# Weakly Supervised Segmentation in Medical Imaging: A Counterfactual Approach

Dmytro Shvetsov, Joonas Ariva [SUP], Marharyta Domnich[SUP], Dmytro Fishman[SUP]

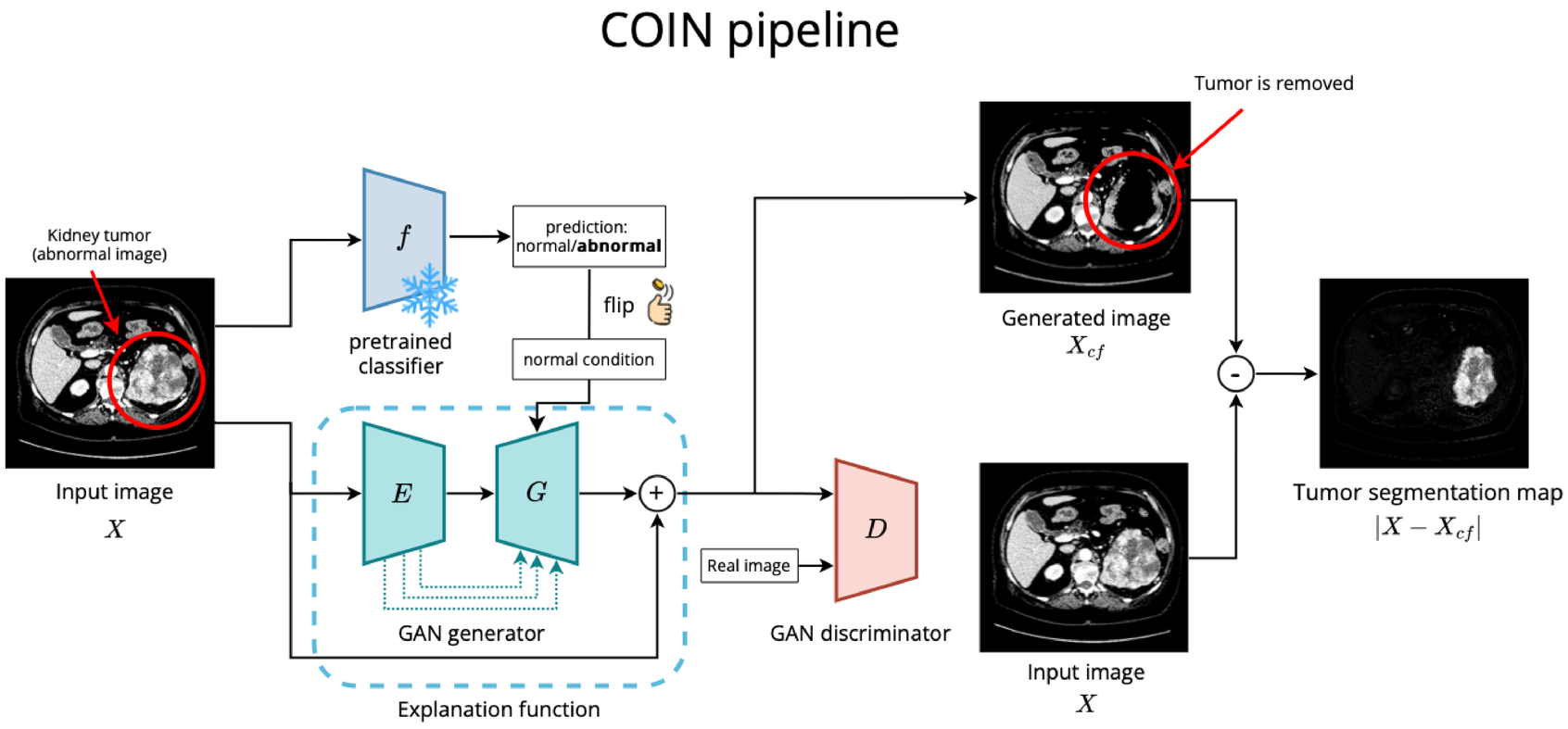


Figure 1: Overview of the proposed Counterfactual Inpainting (COIN) pipeline [1]. Given the input image $X$ and image classifier $f$, the GAN model inpaints the pathology. The absolute difference of the original image $X$ and counterfactual image $X_{cf}$ results in a weak tumor segmentation map.

## Comparison of COIN with other methods

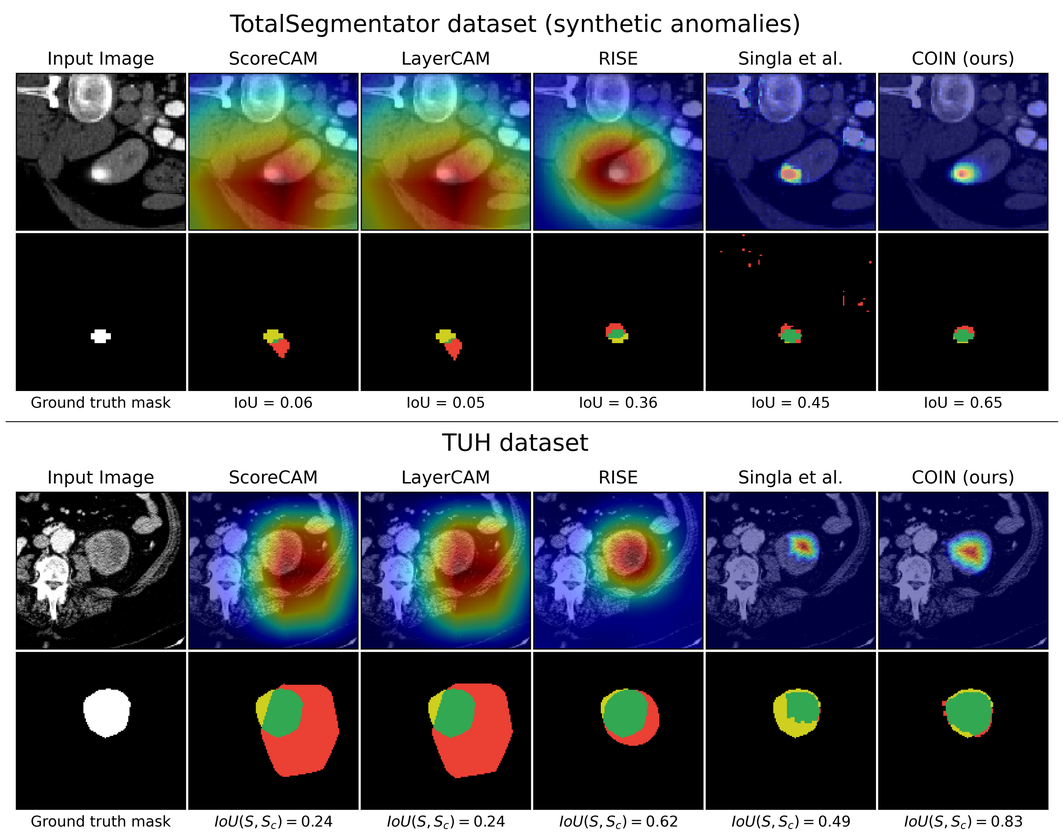


Figure 2: Visualization of the attribution and COIN pipeline methods' predictions on TotalSegmentator and TUH datasets [1]. The colors represent **ground truth** (white), **true positives** (green), **false positives** (red), and **false negatives** (yellow).

| Datasets | Methods | FID ↓ | CV ↑ | IoU ↑ |
|---|---|---|---|---|
| TotalSegmentator | ScoreCAM | - | - | 0.030 |
| | LayerCAM | - | - | 0.026 |
| | RISE | - | - | 0.397 |
| | Singla et al.* | 0.047 | 0.998 | 0.445 |
| | **COIN** | **0.003** | **0.997** | **0.646** |
| Tartu University Hospital | ScoreCAM | - | - | 0.293 |
| | LayerCAM | - | - | 0.296 |
| | RISE | - | - | 0.294 |
| | Singla et al.* | 0.203 | 0.992 | 0.352 |
| | **COIN** | **0.036** | **0.980** | **0.432** |

Table 1: Metric results for the attribution methods and the COIN pipeline on TotalSegmentator and TUH datasets [1].

## Conclusion

Our innovative Counterfactual Inpainting (COIN) approach inspired by the work of Singla et al. [2] accurately segments pathology regions in CT scans without reliance on the existing segmentation masks. COIN's main architectural improvements are:

1) **Perturbation-based** image generation yields great fidelity counterfactual images.
2) **Simplified conditioning** to focus only on the inpainting counterfactual generation.
3) **Skip-connections** for improving the image generation quality of the counterfactuals.
4) **New loss function** for enforcing smoothness among counterfactual images.

## COIN examples on Tartu University Hospital dataset

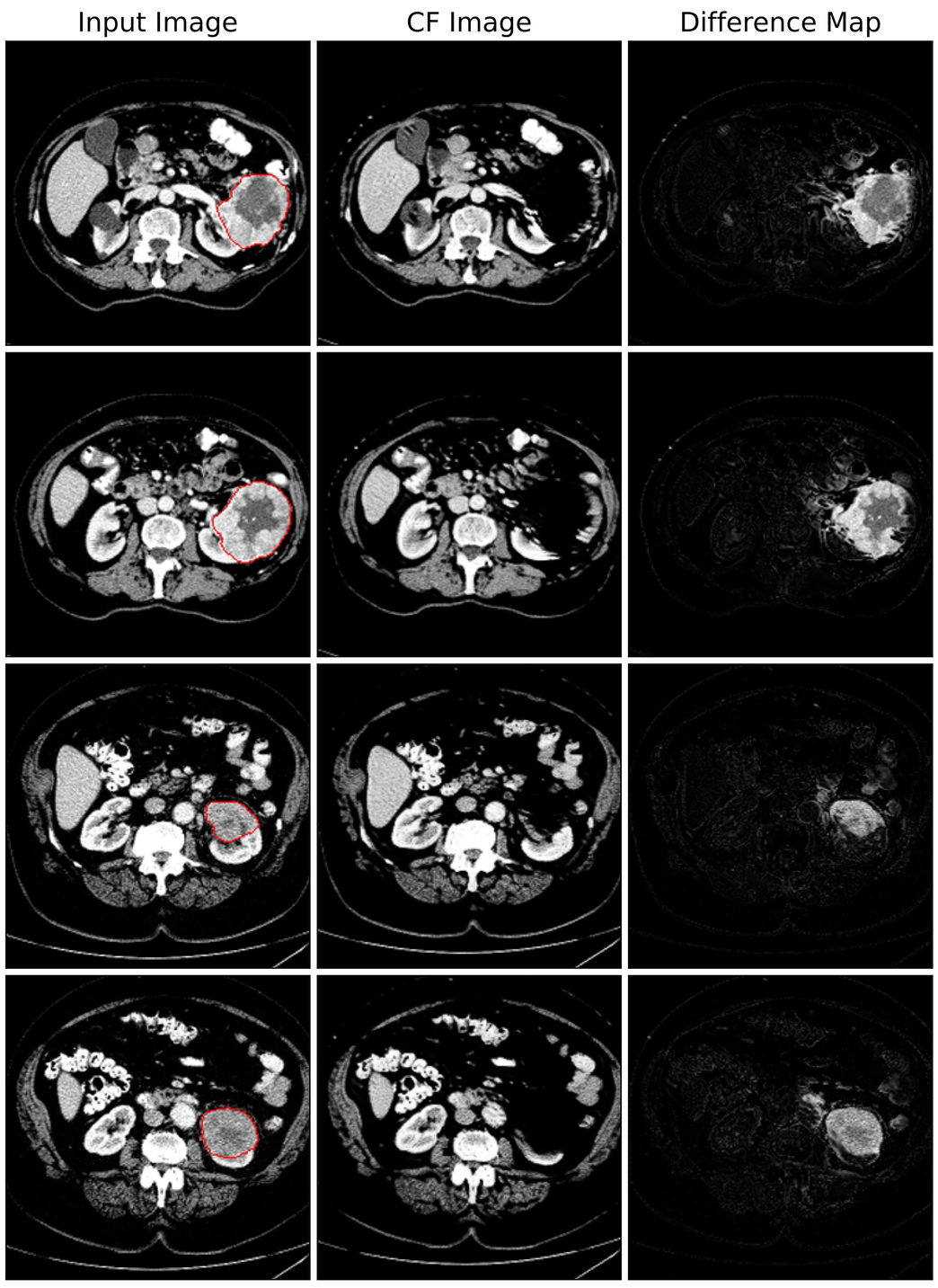


Figure 3: Visualization of counterfactual examples generated with COIN for TUH validation set [1]. The model correctly inpaints majority of the tumors (red contours).

## Acknowledgements

We thank Better Medicine for the provided dataset, as well as the High Performance Computing Center of the Institute of Computer Science at the University of Tartu for the provided computing resources.

## References

[1] D. Shvetsov et al. *COIN: Counterfactual inpainting for weakly supervised semantic segmentation for medical images*. Apr. 19, 2024. arXiv: 2404.12832[cs].

[2] S. Singla et al. *Explaining the Black-box Smoothly- A Counterfactual Approach*. Nov. 18, 2022. arXiv: 2101.04230[cs,eess].