipend of approximately 680€ will be provided. The working language may be French or English. We are seeking outstanding candidates with strong skills in applied mathematics and multimedia analysis, high motivation regarding the proposed research topic, and proficiency in Python programming. Priority will be given to candidates with a strong interest in pursuing a Ph.D. To apply, please send a CV, a brief cover letter, and M1 and M2 transcripts (not mandatory but highly recommended) to the following email address.

kai.wang@gipsa-lab.grenoble-inp.fr

The expected start date for the internship is February 1, 2027, but it may be postponed depending on the candidate's training program schedule.

References:

[1] W. Quan, P. Deng, K. Wang, and D.-M. Yan, CGFormer: ViT-based network for identifying computer-generated images with token labeling, *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 235-250, 2024.

[2] J. Li and K. Wang, Building a universal detector of AI-generated images without training on them, *International Journal of Intelligent Systems*, article no. 8530953, pp. 1-23, 2025.

[3] Z. Yuan, K. Wang, W. Quan, D.-M. Yan, and T. Wu, CLIP-Flow: A universal discriminator for AI-generated images inspired by anomaly detection, *Proceedings of the ACM Deepfake Forensics Workshop*, Dublin, Ireland. pp. 1-9, 2025.

[4] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, 3D Gaussian splatting for real-time radiance field rendering, *ACM Transactions on Graphics*, vol. 42, no. 4, article no. 139, pp. 1-14, 2023.

[5] D. Di Nucci, R. Catalini, G. Borghi, and R. Vezzani, Fake3DGS: A benchmark for 3D manipulation detection in neural rendering, *Proceedings of the International Conference on Pattern Recognition*, pp. 1-15, 2026.

[6] H. Han, Z. Luo, J. Qi, A. Rocha, and R. Wan, GS-Checker: Tampering localization for 3D Gaussian splatting, *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 40, no. 6, pp. 4556-4564, 2026.