**Report**

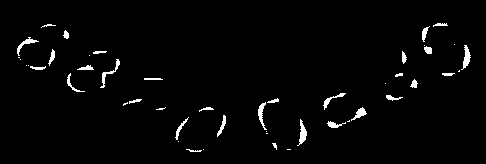
In this experiment, two primary approaches were explored: a classification-based pipeline and a segmentation-based pipeline, each designed to leverage state-of-the-art architectures while addressing the constraints of a relatively limited dataset.

**Classification Approach:**  
The classification methodology employed a modular model design, using CLIP-based backbones and allowing for variable numbers of trainable parameters. Two configurations were tested—one with the ViT-B/32 architecture integrated through CLIP, and the other using a ConvNeXt-Base-W backbone also provided by CLIP. This modular design enabled dynamic adjustment of network complexity relative to the dataset size, effectively preventing overfitting and ensuring efficient parameter utilization. Preliminary results indicate that both configurations achieved competitive precision and accuracy, demonstrating the robustness of the classification approach. The flexible design also made it easier to fine-tune the model to the dataset’s scale and complexity. The resulting performance metrics suggest that the classification strategy can be readily adapted and improved upon with minimal computational overhead.

**Segmentation Approach:**  
In contrast, the segmentation-based approach utilized a U-Net architecture enhanced with custom activations, attention blocks, and a novel skip-connection design intended to strengthen feature propagation. This involved creating a more complex and parameter-heavy model, with approximately three million parameters to be tuned. Although the architecture was conceptually sound and demonstrated theoretical benefits in feature refinement and spatial precision, the relatively small dataset size posed a significant challenge. The complexity and parameter count of the U-Net-based model, coupled with the limited training samples, made it difficult to achieve optimal convergence and generalization. As a result, the segmentation approach did not perform as favorably as the classification approach under the given constraints.

**Conclusion:**  
Overall, the classification pipeline outperformed the segmentation approach, likely due to its modular nature and parameter flexibility, which better suited the modest dataset size. While the segmentation approach offers richer feature extraction and spatial detail, it was hampered by its parameter-intensive architecture and insufficient data. Future work will focus on reducing the parameter count, exploring more aggressive regularization strategies, and potentially incorporating data augmentation or transfer learning techniques for the segmentation pipeline. The attached results and images illustrate the respective outcomes and highlight areas for further refinement.

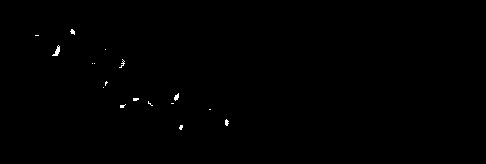
P.S. I have generated above report with chatgpt after explaining it my approach that I have used for better understanding of reader



PREDICTION



ACTUAL

PREDICTION ACTUAL

VIT-B-32(CLASSIFICATION)

**Class: no\_scratches**

**Precision: 0.9390**

**Recall: 0.9625**

**F1 Score: 0.9506**

**Class: scratches**

**Precision: 0.8333**

**Recall: 0.7500**

**F1 Score: 0.7895**

Convnext\_base\_w(CLASSIFICATION)

