Segmentation of Abdominal Organs Using Multitasking Deep Neural Networks

Sai Venkat Malreddy
University of California, Santa Cruz

Akashleena Sarkar
University of California, Santa Cruz

Abstract

Accurate segmentation of abdominal organs from CT scans is critical for various clinical applications. This report explores the use of boundary prediction as an extra task to improve the performance of abdominal organ segmentation. We propose a 3D encoder-decoder network designed to segment abdominal organs and their boundaries simultaneously through multi-task learning. Using the Beyond the Cranial Vault (BTCV) Abdomen dataset, we demonstrate that incorporating boundary information enhances segmentation accuracy. Our findings suggest that boundary-constrained networks provide a robust solution for improving the accuracy of abdominal organ segmentation. (https://github.com/SaiVenkatM/DL244_SegmentationProject)

1 Introduction

Segmenting multiple organs in abdominal CT scans is vital for computer-assisted surgery and organ transplantation,[1],[2] as it facilitates precise organ dose calculations required for various radiotherapy treatments. Manual segmentation by physicians, despite being common, is time-intensive, subjective, and prone to errors. With the increasing volume of CT scans, there is a critical need for automated, robust, and efficient organ-delineation tools[2],[3],[4]. These automated tools can quickly and accurately delineate abdominal structures, addressing the variability and subjectivity inherent in manual methods.

Segmenting abdominal organs from CT scans is particularly challenging due to:

- **Low Contrast**: Organs often exhibit similar intensities, making it difficult to distinguish boundaries.
- **Motion Artifacts**: Patient movement during scanning can introduce artifacts that obscure organ boundaries.

- **Anatomical Variability**: Variations in organ size, shape, and position across patients require robust segmentation models.
- **Complex Background**: The presence of surrounding tissues and organs adds complexity to the segmentation task.

Historically, methods for abdominal multi-organ segmentation relied on multi-atlas[5],[6] or statistical models,[7],[8] and some utilized handcrafted features[9],[10]. However, recent approaches using Fully Convolutional Networks (FCNs) have demonstrated superior results due to enhanced organ representation learning. FCNs, particularly those with encoder-decoder architectures, are now considered the state-of-the-art for this task[2]. These networks extract low-level features with shallow layers and high-level features with deep layers, using skip connections to retain essential information during decoding.

Current FCN-based methods use either 2D or 3D convolutional architectures[11][12]. While 2D methods are efficient in terms of memory and parameters, they lack the ability to fully utilize 3D contextual information. In contrast, 3D methods process entire CT volumes, offering better segmentation performance by leveraging rich volumetric context.

Despite these advancements, existing 3D methods often treat all parts of an organ equally, focusing solely on voxel-level information. This approach does not specifically enhance the segmentation of vulnerable regions. For instance, adjacent organs with weak boundaries can be challenging to separate, as illustrated in Figure 1, where organs like the stomach and pancreas have low-contrast, touching boundaries.

1.1 Motivation

Recognizing boundaries in medical scans is critical for both manual and automated segmentation. Recent deep learning studies have leveraged multitask learning to improve medical image segmentation by incorporating boundary features. This approach has proven effective in solving multiple tasks simultaneously and learning better representations through multiple supervisory signals.

This project proposes improving abdominal organ segmentation on CT scans by enhancing boundary delineation. We train 3D deep learning networks to predict both the boundary and the entire region of organs. This multitask approach addresses the challenge of ambiguous appearances and complex relationships between adjacent organs.

Our contributions include:

- Developing an end-to-end trainable 3D multitask learning framework that predicts both organ voxel-labels and their boundaries. This approach ensures accurate edge prediction alongside whole-organ segmentation.
- Utilizing state-of-the-art 3D encoder-decoder architectures (UNet and Attention-UNet) as baselines, we modify each to include boundary information, demonstrating significant performance improvements with minimal increase in trainable parameters.

- Validating our models on public dataset (BTCV) using metrics like Dice Score, Average Hausdorff Distance. Experiments confirm that boundary networks enhance segmentation accuracy, especially around organ borders.

2 Method

2.1 Boundary-aware Segmentation

Recent approaches have incorporated boundary information to improve segmentation accuracy[13][14][15]. By training networks to predict both organ and boundary maps, these methods achieve better delineation of organ edges. Techniques such as multi-task learning, where a single network learns to perform multiple related tasks simultaneously, have been shown to improve performance by leveraging shared features and providing additional supervisory signals.

2.2 Network Architecture

Our approach uses a 3D encoder-decoder network to segment abdominal organs and their boundaries simultaneously. The network is trained with a combined loss function that includes organ segmentation and boundary prediction losses. We explore architecture with varying degrees of weight sharing between tasks, specifically:

- **UNet-MTL-D**: A network with shared weights for the entire encoder, with separate decoders for each task.

2.3 Dataset and Preprocessing

2.3.1 BTCV Dataset

The BTCV dataset is used to evaluate our method. It includes 3D CT scans with annotations for eight organs: liver, esophagus, stomach, gallbladder, spleen, pancreas, left kidney, and duodenum.

The dataset is divided into 32 training, 5 validation, and 10 testing scans. Regions of interest are cropped and resampled to a common size of 144 $\times$ 144 $\times$ 144 voxels to standardize the input size for the network.

- Liver: 62.19%
- Esophagus: 0.5%
- Stomach: 14.59%
- Gallbladder: 0.83%
- Spleen: 9.97%
- Pancreas: 2.96%

- Left kidney: 5.83%
- Duodenum: 3.12%

2.4 Proposed Methodology

We describe the boundary loss for training 3D encoder-decoder networks to predict both boundaries and entire abdominal organ regions via multi-task learning, followed by our proposed network topology, the baseline 3D network architectures, and the architectural design of the boundary models.

2.4.1 Boundary Loss

Consider a 3D encoder-decoder network predicting voxel labels of an abdominal CT scan with dimensions $W \times H \times Z$. The network outputs a labeled voxel map of the same size as the input. To enhance representation learning using boundary information, the network predicts both the 3D organ-semantic masks and 3D organ-boundaries in a single pass. This is framed as a multi-task learning problem with shared and task-specific representations. The loss L combines organ segmentation loss L_{RS}[16] and organ boundary detection loss L_{BD}.

The organ segmentation loss L_{RS} is evaluated using multi-class dice loss, while the boundary detection loss L_{BD} employs binary cross-entropy loss. The total loss L is minimized concerning the network parameters, aiming to improve boundary accuracy and overall segmentation by penalizing erroneous contour voxels.

$$L = L_{RS} + \lambda L_{BD}$$

Where λ is a weighting factor balancing the two losses.
The organ segmentation loss L_{RS} using the dice coefficient is defined as:

$$L_{RS} = \sum_{c=0}^{C-1} \frac{2(\hat{y}_{i,c} \times y_{i,c})}{\hat{y}_{i,c}^2 + y_{i,c}^2}$$

where:

- $\hat{y}_{i,c}$ and $y_{i,c}$ denote the predicted 3D multi-organ probability map and ground-truth mask, respectively, of the ith abdominal CT scan.
- C denotes the number of organ classes.

The boundary detection loss L_{BD} using binary cross-entropy is defined as:

$$L_{BD} = -\frac{1}{N} \sum_{i=1}^{N} [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)]$$

Where y_i and $\hat{y}_i$ are the true and predicted boundary labels respectively.

2.4.2 Boundary-Constrained Network Topology

Multi-task learning can be implemented via hard-parameter sharing (shared subset of parameters for multiple tasks) or soft-parameter sharing (task-specific parameters regularized jointly). We use hard-parameter sharing to train the encoder-decoder network for organ segmentation and boundary detection.

We explore network topology:

1. Task-Specific Decoders: A shared encoder with separate decoding arms for regions and boundaries.

2.4.3 3D Baseline Models

We use 3D UNet, and Attention-UNet as baselines, extending them to process 3D volumes. These models feature encoder-decoder designs with convolutional blocks, batch normalization, ELU activation[17], max pooling for downsampling, and bilinear interpolation for upsampling. The final convolutional layer maps features to class labels followed by softmax activation.

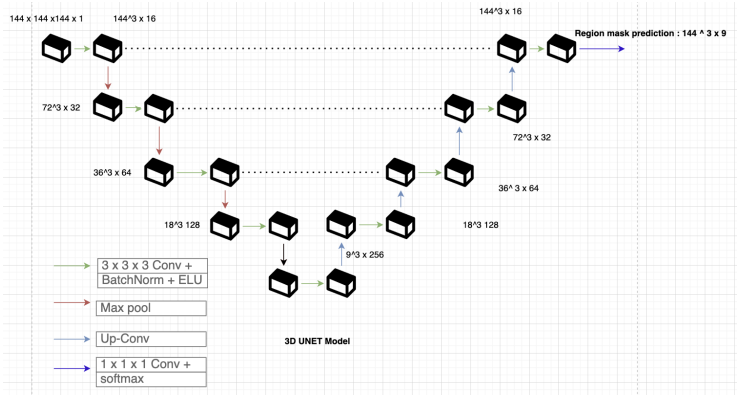

Figure 1: UNet.

2.4.4 3D Boundary Model

To incorporate boundary information, baseline models are trained to predict both organ boundaries and organs using two multi-task topologies: TSD: Features separate decoding paths for boundaries and organs, with modifications to constrain parameter count and avoid attention mechanisms in boundary-decoding.

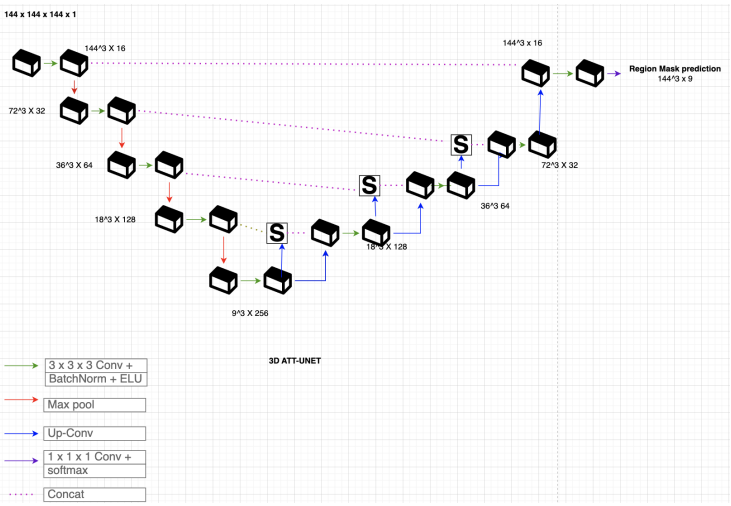

Figure 2: Attention - UNet

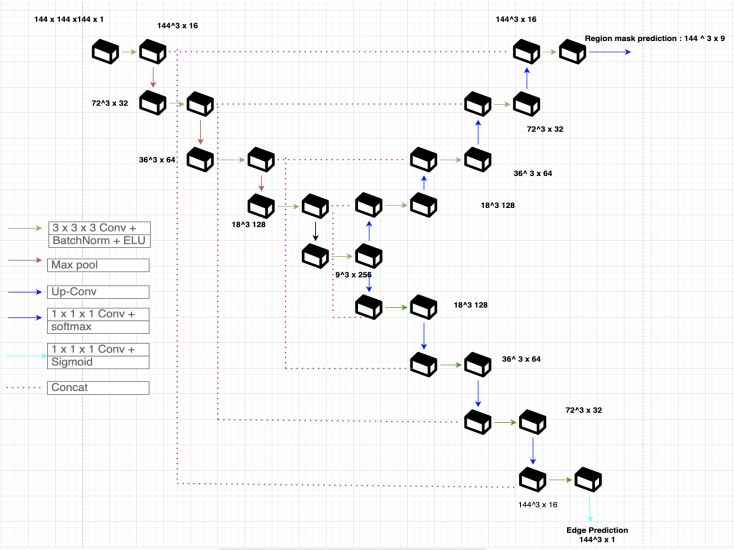

Figure 3: Boundary Attention - UNet, Extended to UNet-Att

2.5 Implementation Details

Experiments are conducted using Pytorch on Nvidia Tesla L4 GPUs. Baseline networks are trained with a batch size of 4. Training uses multi-class dice loss and the Adam optimizer with an initial learning rate of 0.001.

For boundary-aware models, we use a combination of multi-class dice loss and binary cross-entropy loss. Ground-truth boundaries are generated by eroding organ labels and comparing them to the original labels. Boundary-constrained models are trained with batch sizes of 2 for UNet-MTL-D.

3 Results

We presents the experimental results from boundary abdominal segmentation model and compares it to baseline models. The models are abbreviated as follows: 3D UNet: Standard 3D UNet. 3D Att-UNet: Attention-based 3D UNet. 3D UNet-MTL-D: 3D UNet with task-specific decoders.

3.1 Quantitative Results

Table 1 presents the segmentation results specifically for the BTCV dataset. It includes the average Dice Score and Average Hausdorff Distance (Avg. HD) obtained by comparing the predicted segmentation with the ground-truth on the test set. Notably, models incorporating boundary constraints show improved multi-organ segmentation on BTCV abdominal CT scans. There's a significant

increase in mean Dice scores across different model comparisons, indicating enhanced segmentation accuracy due to the inclusion of boundary information.

Moreover, boundary-constrained models exhibit lower Avg. HDs compared to baseline models, suggesting improved boundary prediction accuracy. This improvement persists even when boundary-constrained models achieve equivalent or lower mean Dice scores, highlighting the effectiveness of boundary constraints in refining segmentation boundaries.

In terms of computational complexity and architectural analysis (detailed in Table 2), boundary-constrained models show slightly increased parameter counts and longer segmentation times compared to baseline models, attributed to the incorporation of boundary prediction. The choice of multi-task network topology varies depending on the baseline architecture, with different topologies showing optimal performance improvements across different metrics and datasets. However, overall, integrating boundary information consistently enhances multi-organ segmentation on the BTCV dataset.

Table 1: Average Dice and HD Scores for Each Model

Model	Average Dice Score	Average HD Score
3D UNet	0.637	1.091
3D UNet-Att	0.655	1.320
3D UNet-MTL-D	0.652	0.920
3D UNet-Att-MTL-D	0.689	0.918

Boundary aware model generally show improved segmentation performance compared to baseline models. Notably, they achieve higher Dice scores and lower Avg. HD values, indicating better overlap with ground-truth masks and more accurate boundary predictions.

3.2 Computational Complexity and Architectural Analysis

Table 2 details the number of parameters and inference time for each model.

Table 2: Parameter Count and Inference Time

Method	No. of Parameters	Inference Time (ms)
3D UNet	5.89M (-)	44
3D UNet-MT-D	8.24M (+2.35M)	68
3D Att-UNet	6.47M	77
3D Att-UNet-MT-D	8.82M	89

Boundary-aware models require more parameters and longer inference times, reflecting the additional complexity and computational demands.

Table 3: Organ-wise Dice Scores

Method	Organ 1 (Dice)	Organ 2 (Dice)	Organ 3 (Dice)	Organ 4 (Dice)	Organ 5 (Dice)	Organ 6 (Dice)	Organ 7 (Dice)	Organ 8 (Dice)
3D UNet	0.678	0.559	0.697	0.567	0.698	0.602	0.713	0.566
3D UNet-Att	0.681	0.560	0.685	0.570	0.708	0.635	0.704	0.573
3D-UNet-MTL-D	0.697	0.583	0.710	0.508	0.732	0.637	0.737	0.597
3D-UNet-Att-MTL-D	0.713	0.592	0.720	0.603	0.720	0.747	0.732	0.608

Boundary-constrained models typically yield higher Dice scores across various organs compared to their baseline counterparts, indicating better segmentation accuracy.

Table 4: Organ-wise Hausdorff Distance (HD) Scores

Method	Organ 1 (HD)	Organ 2 (HD)	Organ 3 (HD)	Organ 4 (HD)	Organ 5 (HD)	Organ 6 (HD)	Organ 7 (HD)	Organ 8 (HD)
3D UNet	0.474	1.741	0.337	1.750	0.770	0.636	0.941	2.032
3D UNet-Att	0.577	2.017	0.404	2.880	0.980	0.435	1.09	2.195
3D UNet-MTL-D	0.329	1.787	0.253	1.076	0.648	0.619	0.778	1.849
3D UNet-Att-MTL-D	0.273	1.377	0.191	1.631	0.887	0.318	0.789	1.880

4 Discussion

Segmentation of abdominal organs from CT scans accurately is crucial for clinical procedures like computer-assisted surgery. However, the low contrast and weak edges in these scans make segmentation challenging. We propose using boundary information to improve 3D abdominal multi-organ segmentation. This multi-task learning model utilizes statistics from multiple ground-truth sources, retaining shared features between tasks. Boundary annotations, easily generated from ground-truth masks, provide additional, cost-free information.

Table 1 shows that our boundary-constrained 3D encoder-decoder models enhance multi-organ segmentation on the BTCV dataset across various baselines (3D UNet and 3D Att-UNet). Improvements in Dice Score, Avg. HD, Recall, and Precision are due to better boundary segmentation. These gains come with minimal increases in parameter count, demonstrating the efficiency of using boundary information for regularization.

The reduction in Avg. HD for the BTCV dataset across all baselines highlights the model's improved understanding of organ structures. Significant decreases in Avg. HD, ranging from 19% to 28%, show that the model effectively

learns features representing entire organs. Training with boundary knowledge helps develop more generalizable and discriminative features. Additional experiments confirm the advantages of boundary information, particularly near organ boundaries.

Tables 3 and 4 show that boundary-regularized models significantly enhance the segmentation of rare organs in the BTCV dataset. Figures 5 demonstrate the positive impact of boundary information, reducing over-segmentation and under-segmentation. The model's ability to learn better representations for multiple organs simultaneously underscores the effectiveness of this approach.

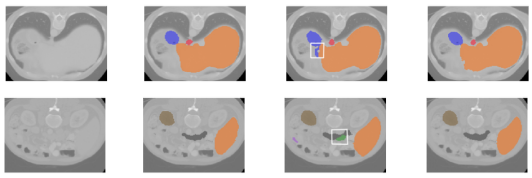

Figure 4: Columns 1-2 display the original image and the corresponding ground-truth mask overlaid on the 2D image and 3-4 columns results for the baseline UNet and its UNet-MT-D counterpart. The white boxes highlight regions where segmentation has been enhanced by incorporating boundary information.

5 Conclusion

We enhance 3D abdominal multi-organ segmentation by integrating organ boundary information to address the challenge of unclear boundaries in low-contrast CT scans. We found that the most effective multi-task topology varies depending on the baseline architecture. Incorporating boundary features improves the segmentation of less prominent organs like the gallbladder, pancreas, and duodenum with minimal additional parameters. These findings suggest a small potential of our approach in enhancing abdominal organ segmentation for clinical applications.

References

1. T. Okada, M. G. Linguraru, M. Hori, Y. Suzuki, R. M. Summers, N. Tomiyama, Y. Sato, Multi-organ segmentation in abdominal CT images, in: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2012, pp. 3986–3989.

2. E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S. P. Pereira, M. J. Clarkson, D. C. Barratt, Automatic multi-organ segmentation on abdominal CT with dense V-networks, IEEE transactions on medical imaging 37 (8) (2018) 1822–1834.

3. P. Hu, F. Wu, J. Peng, Y. Bao, F. Chen, D. Kong, Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets, International journal of computer assisted radiology and surgery 12 (3) (2017) 399–411.

4. H. Kim, J. Jung, J. Kim, B. Cho, J. Jang, S.-w. Lee, J.-G. Lee, S. Yoon, Abdominal multi-organ auto-segmentation using 3d-patch-based deep convolutional neural network, Scientific Reports 10 (2020) 6204. doi: 10.1038/s41598-020-63285-0.

5. Z. Xu, R. P. Burke, C. P. Lee, R. B. Baucom, B. K. Poulose, R. G. Abramson, B. A. Landman, Efficient multi-atlas abdominal segmentation on clinically acquired CT with simple context learning, Medical image analysis 24 (1) (2015) 18–27.

6. M. Suzuki, M. G. Linguraru, K. Okada, Multi-organ segmentation with missing organs in abdominal CT images, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2012, pp. 418–425.

7. J. J. Cerrolaza, M. Reyes, R. M. Summers, M. A. González-Ballester, M. G. Linguraru, Automatic multi-resolution shape modeling of multi-organ structures, Medical image analysis 25 (1) (2015) 11–21.

8. T. Okada, M. G. Linguraru, M. Hori, R. M. Summers, N. Tomiyama, Y. Sato, Abdominal multi-organ segmentation from CT images using conditional shape–location and unsupervised intensity priors, Medical image analysis 26 (1) (2015) 1–18.

9. P. Campadelli, E. Casiraghi, S. Pratissoli, G. Lombardi, Automatic abdominal organ segmentation from CT images, ELCVIA: electronic letters on computer vision and image analysis 8 (1) (2009) 1–14.

10. M. A. Selver, Segmentation of abdominal organs from CT using a multilevel, hierarchical neural network strategy, Computer methods and programs in biomedicine 113 3 (2014) 830–52.

11. M. P. Heinrich, O. Oktay, N. Bouteldja, Obelisk-net: Fewer layers to solve 3d multi-organ segmentation with sparse deformable convolutions, Medical image analysis 54 (2019) 1–9.

12. M. F. Bobo, S. Bao, Y. Huo, Y. Yao, J. Virostko, A. J. Plassard, I. Lyu, A. Assad, R. G. Abramson, M. A. Hilmes, et al., Fully convolutional neural networks improve abdominal organ segmentation, in: Medical Imaging 2018: Image Processing, Vol. 10574, International Society for Optics and Photonics, 2018, p. 105742V.

13. A. Sinha, J. Dolz, Multi-scale self-guided attention for medical image segmentation, IEEE Journal of Biomedical and Health Informatics (2020).

14. H. Lu, S. Tian, L. Yu, L. Liu, J. Cheng, W. Wu, X. Kang, D. Zhang, DCACNet: Dual context aggregation and attention-guided cross deconvolution network for medical image segmentation, Computer Methods and Programs in Biomedicine 214 (2022) 106566. doi:10.

15. H. Chen, X. Qi, L. Yu, P.-A. Heng, DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2487–2496.

16. C. Tan, L. Zhao, Z. Yan, K. Li, D. Metaxas, Y. Zhan, Deep multi-task and task-specific feature learning network for robust shape preserved organ segmentation. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 1221–1224. doi: 10.1109/ISBI.2018.8363791.

17. R. Zhao, W. Chen, G. Cao, Edge-boosted U-Net for 2D Medical Image Segmentation. IEEE Access, 7, 171214–171222. doi: 10.1109/ACCESS.2019.2953727.

18. H. J. Lee, J. U. Kim, S. Lee, H. G. Kim, Y. M. Ro, Structure boundary preserving segmentation for medical image with ambiguous boundary. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.