BRNO FACULTY
UNIVERSITY OF INFORMATION
OF TECHNOLOGY TECHNOLOGY

Excel@FIT 2025
http://excel.fit.vutbr.cz

Fast Adaptation of Segmentation Models for Tire Sidewall Symbols

Michal Balogh*

Abstract
The tire manufacturing industry needs efficient quality control systems that can adapt quickly to new tire sidewall symbols. Current industrial segmentation solutions require 72 hours of model retraining when new symbols are introduced. We employ state-of-the-art OneFormer architecture with transformer backbone and apply efficient fine-tuning techniques including mixed precision training, selective parameter freezing, and Low Rank Adaptation (LoRA). Our approach reduces the adaptation time from 72 hours to less than 1 hour without sacrificing segmentation quality (F1 Score: 0.94, Panoptic Quality: 0.91). The solution enables rapid deployment of quality control systems for new tire types and opens possibilities for customer-specific segmentation model adaptation.

*baloghmichal03@gmail.com, *Faculty of Information Technology, Brno University of Technology*

1. Introduction

Tire sidewalls contain engraved symbols that provide crucial information about tire capabilities, usage conditions, and manufacturing details. These symbols must meet strict quality standards. Industrial inspection systems use computer vision to segment and verify these symbols, but adapting to new symbols is time-consuming with current methods.

The current industrial solution at Micro-Epsilon Inspection requires approximately 72 hours to retrain their Mask R-CNN model when new symbols are introduced. This creates a significant bottleneck in production when new tire types arrive. An ideal solution would maintain or improve segmentation quality while drastically reducing adaptation time.

Current approach involve complete retraining of segmentation models, which is computationally expensive and time-consuming. Some research explores few-shot learning and zero-shot segmentation methods, but these often sacrifice precision for adaptability.

We implement a solution based on the state-of-the-art OneFormer [1] architecture with efficient fine-tuning techniques. By combining mixed precision training [2], selective parameter freezing, and Low Rank Adaptation [3], we enable rapid adaptation to new symbols while preserving knowledge about previously learned ones. Our approach requires only a small sample of new symbols and minimal computational resources.

We demonstrate that mixed data training with a 1:6 ratio of new to original data preserves performance on existing symbols (0% quality loss on all metrics) while achieving excellent results on new symbols (F1 Score of 0.97). Our approach works on consumer-grade hardware (8GB GPU) and can be applied to various segmentation architectures.

2. Industrial Problem Context

The sections titled "Experiments" and "Methods" on the poster show an artificially created tire sidewall featuring embossed symbols and text. Real examples of these symbols are displayed in the poster's title. These markings are engraved into the same black material as the tire itself, making them difficult to distinguish with standard RGB cameras. To address this, depth scanning technology is used to produce grayscale images where pixel intensity corresponds to surface depth, making the symbols clearly visible.

The industrial solution processes these depth images to detect, segment, and verify the symbols. However, when new tires are introduced, they often include new symbols or variations that were not part of the original training data. The current system, based on Mask

R-CNN, requires a full retraining process, which takes approximately 72 hours on industrial-grade hardware.

The primary challenge is adapting quickly to new symbols while maintaining recognition of previously learned ones. This issue, known as catastrophic forgetting, occurs when a model loses existing knowledge while learning from new data.

3. Technical Approach

Our solution is built upon the OneFormer architecture, which unifies different segmentation types (semantic, instance, and panoptic) in a single framework. As shown in Table 1, instance segmentation provides the most accurate masks for tire sidewall symbols, showing a significant advantage over semantic segmentation (+23.01% in Panoptic Quality).

The core of our approach involves three key techniques:

3.1 Selective Parameter Freezing

By freezing the Swin [4] Transformer backbone (which comprises 83% of model parameters), we preserve the general visual feature extraction capabilities while allowing adaptation in the segmentation and classification heads.

This selective freezing approach significantly reduces the severity of catastrophic forgetting compared to full model fine-tuning. With only 10% of parameters being trainable, adaptation becomes much faster while maintaining high-quality segmentation masks.

3.2 Mixed Data Training Strategy

Our experiments revealed that the optimal strategy for preserving performance on original symbols while learning new ones involves training on a mixture of both data types. Figure 1 shows the impact of different ratios (1:1, 1:2, 1:3, and 1:6 new:original) on model performance on the old data.

The 1:6 ratio achieves the best balance - maintaining performance on original data (0% degradation in metrics) while achieving excellent results on new symbols (Dice Score of 0.974). This approach addresses catastrophic forgetting without requiring complex architectural modifications.

3.3 Low Rank Adaptation (LoRA)

LoRA adds small trainable matrices to the attention mechanism of the transformer. These matrices have significantly fewer parameters than the original model components but can effectively adapt the model's behavior.

While LoRA showed comparable performance to selective freezing with similar training times, it offered additional benefits for smaller models. When using smaller Swin backbone, LoRA achieved good adaptation with only 19.65% of parameters being trainable compared to 34.1% with selective freezing.

4. Experimental Results

Fine-tuned model maintains performance on original symbols while achieving excellent results on new ones.

The mixed precision training technique further accelerated the training process by 27%, allowing adaptation in just 37 minutes for a new dataset of 38 images mixed with 228 original images. This rapid adaptation time represents a dramatic improvement from the original 72 hours.

5. Conclusions

Our solution drastically reduces the time required to adapt segmentation models to new tire sidewall symbols - from 72 hours to less than one hour - while maintaining high-quality segmentation results. This improvement enables tire manufacturers to deploy quality control systems for new tire types much more rapidly.

The implemented methods are architecture-agnostic and could be applied to other models.

The approach enables more personalized quality control systems tuned to specific customer requirements, potentially opening new business opportunities for industrial inspection system providers.

References

[1] Jitesh Jain, Jiachen Li, MangTik Chiu, Ali Hassani, Nikita Orlov, and Humphrey Shi. Oneformer: One transformer to rule universal image segmentation, 2022.

[2] NVIDIA Corporation. *Train With Mixed Precision*, 2023.

[3] Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models, 2021.

[4] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.