A cascaded deep network for automated tumor detection and segmentation in clinical PET imaging of diffuse large B-cell lymphoma

S. Ahamed [1, 2], N. Dubljevic [1], I. Bloise [4], C. Gowdy [5], P. Martineau [4], D. Wilson [4], C. F. Uribe [3, 4], A. Rahmim [1, 2, 3], F. Yousefirizi [2]

[1] Department of Physics & Astronomy, University of British Columbia, Vancouver, BC, Canada
[2] Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, BC, Canada
[3] Department of Radiology, University of British Columbia, Vancouver, BC, Canada
[4] BC Cancer, Vancouver, BC, Canada
[5] BC Children's Hospital, Vancouver, BC, Canada

Quantitative Radiomolecular Imaging & Therapy

INTRODUCTION

- Accurate **detection** and **segmentation** of diffuse large B-cell lymphoma (DLBCL) tumors in PET images is important for **total metabolic tumor volume (TMTV)** calculation [1].
- **Manual segmentation** by physicians is time-consuming, labor intensive, operator dependent, with high intra- and inter-operator variability [2].
- A **single** end-to-end segmentation network such as U-Net on whole-body PET images is not very efficient, especially when the disease is generalized or the tumors are small and occur in close proximity to one another [3].
- Removing axial slices not intercepting a tumor (**background slices**) from the PET images and predicting **crude regions of interest** (ROIs) around the suspicious regions in the slices intercepting tumors (**foreground slices**) before the segmentation step would improve performance by removing **extraneous** part of the image and **centering** the object of interest in the receptive field of view.

OBJECTIVES

- Implementing a fast and efficient deep-learning based three-step network that takes in a 3D PET image as input and outputs the segmentation contours for each of the 2D axial foreground slices.
- In the first step, a **Slice classifier** (Module-1) classifies the 3D PET axial slices into foreground and background slices.
- In the second step, a **Tumor detector** (Module-2) localizes tumors via bounding boxes on the foreground axial slices obtained from the first step.
- In the third step, a **Tumor segmentor** (Module-3) predicts the segmentation contours inside the bounding boxes obtained from the second step (Figure 1).

METHODS

Dataset

- Our lymphoma PET images dataset consisted of **126 annotated cases of primary mediastinal large B-cell lymphoma (PMBCL)** (for pre-training) and **50 annotated cases of DLBCL** [4].
- Training, validation and testing were performed on the **axial slices** of 3D whole-body PET images. A **60%:20%:20%** splitting was used for training, validation and testing, respectively.
- Class imbalance (proportion of foreground slices and background slices) for both PMBCL and DLBCL datasets was about 10%:90%, across all training, validation and testing sets.

METHODS

Network architecture and training

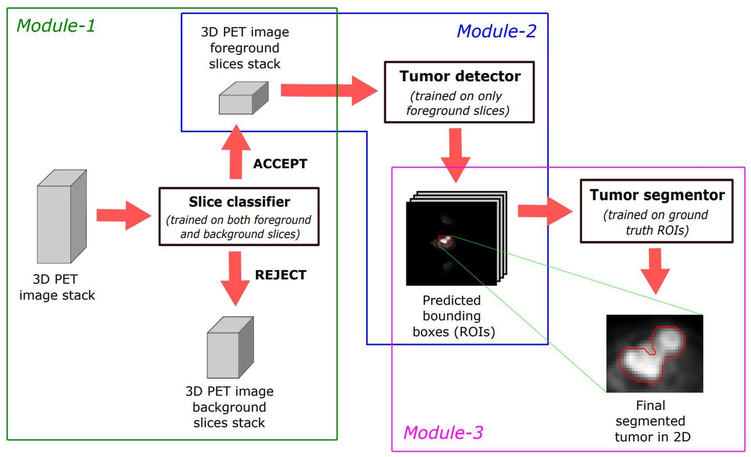

Figure 1: The proposed cascaded 3-step model for PET tumor segmentation.

***1. Module-1 (Slice classifier)*:**

- This is a binary classification network that uses **ResNet152** backbone pre-trained on ImageNet dataset.
- Pre-training and testing on 31126 PMBCL axial slices, followed by fine-tuning and testing on 13185 DLBCL axial slices.
- The classification loss function used was **focal loss**, $\mathcal{L}_{Focal}$ [5].

$$\mathcal{L}_{\text{classification}} = \mathcal{L}_{\text{Focal}}$$

***2. Module-2 (Tumor detector)*:**

- This is an object detection network based on **Faster R-CNN with a ResNet50** backbone pre-trained on ImageNet dataset.
- Pre-training and testing on 2958 PMBCL foreground axial slices, followed by fine-tuning and testing on 1217 DLBCL foreground axial slices.
- The detection loss function was a weighted-sum of **cross-entropy loss**, $\mathcal{L}_{class}$ (for class prediction) and **L1 regression loss**, $\mathcal{L}_{box}$ (for bounding box coordinates prediction) as given below (with $\lambda = 10.0$).

$$\mathcal{L}_{\text{detection}} = \mathcal{L}_{\text{class}} + \lambda \mathcal{L}_{\text{box}}$$

***3. Module-3 (Tumor segmentor)*:**

- This is a segmentation network based on **2D U-Net** architecture modified for smaller images.
- Pre-training and testing on 4057 PMBCL ROIs, followed by fine-tuning and testing on 1679 DLBCL ROIs.
- The segmentation loss function was a weighted-sum of **generalized Dice loss** $\mathcal{L}_{generalizedDice}$ [6] and **focal loss**, $\mathcal{L}_{Focal}$ as given below (with $\lambda = 10.0$).

$$\mathcal{L}_{\text{segmentation}} = \mathcal{L}_{\text{generalizedDice}} + \lambda \mathcal{L}_{\text{Focal}}$$

RESULTS AND DISCUSSION

- The performance of the three modules on the DLBCL test set are summarized in Table 1 (a), (b), and (c), respectively. The proposed three-step model achieved a 2D Dice score of **77.9% ± 13.2%** and a 3D Dice score (upon aggregation of 2D predictions) of **78.1% ± 8.6%** (Table 1 (c)).
- The segmentation performance of our model was compared to that of a single end-to-end **3D U-Net model** [7], which achieved a 3D Dice score of **58.9% ± 16.1%**. We demonstrate a segmentation **performance improvement by about 19%** across all the cases in the DLBCL test set.

- **Rejecting the background slices** before detection eliminates the need to train the tumor detector on the background slices.
- **Segmentation inside bounding boxes** (instead of whole-body images) improves performance by removing extraneous parts of the image and centering the tumor in the receptive field-of-view.

(a) Metric	Score
Specificity	90%
Sensitivity (Recall)	83%
Precision	43%
F1-score	57%
Area under ROC curve	0.93

(b) Metric	Score
Detection accuracy	81%
Mean average precision	69%

(c) Model	Dice score
3D U-Net	58.9% ± 16.1% (3D)
Proposed model	77.9% ± 13.2% (2D)
	78.1% ± 8.6% (3D)

Table 1: Performance of the Slice classifier (a), the Tumor detector (b), and the Tumor segmentor (c) on the DLBCL test set.

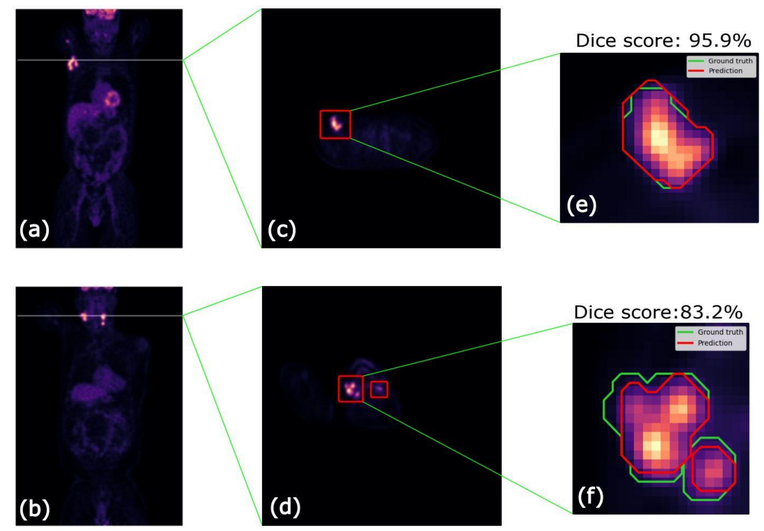

Figure 2: Performance of the tumor detector and segmentor modules: (left column) (a)-(b) show 2 representative DLBCL PET images in the coronal view. (middle column) (c)-(d) show the corresponding selected axial slices (white horizontal lines in (a)-(b)) with the predicted bounding boxes around the tumors (shown in red). (right column) (e)-(f) show the corresponding ground truth (shown in green) and predicted (via our implemented 2D U-Net, shown in red) segmentation contours in 2D for the detection tumors in (c)-(d).

REFERENCES

[1] *Vercellino, L.*, *et al,* Blood, vol.135, no. 16, 2020
[2] *Willemink, M.J.*, *et al,* Radiology, 2020.
[3] *Weisman, A.*, *et al*, Phys Med Biol 65, 2020.
[4] *Yousefirizi, F.*, *et al*, Eur J Nucl Med Mol Imaging 48 (1), 2020.
[5] *Lin*, *T.Y.*, *et al*, IEEE Trans Pattern Anal Mach Intell 42 (2), 2020.
[6] *Sudre*, *C.H.*, *et al*, DLMIA 2017, ML-CDS 2017.
[7] *Ronneberger*, *O.*, *et al*, MICCAI 2015.