**Original Manuscript ID:** Access-2024-41876

**Original Article Title: “**Unsupervised Geometric-guided Industrial Anomaly Detection”

**To:** IEEE Access Editor

**Re:** Response to reviewers

Dear Editor,

Thank you for allowing a resubmission of our manuscript, with an opportunity to address the reviewers’ comments.

We are uploading (a) our point-by-point response to the comments (below) (response to reviewers, under “Author’s Response Files*”*), (b) an updated manuscript with yellow highlighting indicating changes (as “Highlighted PDF*”*), and (c) a clean updated manuscript without highlights (“Main Manuscript”*).*

Best regards,

Dinh-Cuong Hoang, et al.

**Reviewer#1, Concern # 1:** If the authors could include more qualitative figures across various types and provide more detail on the datasets used, it would enhance the paper further. Additionally, citing peer-reviewed publications rather than preprints from arXiv would improve the reliability and credibility of the references.

**Author response:** Thank you for this suggestion. We acknowledge the importance of including qualitative figures across various types of anomalies to better illustrate the model’s effectiveness. We will provide additional qualitative results showcasing the model's performance, including both anomaly maps and ground-truth segmentations across diverse object categories. Detailed descriptions of the datasets, such as the specific defect types, their distribution, and the sensor characteristics used for data acquisition, will also be incorporated into the revised manuscript.

**Author action:**

* Added qualitative figures illustrating anomalies for diverse object categories from the MVTec 3D-AD dataset.
* Expanded the dataset description section to include details about defect categories, data acquisition methods, and pre-processing steps.

**Reviewer#1, Concern # 2:** The authors may consider citing additional state-of-the-art works that use GANs in the literature review, particularly from IEEE Access, relevant to this research area. Suggested references include:

T. Ganokratanaa, S. Aramvith, and N. Sebe, "Unsupervised Anomaly Detection and Localization Based on Deep Spatiotemporal Translation Network," IEEE Access, vol. 8, pp. 50312-50329, 2020. doi: 10.1109/ACCESS.2020.2979869.

Thittaporn Ganokratanaa, Supavadee Aramvith, Nicu Sebe, "Video anomaly detection using deep residual-spatiotemporal translation network," Pattern Recognition Letters, vol. 155, pp. 143-150, 2022. ISSN 0167-8655. https://doi.org/10.1016/j.patrec.2021.11.001.

**Author response:** We have reviewed the suggested papers and incorporated them into the literature review to provide a more comprehensive overview of related GAN-based anomaly detection techniques.

**Reviewer#2, Concern # 1:** The title should be refined to provide precise information about the focus of the research study by using "image-based anomaly detection" or "Industrial quality inspection" instead of "Industrial Anomaly Detection", which is actually a very broad topic dealing with anomaly detection using various sensors embedded in machinery and equipment to capture detailed information about the process, equipment's sensors and/or the quality of products.

**Author response:** We agree that the title could be refined to better reflect the specific focus on image-based industrial anomaly detection. We will revise the title to emphasize the use of 2D and 3D data in quality inspection rather than the broader field of anomaly detection.

**Author action:**  Revised the title to: "Unsupervised Visual-to-Geometric Feature Reconstruction for Image-Based Industrial Quality Inspection."

**Reviewer#2, Concern # 2:** The technical aspects of the proposed method are not strong enough and lack pertinent experimental tests and an in-depth analysis, in particular the Transformer-based Visual-to-Geometric feature reconstruction.

**Author response:** We acknowledge the need for additional experimental validation and an in-depth analysis of the Transformer-based Visual-to-Geometric feature reconstruction. We will enhance the technical explanation by providing a detailed mathematical formulation of the reconstruction process and further experiments, such as ablation studies, to evaluate its contribution to anomaly detection performance.

**Author action:**

* Expanded the methodology section with a more detailed explanation of the Visual-to-Geometric reconstruction network, including equations and implementation specifics.
* Added an ablation study evaluating the impact of components such as non-local attention and GCN on the reconstruction quality.

**Reviewer#2, Concern # 3:** The claim of a substantial improvement of the proposed method requires more evidence to be credible. To get reproducible results, authors should expand simulation tests and provide a better description of the implementation of the Transformer-based Visual-to-Geometric feature reconstruction, which is the core of this research study.

**Author response:** We agree that reproducibility is critical and will provide more detailed implementation specifics for the Transformer-based reconstruction module, including training parameters, hardware specifications, and hyperparameter tuning. Additionally, we will conduct more simulation tests to demonstrate the method’s robustness across various scenarios.

**Author action:**

* Detailed the implementation of the Transformer-based module, including configurations of Vision Transformers and training protocols, in a new subsection under the methodology.
* Included additional simulation results using variations in dataset size and defect characteristics.

**Reviewer#2, Concern # 4:** The Why the global anomaly score for each sample is based on the maximum value in the anomaly map? Apparently this decision is sensitive to noise, increasing false alarm rates. How to determine the best threshold to avoid a lot of false positives?

**Author response:** We appreciate this important observation. To address the issue of sensitivity to noise, we conducted additional experiments using alternative anomaly score formulations, such as mean or weighted aggregate values from the anomaly map. Furthermore, we will include a threshold optimization analysis to determine the trade-off between false positives and detection sensitivity.

**Author action:**

* Added a new subsection analyzing the use of alternative aggregation methods (e.g., mean, weighted mean) for computing the global anomaly score.
* Included a quantitative evaluation of threshold selection strategies, such as ROC curve analysis, to identify the optimal threshold for minimizing false positives.