Hybrid Semi-Supervised Framework for Sparse Annotation–Efficient Abdominal CT Multi-Organ Segmentation

Advancing medical imaging with minimal data.

Abstract

Addressing Annotation Challenges in Medical Imaging

The Problem

Manual segmentation of abdominal CT scans is highly time-consuming, suffers from inter-observer variability, and demands extensive expert annotations. Leveraging sparse annotations (points, scribbles) is a realistic yet underutilized approach.

Our Solution

We propose a novel hybrid deep learning framework. It integrates CNN–Transformer backbones, a robust semi-supervised learning strategy (teacher–student with consistency regularization), and advanced shape priors (distance map and topological losses). This framework efficiently uses limited sparse labels alongside large volumes of unlabeled CT data.

Expected Impact

- Achieve segmentation accuracy comparable to fully-supervised nnU-Net, utilizing only 10–20% of labeled data.
- Enhance generalization across diverse abdominal organs (pancreas, liver, kidneys, spleen, bowel).
- Enable efficient, clinically deployable segmentation, significantly reducing annotation costs.

Introduction

Revolutionizing Abdominal CT Segmentation

Accurate abdominal CT segmentation is paramount for precise diagnosis, effective therapy planning in oncology (e.g., radiotherapy and surgery), and comprehensive disease monitoring. However, the current reliance on manual delineation presents significant hurdles.

1

Manual Limitations

Manual delineation is labor-intensive, often requiring 20+ minutes per patient, and prone to inconsistencies across different medical professionals.

2

Data Scarcity

Existing supervised deep learning models demand impractically large, fully labeled datasets, a significant bottleneck in clinical translation.

3

Our Contribution

Our framework is a novel hybrid model (CNN + Transformer + semi-supervision + shape priors) specifically designed for sparse-label abdominal CT segmentation.

Literature Review

Deep Multi-Planar Co-Training (DMPCT, MICCAI 2020)

Problem: Scarcity of labeled CT scans for multi-organ segmentation, with 3D annotation being prohibitively costly.

Contribution: Introduced a pioneering semi-supervised framework that effectively leverages unlabeled CT scans through multi-planar co-training.

Dataset Used: BTCV (Beyond the Cranial Vault, MICCAI 2015).

Method: Utilized three plane-specific 2D CNNs (axial, sagittal, coronal) that co-train by exchanging pseudo-labels, significantly reducing dependence on extensive manual labels.

Results: Achieved an average Dice score of 85.7% using only 12 labeled volumes, demonstrating performance comparable to fully supervised training.

Future Scope: Potential extensions include integrating hybrid 3D models, refining pseudo-label generation quality, and adapting the approach to more diverse organs and imaging modalities.

Literature Review

AMOS Benchmark (MICCAI 2022)

Problem: A critical gap in large-scale, multi-center datasets specifically designed for abdominal multi-organ segmentation hindered progress in the field.

Contribution: Released the comprehensive AMOS benchmark, a large dataset featuring labels for 15 abdominal organs, supporting both CT & MRI modalities.

Dataset Used: Comprised of 500 labeled scans and 1,000 unlabeled scans, collected from multiple centers and across various modalities.

Method: Established a robust baseline evaluation by testing multiple state-of-the-art models including 3D U-Net, nnU-Net, and Swin-UNETR.

Results: nnU-Net and Swin-UNETR achieved the highest Dice scores (above 0.9 for large organs), though segmentation of smaller structures remained challenging.

Future Scope: Positioned to serve as a vital benchmark for semi-supervised and weakly supervised approaches, encouraging cross-modality and generalization research.

Literature Review

TotalSegmentator (Scientific Reports 2023)

Problem: The absence of a general-purpose CT segmentation model capable of covering the full human anatomy (including organs, bones, vessels, and muscles).

Contribution: Provided a robust, open-source, and pre-trained model capable of segmenting 104 distinct anatomical structures across the entire body.

Dataset Used: Leveraged 1,228 whole-body CT scans, meticulously annotated at the voxel level for comprehensive coverage.

Method: Employed an extended nnU-Net framework, specifically adapted for highly complex multi-class segmentation tasks.

Results: Achieved Dice scores exceeding 0.9 for large structures, demonstrating remarkable robustness across varying anatomies and diverse imaging protocols.

Future Scope: Opportunities for fine-tuning for specialized tasks (e.g., abdominal segmentation), extension to other modalities like MRI/PET, and scalable clinical deployment.

Literature Review

Probabilistic V-Net with Hierarchical Spatial Feature Transform (2022)

Problem: Standard Convolutional Neural Networks (CNNs) often struggle with the significant anatomical variability and inherent uncertainty in organ boundaries, leading to less reliable segmentations.

Contribution: Introduced a novel approach incorporating probabilistic modeling with Hierarchical Spatial Feature Transform (HSFT) to enforce anatomical plausibility and enhance segmentation accuracy.

Dataset Used: Evaluated on AbdomenCT-1K (1,112 scans from 12 centers), BTCV, and the TCIA pancreas dataset.

Method: Utilized a Conditional Variational Autoencoder (VAE) to model organ-specific latent distributions, combined with HSFT for injecting spatial priors directly into the segmentation process.

Results: Significantly outperformed nnU-Net and CoTr, particularly for complex organ segmentation, while also achieving 7 times faster inference speeds.

Future Scope: Explores the application of probabilistic priors in semi-supervised learning settings, aims to improve the segmentation of small organs, and expand to multi-modal imaging scenarios.

Literature Review

CoTr – Hybrid CNN–Transformer for Multi-Organ Segmentation (2021)

Problem: Conventional CNNs often fail to capture long-range dependencies crucial for accurate medical image segmentation, while Transformers typically require massive datasets and are computationally expensive.

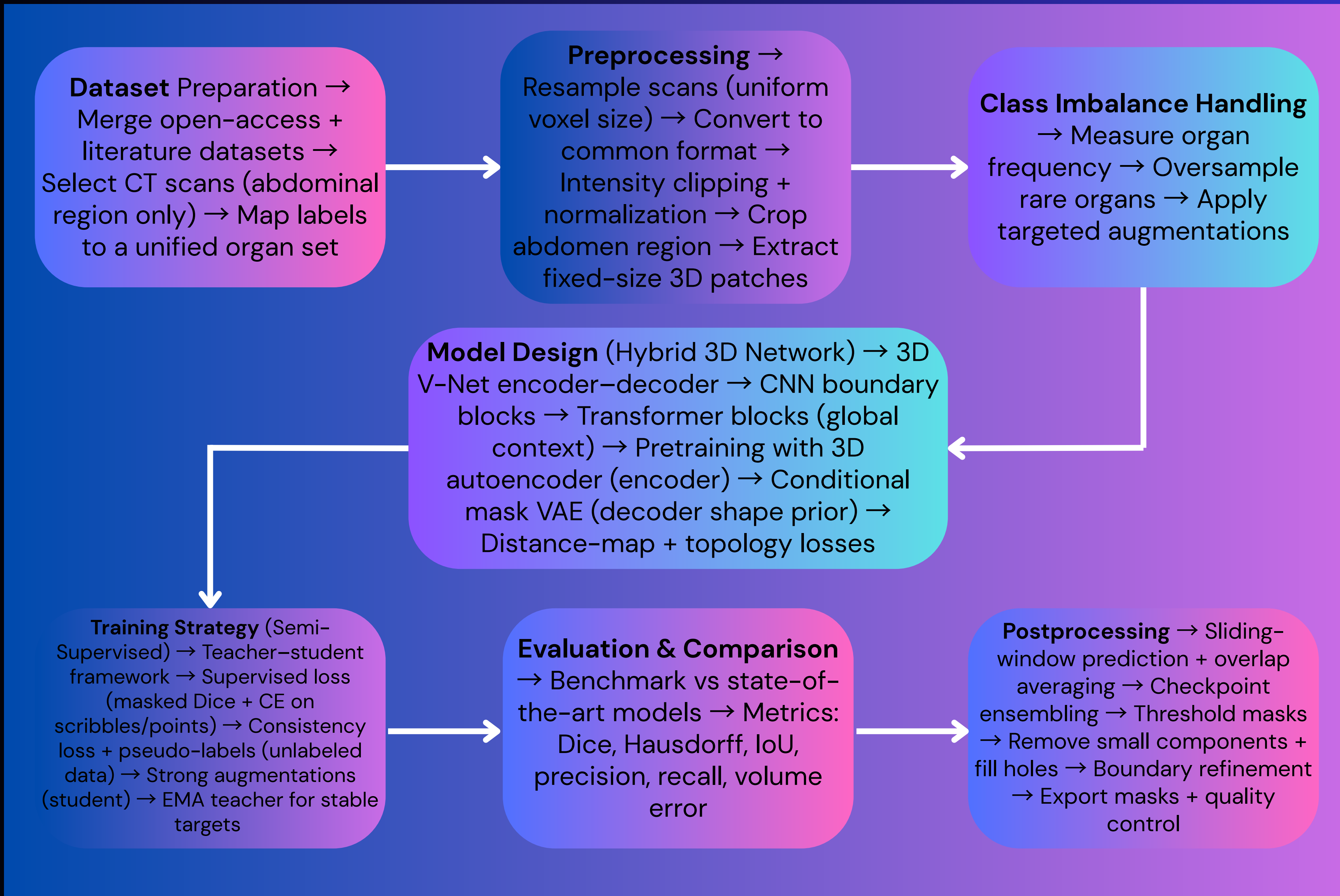
Contribution: Proposed CoTr, a hybrid model that synergistically combines the strengths of CNNs (local feature extraction) with Transformers (global context modeling).

Dataset Used: Evaluated and validated on the BTCV and AMOS benchmark datasets.

Method: Integrated Transformer layers directly within a CNN encoder–decoder architecture to enable efficient modeling of global contextual information.

Results: Demonstrably outperformed 3D U-Net and nnU-Net, especially in segmenting challenging organs like the pancreas and gallbladder.

Future Scope: Further optimization of training efficiency, adaptation for various semi-supervised learning paradigms, and exploration of lightweight Transformer variants for practical clinical deployment.



Proposed Hybrid Methodology

Looking Ahead: Impact and Future Directions

Cost-Efficient Annotation

Our proposed framework directly addresses the critical bottleneck of annotation costs in medical imaging, offering a path to significantly reduce expenses.

Enhanced Segmentation Performance

We anticipate matching or surpassing the performance of fully-supervised models, but with a fraction of the labeled data required.

Improved Clinical Utility

This work will enable more robust and efficient abdominal organ segmentation, even with sparse labels, enhancing its practical value in clinical settings.

Broad Clinical Impact

The potential impact extends to faster radiotherapy planning, more precise surgical navigation, and facilitating large-scale population imaging studies.

This framework is poised to transform the efficiency and accessibility of high-quality medical image segmentation, accelerating advancements in patient care and research.

References:

Y. Zhou et al., "Semi-Supervised 3D Abdominal Multi-Organ Segmentation via Deep Multi-Planar Co-Training," in *Proc. MICCAI*, Cham: Springer, 2020, pp. 297–306, doi: 10.48550/arXiv.1804.02586.

Y. Ji et al., "AMOS: A Large-Scale Abdominal Multi-Organ Benchmark for Versatile Medical Image Segmentation," in *Proc. MICCAI*, Cham: Springer, 2022, pp. 91–102, doi: 10.48550/arXiv.2206.08023.

J. Wasserthal et al., "TotalSegmentator: Robust Segmentation of 104 Anatomical Structures in CT Images," *Scientific Reports*, vol. 13, no. 1, p. 101, Jan. 2023, doi: 10.48550/arXiv.2208.05868.

W. Zhang et al., "Probabilistic V-Net With Hierarchical Spatial Feature Transform for Multi-Organ Segmentation," *Medical Image Analysis*, vol. 77, p. 102368, Feb. 2022, doi:10.48550/arXiv.2208.01382.

Y. Xie, J. Zhang, Z. Shen, and C. Shen, "CoTr: Efficiently Bridging CNN and Transformer for 3D Medical Image Segmentation," in *Proc. MICCAI*, Cham: Springer, 2021, pp. 171–180, doi: 10.48550/arXiv.2103.03024.

THANK YOU !!!

ANY QUESTIONS !?