

Hybrid Tumor Segmentation Framework Integrating Generative Adversarial Networks with Topological Active Nets

Shruthi N¹ and Manju N²

^{1,2}Department of Information Science and Engineering, JSS Science and Technology University, Mysuru, India
Email: shruthin@jssstuniv.in

Abstract—Earlier detection of Lung cancer involves segmenting tumor regions in whole body Positron Emission Tomography (PET)/ computed tomography (CT) images. Compared to segmenting tumor on isolated regions, segmenting tumor on whole body region is complex due to its higher metabolic similarity with surrounding tissues. Though many methods have been proposed to segment lung tumors, their tumor detection performance outside thorax region is poor. Thus, they cannot be used successfully for segmenting tumor in whole body regions. This work proposes a deep learning integrated with topological active net (TAN) for accurate segmentation of tumor from whole body scans. TAN is applied at first level to get approximate boundaries for legions which is then processed by deep learning Generative Adversarial Network (GAN) to get the accurate location of legion by filtering the noises.

Keywords—Lung cancer, Tumor segmenation, whole body PET/CT, Camouflaged object detection, Deep learning.

I. INTRODUCTION

One of the main causes of cancer-related fatalities globally is lung cancer. Among the cancer related deaths, 18% is contributed due to Lung cancer [1]. Earlier detection and treatment is the way to reduce the mortality. Metabolic Tumor Volume (MTVwb) is an important indicator for earliest detection of Lung cancer. This indicator is calculated by segmenting the tumor regions in whole body PET/CT images. Tumor regions around thorax regions appear in later stages and thus segmenting tumor in whole body region is the way forward for earlier detection of Lung tumor. Segmenting tumor regions from whole body PET/CT images and calculating the MTVwb from isolated segments through manual effort is time consuming and can be erroneous. These presents the case for automatic segmentation of tumor regions and calculate MTVwb [3]. Various automatic tumor segmentation approaches based on active shape models, texture features based classification, graph based methods etc have been proposed in earlier works with limited accuracy. Recently deep learning methods have been proposed to segment tumor regions. Effectiveness of deep learning solutions depends on the training dataset volume. The problem with training deep learning classifier with whole body PET/CT images for tumor segmentation is that they become overfit and has higher false positives. It is challenging to segment tumor in whole body region due to presence of similar tissues density surrounding tumor regions. Though deep learning methods provide higher accuracy in detecting tumor around thorax regions, their accuracy when applied to whole body region is less due higher tissue similarity.

This work proposes an integrated solution combining topological active net with GAN to segment the tumor region. The input CT and PET images are first segmented by TAN to get characteristic regions map. GAN is trained to identify the characteristic regions containing the tumor legions in form of mask. The masks generated for CT and PET images using GAN are fused to get the final segmentation map. By use of TAN, the effective region for tumor legion identification is reduced. Through use of fusion of GAN generated masks, the legions are detected with higher accuracy and lower false positives.

Following are the novel contributions of this work

- A novel improved TAN technique to identify characteristic region in the image where the probability of tumor legion localization is higher.
- A fusion algorithm to identify the segmentation mask using the probable characteristic region mask generated for the PET and CT images by the GAN.

II. LITERATURE REVIEW

Jemma et al [2] used convolutional neural networks(CNN) to detect tumor from whole body PET/CT scans. 2D and 3D CNN was used to segment tumor from eyes to thighs in the whole body scans. Though the method was able to identify tumor regions, the false positives were higher as it could not differentiate the similarity of tissues near tumor region. He et al [3] transformed whole body PET/CT scans to maximum intensity projections. Deep neural network is applied at first step to segment tumor regions. In the second step, COD is applied to differentiate tumor regions from surrounding tissues. The accuracy is low at 70% in this method. Meng et al [4] proposed a dual path convolutional network to detect tumor regions. Feature map from two networks are fused to detect tumor segments. To reduce the false positives, conditional random fields are applied. The method works well for thorax regions but for whole body scans, the impact of surrounding tissues is higher and false positives cannot be removed by conditional random field alone. Liu et al [5] proposed a spatial CNN to segment tumor regions from CT images. Optimizations were done in the CNN inform of skip connections, feature fusion, multi scale features etc to improve the effectiveness of segmentation. 3D conditional random fields were applied to reduce false positives. The precision is only 38% in this method. Alirr et al [6] combined both convolutional neural network and region level set method to segment tumor. In the first stage CNN is applied to segment tumor at coarse level. In the second stage localized region level set function is applied to refine the segment region. The method works only for tumor detection in liver envelope and does not work for whole body scans.

Shiri et al [7] combined convolutional neural learning with output level voting based fusion to segment tumor regions from PET/CT images. The method worked well for selected regions like neck with 86% accuracy but without consideration for similarity tissues around tumor region, it cannot be used for whole body scans. Li et al [8] extracted cross model features at multiple resolutions from CT images and used it for segmenting tumor regions. The segmentation accuracy is improved in this method by using cross model information learnt from CT and PET images. But the cross modality approach did not consider tumors hidden by similar tissues, so the accuracy in whole body scans will be low. Kevin et al [9] proposed a semi supervised transfer learning based approach to detect tumor regions from PET/CT images. The method works well for thorax region. Without consideration for similar tissue noises around tumor regions, the accuracy for whole body scans will be low. Li t al [10] proposed two stage variational deep learning model to segment tumor regions. In the first stage, 3D CNN was used to generate a tumor probability map from the CT image. In the second stage, the probability map and the intensity information extracted from PET image are fused. This fused image is used for segmenting the tumor region. Though the method works well for thorax regions, its performance was not tested for whole body scans. Also the method has not considered the problem of similarity of tissues affecting accuracy of tumor detection. Sibille et al [11] proposed a cascaded approach for tumor segmentation. Stacked ensemble of 3D CNN processed the PET/CT images to detect tumor regions.. The accuracy is only 68% in this method. Yuan et al [12] proposed a deep learning segmentation network using encoder decoder transformers. The method fused images at multiple scales to generate fused features which are then used for segmentation. Dynamic scale attention mechanism was used to fuse multi scale features. The accuracy is only 73% in this method. Dong et al [13] utilized a 3D UNet structure and took the concatenation of PET, CT, and Gaussian noise volumes as the network input to generate the tumor mask. The method works only for certain regions in PET/CT scans.To exploit low-level and high-level semantics at full sizes, Bourigault et al. [14] suggested a new network based on an encoder-decoder architecture with complete inter- and intra-skip connections. In order to improve the anticipated segmentation maps, the authors also employed Conditional Random Fields as a post-processing step. But the accuracy is only 62% in this method. For HNC patients, Naser et al. [15] suggested a unique deep learning architecture for primary tumour auto-segmentation. In order to reduce overfitting, the authors preprocessed

PET/CT images by normalising intensities and using data augmentation. Both 2D and 3D convolutional neural networks based on the U-net architecture were put into practice. These networks were optimised using a model loss function that combined binary cross entropy and dice similarity coefficient (DSC). The accuracy of the approach was only 69%. Leung et al [16] proposed a transfer learning approach for segmenting tumor from PET/CT images. A uu-Net deep learning network is used to identify probable tumor segments. From this probable map, further refinement is done higher order radiomics to mark the segments with confidence. The method worked for head and neck regions, but its performance near metabolic similar tissue regions were not tested. Froelich et al [17] trained CNN with various hyperparameter optimizations to detect tumor from PET/CT images. Authors also applied refinement to base data to remove bias. The method was able to achieve only 70% accuracy. The false positives were higher due to in ability to differentiate similar tissue regions from tumor. Häggström et al [18] used maximum intensity project, 3D CT and 3D PET images as input to Resnet24 deep learning model to detect tumor. Though the model is able to provide accuracy in detection of tumor, fine grained test at metabolic similar tissues were not tested. Durand et al [19] used a 3D U-net CNN to segment tumor from PET/CT images. Two channel U-net was used to process PET and CT images separately and features fused to detect segments. Authors used summed cross entropy and dice similarity coefficient loss as training functions. The work did not consider the influence of surrounding similar tissues in tumor segmentation. Yuan et al [20] proposed hybrid learning feature fusion method for tumor segmentation. Multi modal features extracted from PET and CT images separately fused to detect tumor segments. But the method did not consider the influence of surrounding similar tissues in tumor segmentation. As result, the false positives will be higher for those cases. Survey discussed various methods that have been proposed for automatic tumor segmentation. Though they perform well for segmenting tumor regions around thoracic regions, their performance for other whole body regions are poor. This is due to following factors

- It can be challenging to distinguish the margins of extra-thoracic tumours on CT scans because their tissue density may be equivalent to that of the surrounding soft tissue.
- On PET visuals, normal organs and tissues may resemble tumours because they are metabolically active (FDG-avid). The brain, spleen, liver, heart, bone marrow, lymphoid tissue, and brown fat among a few examples.

Thus, an automatic tumor segmentation algorithm, solving the above problems to achieve higher accuracy and lower false positives for whole body scan is needed.

III. METHODOLOGY

The architecture of the proposed solution integrating TAN with GAN for segmentation of tumor legions from whole body PET/CT scan is given in Figure 1. The segmentation is built on two concepts of TAN and GAN. TAN is geometric deformable model. A grid mesh placed over the object is deformed in such a way mesh grids are removed where objects are present. This is done by calculating energy function between the interrelated nodes in the mesh based on intensity of pixels near their boundaries. Deformation is done continuously till the energy function is minimized between the interrelated nodes. When the minimization is completed, the mesh is over the object.

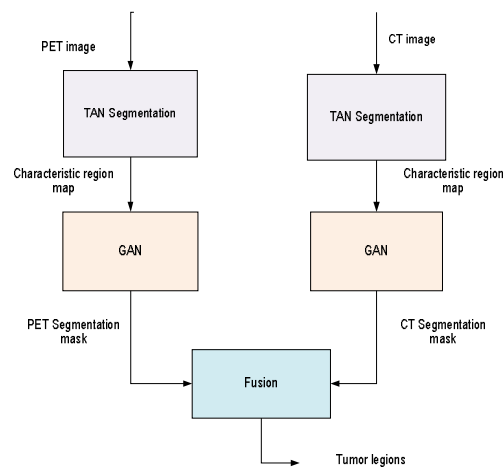


Figure 1: Integrated TAN-GAN Architecture for Segmentation

Algorithm

Algorithm: Integrated TAN-GAN Architecture for Segmentation

Step 1: Input: PET image, CT image

I_{PET} : Input PET image

I_{CT} : Input CT image

Step 2: TAN Segmentation

$$C_{r1} = \text{TAN}(I_{PET})$$

$$C_{r2} = \text{TAN}(I_{CT})$$

Where, $C_{r1} \leftarrow \text{TAN segment (PET image)}$

$C_{r2} \leftarrow \text{TAN segment (CT image)}$

$C_{r1}, C_{r2} \in [0, 1]^{H \times W}$ are the tissue-aware coarse segmentation maps, where H and W are the height and width of the image.

Step 3: GAN Segmentation

$$G_1 = \text{GAN}(C_{r1})$$

$$G_2 = \text{GAN}(C_{r2})$$

Where, $G_1 \leftarrow \text{GAN segment (Cr1)}$

$G_2 \leftarrow \text{GAN segment (Cr2)}$

This can be described in adversarial training form:

Generator: $G : C_r \rightarrow S$

Discriminator: $D : S \rightarrow [0, 1]$ where S is the real or generated segmentation mask.

$$\min_G \max_D \mathbb{E}_{S \sim P_{\text{real}}} [\log D(S)] + \mathbb{E}_{C_r \sim P_{\text{TAN}}} [\log(1 - D(G(C_r)))]$$

G_1 and $G_2 \in \{0, 1\}^{H \times W}$ are the binary segmentation masks.

Step 3: Intersection Fusion

Final mask is obtained by fusing G_1 and G_2 using logical intersection:

$\text{segmask} \leftarrow \text{Intersection fusion (G1, G2)}$

$$\text{Segmask}(x, y) = G_1(x, y) \cdot G_2(x, y)$$

return Segmask

Thus, TAN solves the problem of segmentation as a function of energy minimization between interrelated nodes in the grid mesh.

In this work, the first level of segmentation on input PET and CT images is done using TAN. A mesh is placed on the image and for each node energy minimization is applied to deform the mesh and fit to objects. The process of deformation is done in steps.

In each step, the node moves to minimal energy value. In the mesh, where grids are not present are made as white color. The mesh of PET and CT image are passed to GAN. GAN takes the mesh image as input and provides an improvised segmented object. GAN processing is necessary as the TAN does not provide correct segmentation boundaries when there are complex shapes and sharp concavities.

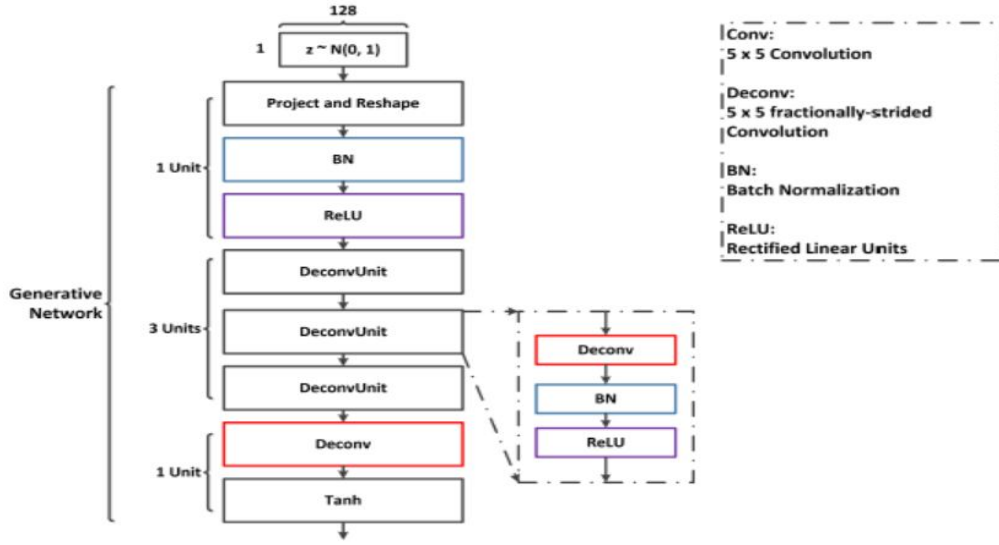


Figure 2: Generative network

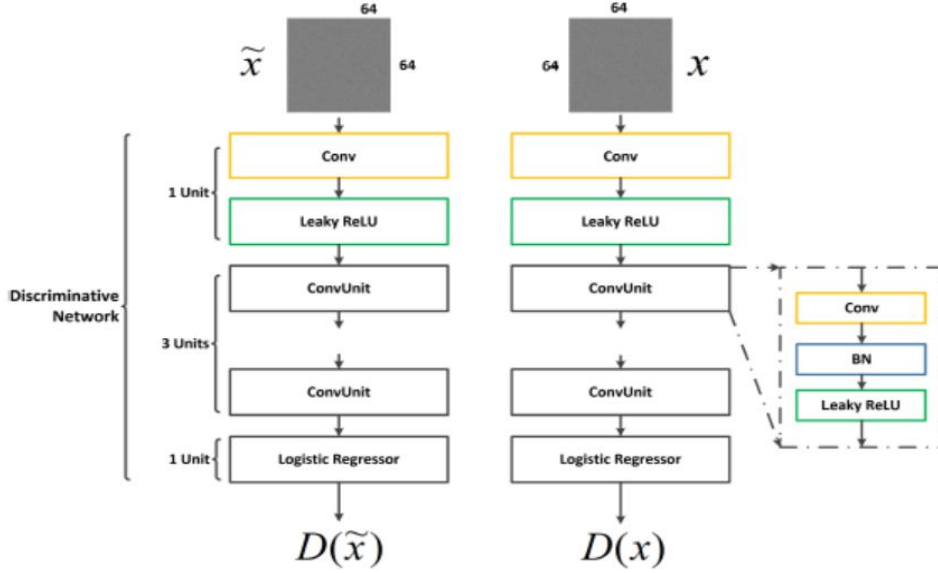


Figure 3: Discriminator network

GAN is a deep learning network that has two networks: generative network and discriminative. It is being used recently to generate synthetic samples as close to original. The Generative network is given masks generated by TAN as input and trained to give the segmentation mask as output. Generative network and discriminative network of GAN compete each other to produce accurate segmentation mask as output. The GAN is trained to provide accurate samples by using the following objective function.

$$L_{GAN} = E_{\tilde{x} \sim P_g} [D(\tilde{x})] - E_{x \sim P_r} [D(x)] + \lambda E_{\tilde{x} \sim P_x} [(||\nabla_{\tilde{x}} D(\tilde{x})||_2 - 1)^2]$$

The distribution over V is given as P_r . The distribution over which generator produces data is given as P_g . The uniform samples over P_r and P_g is given as P_x . The configuration of the GAN used in this work for generating segmentation mask from the characteristic region map is given in Figure 3 and Figure 4. GAN takes the TAN generated characteristic region map as input and provides the segmentation mask as output. The segmentation masks corresponding PET and CT images are then fused by keeping the intersection regions alone in the fused results. The fused result is the tumor legions mask. The steps in the algorithm for segmenting the tumor legions from the input PET and CT image is given in Figure 1.

IV. RESULTS AND DISCUSSION

The proposed solution's performance assessment is done using the F-FDG PET/CT dataset that is available in Cancer Imaging Archive. The tumor detection performance comparison is done by analyzing True-positive Rate (TPR), Positive Predictive Value (PPV), and False-Discovery Rate (FDR). The solution is also compared with the approach that is used to measure deep semi-supervised transfer learning described by Leung et al [16] for Lung cancer. The results for TPR, PPV, and PDR for lung cancer are shown in Figure 4.

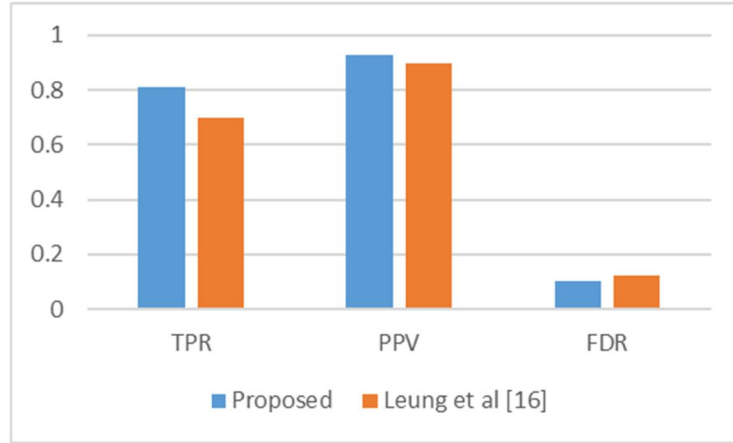


Figure 4: True-positive Rate (TPR), Positive Predictive Value (PPV) and False-Discovery Rate (FDR)

From the results, it can be seen that TPR is atleast 10% higher in proposed solution. Also, the FDR has reduced by 20% in proposed solution. Leung et al [16] used the entire whole-body image to train the deep learning network and this necessitates higher training volume to detect the tumor accurately. But the proposed solution narrowed down the area for tumor location using TAN and this reduced the area of search for tumor localization. This has increased TPR and reduced FDR in the proposed solution.

The segmentation performance of the proposed solution is compared with Leung et al [16] and the results are given in Table 1.

TABLE I: SEGMENTATION PERFORMANCE

Model	Accuracy	True-positive rate	True-negative rate
Leung et al [16]	0.72	0.50	0.81
Proposed	0.84	0.82	0.83

From the results, it can be seen that predicted segments are close to ground truth in proposed solution compared to Leung et al [16]. Use of two level segmentation with TAN at first level and refinement of TAN results with fusion at second level has increased the segmentation performance of the proposed solution.

V. CONCLUSION AND FUTURE SCOPE

Detecting tumor at earlier stages can reduce the mortality rates. Detecting tumor from PET/CT whole body scans is the best solution for earlier detection. But segmenting tumor from whole body scans is challenging due to noises of similar tissues surrounding tumor regions. This work proposed an integration of TAN with GAN for segmentation of tumor legions from PET/CT images. The proposed solution increased the true positives rate and reduced false discovery rate in tumor localization compared to existing work. The segmentation accuracy is also higher in the proposed solution.

Examining the performance of proposed tumor detection techniques against more noisy images is in scope of future work.

REFERENCES

- [1] <https://www.who.int/news-room/fact-sheets/detail/lung-cancer..>
- [2] Jemaa S, Fredrickson J, Carano RAD, Nielsen T, de Crespigny A, Bengtsson T. Tumor Segmentation and Feature Extraction from Whole-Body FDG-PET/CT Using Cascaded 2D and 3D Convolutional Neural Networks. J Digit Imaging. 2020 Aug;33(4):888-894.

- [3] He J, Zhang Y, Chung M, Wang M, Wang K, Ma Y, Ding X, Li Q, Pu Y. Whole-body tumor segmentation from PET/CT images using a two-stage cascaded neural network with camouflaged object detection mechanisms. *Med Phys*. 2023 Oct;50(10):6151-6162. doi: 10.1002/mp.16438. Epub 2023 May 3. PMID: 37134002.
- [4] Meng L, Tian Y, Bu S. Liver tumor segmentation based on 3D convolutional neural network with dual scale. *J Appl Clin Med Phys*. 2020 Jan;21(1):144-157. doi: 10.1002/acm2.12784. Epub 2019 Dec 2. PMID: 31793212; PMCID: PMC6964770.
- [5] Liu T, Liu J, Ma Y, He J, Han J, Ding X, Chen CT. Spatial feature fusion convolutional network for liver and liver tumor segmentation from CT images. *Med Phys*. 2021 Jan;48(1):264-272. doi: 10.1002/mp.14585. Epub 2020 Nov 27. PMID: 33159809.
- [6] Alirri OI. Deep learning and level set approach for liver and tumor segmentation from CT scans. *J Appl Clin Med Phys*. 2020 Oct;21(10):200-209. doi: 10.1002/acm2.13003. Epub 2020 Aug 10. PMID: 33113290; PMCID: PMC7592966.
- [7] Shiri I, Amini M, Yousefirizi F, Vafaei Sadr A, Hajianfar G, Salimi Y, Mansouri Z, Jenabi E, Maghsudi M, Mainta I, Becker M, Rahmim A, Zaidi H. Information fusion for fully automated segmentation of head and neck tumors from PET and CT images. *Med Phys*. 2024 Jan;51(1):319-333. doi: 10.1002/mp.16615. Epub 2023 Jul 20. PMID: 37475591.
- [8] Li GY, Chen J, Jang SI, Gong K, Li Q. SwinCross: Cross-modal Swin transformer for head-and-neck tumor segmentation in PET/CT images. *Med Phys*. 2024 Mar;51(3):2096-2107. doi: 10.1002/mp.16703. Epub 2023 Sep 30. PMID: 37776263; PMCID: PMC10939987.
- [9] Kevin H. Leung. 2023. Automated Whole-Body Tumor Segmentation and Prognosis of Cancer on PET/CT. In *Proceedings of the SC '23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis (SC-W '23)*. Association for Computing Machinery, New York, NY, USA, 124–133.
- [10] Li, Laquan & Zhao, Xiangming & Lu, Wei & Tan, Shan. (2019). Deep Learning for Variational Multimodality Tumor Segmentation in PET/CT. *Neurocomputing*. 392. 10.1016/j.neucom.2018.10.099.
- [11] Ludovic Sibille, Xinrui Zhan, Lei Xiang, "Whole-body tumor segmentation of 18F -FDG PET/CT using a cascaded and ensembled convolutional neural networks", *Computer Vision and Pattern Recognition (cs.CV)*;2022.
- [12] Yuan, Yading. (2021). Automatic Head and Neck Tumor Segmentation in PET/CT with Scale Attention Network. 10.1007/978-3-030-67194-5_5.
- [13] Yafei Dong, "Head and Neck Tumor Segmentation from F-FDG PET/CT Images Based on 3D Diffusion Model", arXiv:2401.17593v1 [eess.IV] 31 Jan 2024.
- [14] Bourigault, Emmanuelle & McGowan, Daniel & Mehranian, Abolfazl & Papież, Bartłomiej. (2022). Multimodal PET/CT Tumour Segmentation and Prediction of Progression-Free Survival Using a Full-Scale UNet with Attention. 10.1007/978-3-030-98253-9_18.
- [15] Naser MA, van Dijk LV, He R, Wahid KA, Fuller CD. Tumor Segmentation in Patients with Head and Neck Cancers Using Deep Learning Based-on Multi-modality PET/CT Images. *Head Neck Tumor Segm* (2020). 2021;12603:85-98. doi: 10.1007/978-3-030-67194-5_10. Epub 2021 Jan 13. PMID: 33724743; PMCID: PMC7929493.
- [16] Leung KH, Rowe SP, Sadaghiani MS, Leal JP, Mena E, Choyke PL, Du Y, Pomper MG. Deep Semisupervised Transfer Learning for Fully Automated Whole-Body Tumor Quantification and Prognosis of Cancer on PET/CT. *J Nucl Med*. 2024.
- [17] Froelich JW, Salavati A. Artificial Intelligence in PET/CT Is about to Make Whole-Body Tumor Burden Measurements a Clinical Reality. *Radiology*. 2020 Feb;294(2):453-454. doi: 10.1148/radiol.2019192425. Epub 2019 Dec 10. PMID: 31825291.
- [18] Häggström I, Leithner D, Alvéen J, Campanella G, Abusamra M, Zhang H, Chhabra S, Beer L, Haug A, Salles G, Raderer M, Staber PB, Becker A, Hricak H, Fuchs TJ, Schöder H, Mayerhoefer ME. Deep learning for [18F]fluorodeoxyglucose-PET-CT classification in patients with lymphoma: a dual-centre retrospective analysis. *Lancet Digit Health*. 2024 Feb;6(2):e114-e125. doi: 10.1016/S2589-7500(23)00203-0. Epub 2023 Dec 21. PMID: 38135556; PMCID: PMC10972536.
- [19] Blanc-Durand P, Jégou S, Kanoun S, et al. Fully automatic segmentation of diffuse large B cell lymphoma lesions on 3D FDG-PET/CT for total metabolic tumour volume prediction using a convolutional neural network. *Eur J Nucl Med Mol Imaging* 2021; 48: 1362–70.
- [20] Yuan C, Zhang M, Huang X, et al. Diffuse large B-cell lymphoma segmentation in PET-CT images via hybrid learning for feature fusion. *Med Phys* 2021; 48: 3665–78.