# Intelligent Anomaly Detection for Lane Rendering Using Transformer with Self-Supervised Pre-Training and Customized Fine-Tuning

Dong, Yongqi ; Lu, Xingmin; Li, Ruohan; Song, Wei; van Arem, Bart ; Farah, Haneen

**Publication date**
2024

**Document Version**
Final published version

**Citation (APA)**
Dong, Y., Lu, X., Li, R., Song, W., van Arem, B., & Farah, H. (2024). *Intelligent Anomaly Detection for Lane Rendering Using Transformer with Self-Supervised Pre-Training and Customized Fine-Tuning*. Poster session presented at Transportation Research Board 103rd Annual Meeting 2024, Washington DC, United States.

# Intelligent Anomaly Detection for Lane Rendering Using Transformer with Self-Supervised Pre-Training and Customized Fine-Tuning

Authors: Yongqi Dong [#,*] | Xingmin Lu [#] | Ruohan Li | Wei Song [*] | Bart van Arem | Haneen Farah
y.dong-4@tudelft.nl

Event Number: 3215, Event Title: Transforming Transportation Planning and Traffic Forecasting with Artificial Intelligence, Event Location: Convention Center, Hall A
Event Time: Tuesday, Jan 09, 2023 3:45PM - 5:30PM, Presentation Number: TRBAM-24-06382

## Background

- The burgeoning navigation services using digital maps provide great convenience to drivers.
- There are anomalies (errors and/or defects), e.g., irregular shapes, and missing edges or corners, in lane-level rendered map images.
- These anomalies will be equivocal for human drivers' understanding and decision-making during their driving routing which might result in critical unsafe situations.

## Aim

- To accurately and effectively detect lane rendering image anomalies;
- To transform the lane rendering anomaly detection problem into a muti-class classification problem and leveraging state-of-the-art AI models;
- To delivery excellent detection performance in regarding various metrics.

## The framework of the proposed pipeline

- **Image pre-processing, which normalizes the inconsistent images into uniform size and format;**
- **Self-supervised pre-training, which is tackled by the masked image modeling (MiM) method.**
- **Customized fine-tuning;**
- **Post-processing;**
- **Tested models:**
  - ViT
  - BEiT
  - Swin-Transformer (Swin-Trans)
  - Swin-Transformer-UniformMasking (Swin-Trans-UM)

## Evaluation Metrics

- **Accuracy**
- **F1-Meassure**
- **Precision**
- **Recall**
- **True Positive Rate**
- **False Negative Rate**

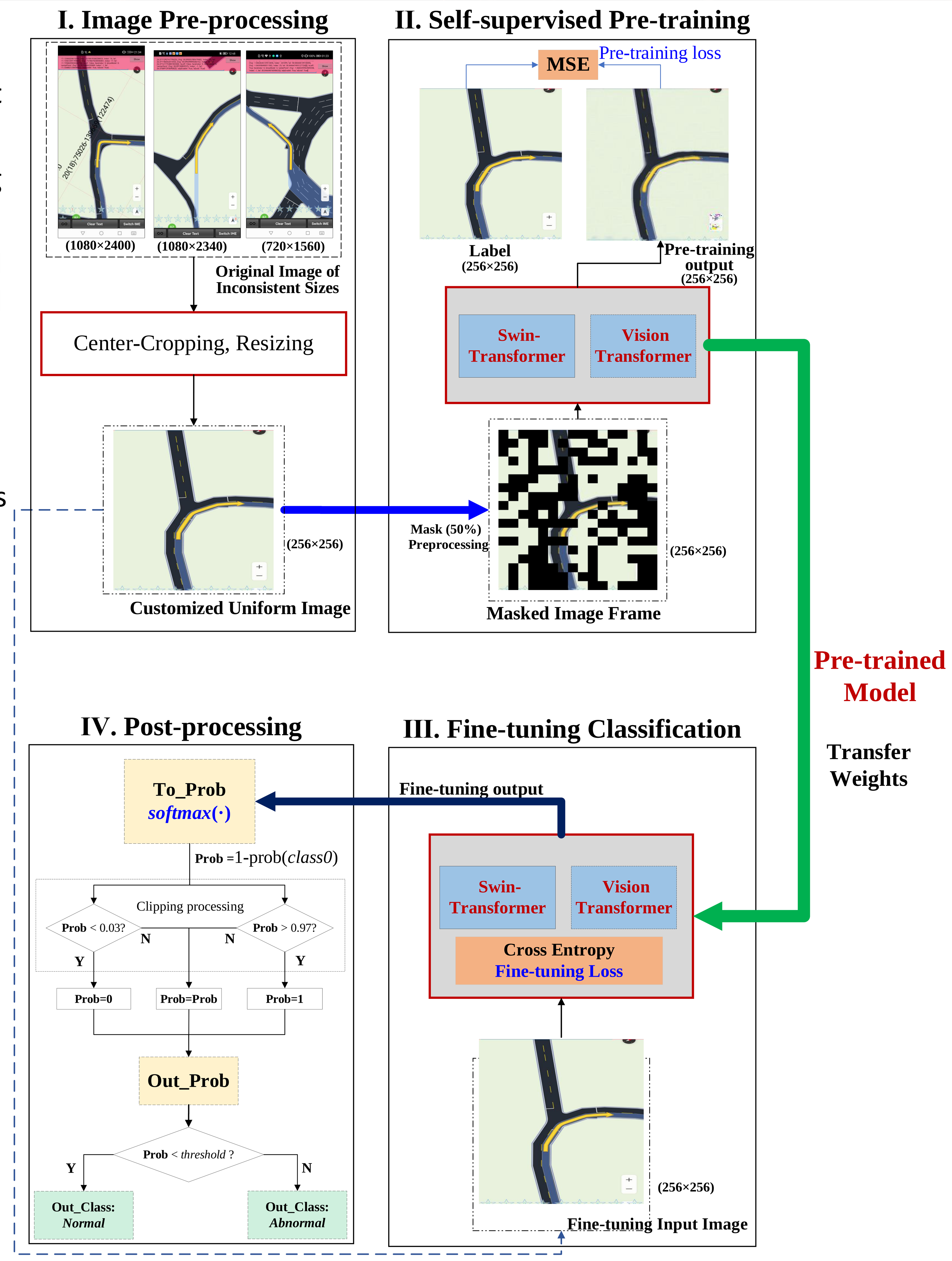


**Figure 1. The architecture of the proposed four-phase pipeline.**

## Results

Converted the problem of lane rendering image anomaly detection into a classification problem;
Various SOTA computer vision techniques and models were adopted and compared.

**Table 1 The model performance regarding different metrics.**

| Model | Acc | AUC | Precision | Recall | F1-measure | Param | Epoch time | Fine-tuning Epoch |
|---|---|---|---|---|---|---|---|---|
| ViT | 0.9489 | 0.9080 | **0.9393** | 0.6178 | 0.7454 | 632.20 | 4210 | 40 |
| BEiT | 0.9413 | 0.9481 | 0.7913 | 0.6996 | 0.7427 | 311.53 | 159 | 15 |
| Swin-Trans | 0.9401 | 0.9498 | 0.8518 | 0.6121 | 0.7123 | 86.90 | 120 | 280 |
| Swin-Trans-UM | **0.9477** | **0.9743** | 0.7743 | **0.8022** | **0.7805** | 194.95 | 223 | 41 |

**Table 2 The performance of the Swin-Trans-UM_2 and Swin-Trans-UM_9.**

| Model | Accuracy | AUC | Precision | Recall | F1-measure |
|---|---|---|---|---|---|
| Swin-Trans-UM_2 | **0.9482** | **0.9756** | **0.7813** | **0.7947** | **0.7879** |
| Swin-Trans-UM_9 | 0.9392 | 0.9731 | 0.6990 | 0.8745 | 0.7770 |
| Swin-Trans-UM_8 | 0.9477 | 0.9743 | 0.7743 | 0.8022 | 0.7805 |

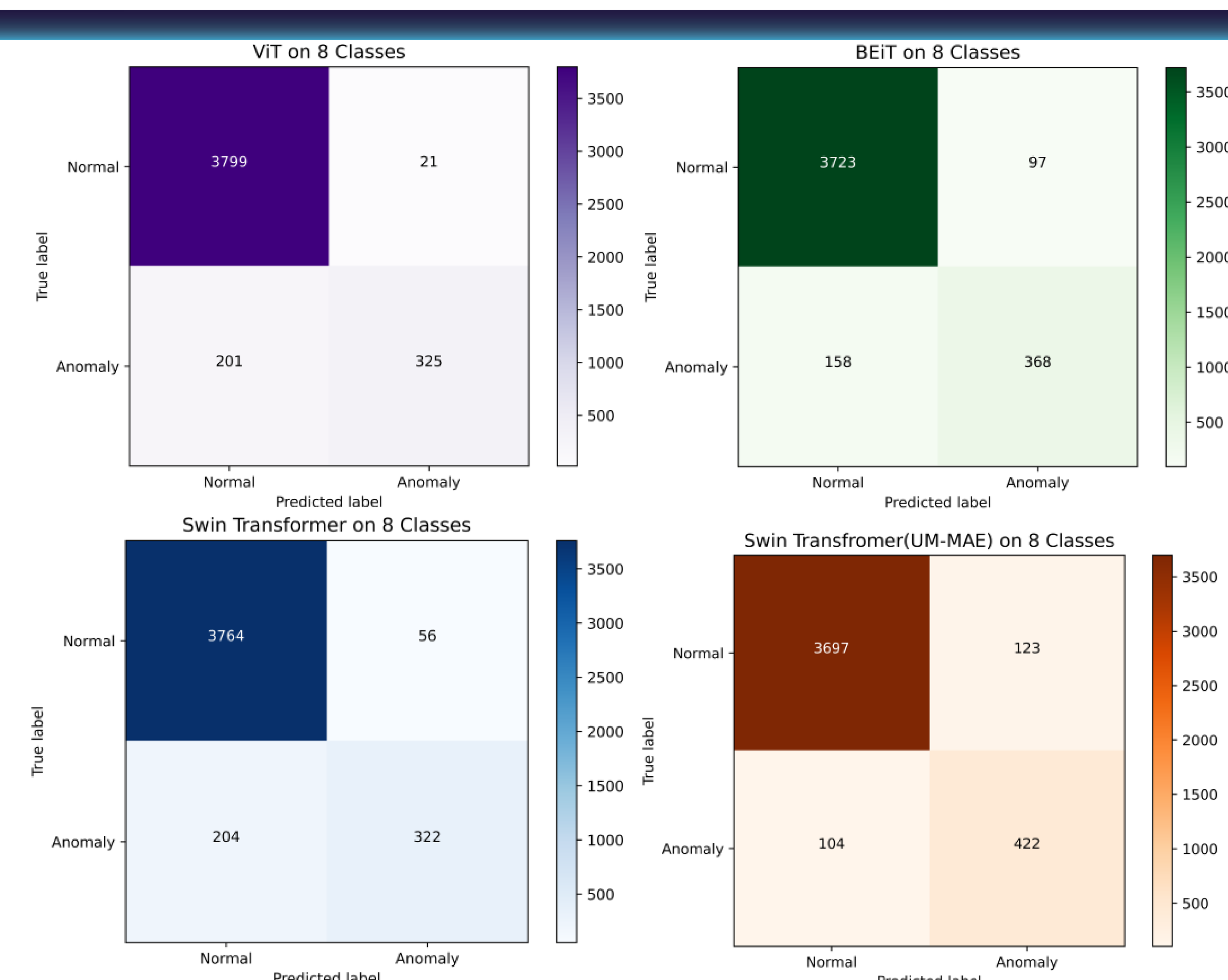


**Figure 2. The confusion matrix of test models.**

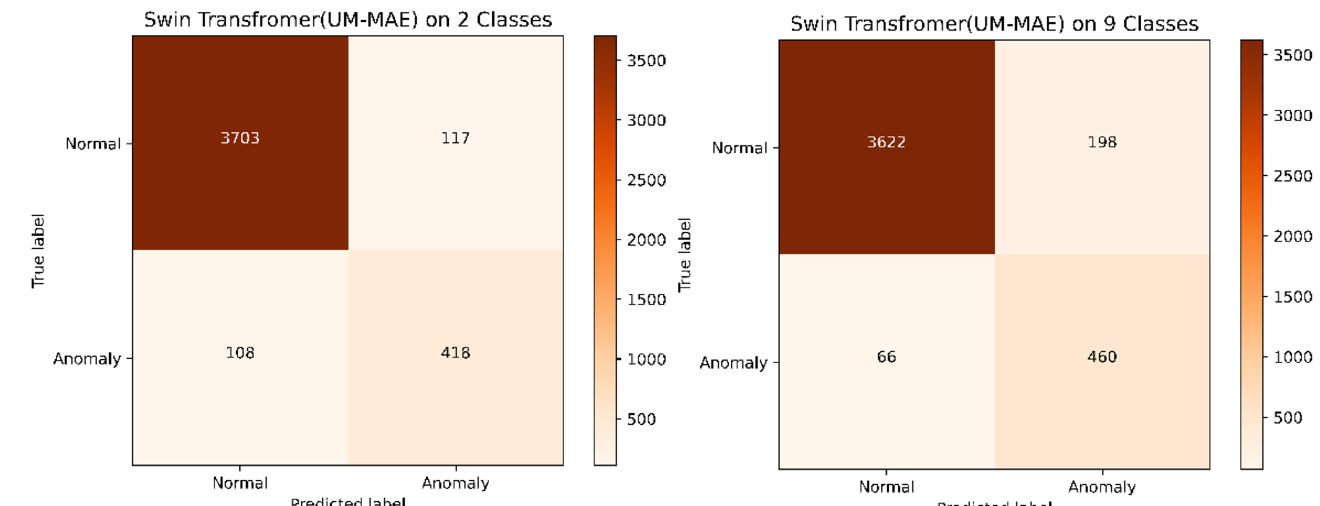


**Figure 3. The confusion matrix of Swin-Trans-UM when treated as a 2-class classification and a 9-class multi-label classification.**

## Conclusions

- **The proposed four-phase pipeline can tackle the lane rendering image anomaly detection task with super performances at high accuracy.**
- **The self-supervised pre-training with MiM can greatly improve the model accuracy.**
- **The proposed method can improve the efficiency of lane rendering image data anomaly detection reducing labor costs while keeping high accuracy.**

TU Delft | Villanova University | NWO Netherlands Organisation for Scientific Research | SAMEN | Traffic and Transportation Safety Lab