Grenze International Journal of Engineering and Technology, June Issue

Survey on Deep Learning Techniques for Automated Tumor Segmentation and Classification in Whole-Body PET/CT Imaging

Shruthi N[1] and Manju N[2]
[1,2]Department of Information Science and Engineering, JSS Science and Technology University, Mysuru, India

***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. Tumor detection performance outside thorax region is indigent among the methods that have been proposed for lung tumor segmentation hence they cannot be used successfully for tumor segmentation in regions of whole body. Deep learning techniques too have issues like over fit and higher false positives when applied to whole body tumor segmentation. This paper critically evaluates the existing deep learning-based tumour segmentation techniques when applied to whole body PET/CT data in order to identify their drawbacks. Future research directions and potential benefits of integrating these techniques into clinical procedures are addressed in the paper's conclusion.**

***Keywords*— Lung cancer, Tumor segmentation, Camouflaged object detection, Deep learning.**

I. INTRODUCTION

Around the world, lung cancer is one of the leading causes of cancer-related death. An estimated 18% of cancer-related deaths are caused by lung cancer. The way to lower mortality is by early detection and treatment. A key metric for the early diagnosis of lung cancer is metabolic tumour volume (MTVwb). 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 from whole body PET/CT images and calculating the MTVwb from isolated segments through manual effort is time consuming and can be erroneous [1]. These leads to the automatic segmentation of tumor regions and calculate MTVwb. 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.Deep learning techniques are effectively used for tumor segmentation. 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 [2]. Classification in PET/CT imaging and tumor segmentation using deep learning techniques is the major highlight in this study.

Grenze ID: 01.GIJET.11.2.754

The work's objective is to highlight the issues with using deep learning methods for tumour segmentation and suggest a framework for a solution.

II. LITERATURE SURVEY

Tumour segmentation using whole-body PET/CT scans has been widely used, with deep learning approaches being used to increase accuracy. Convolutional neural networks (CNNs) in both 2D and 3D were used in early techniques to segregate tumour regions from PET/CT scans, especially those extending from the eyes to the thighs [3]. While these methods could locate tumors, high false-positive rates persisted due to challenges in distinguishing tumor-like tissues from normal surrounding tissues. To address this, Maximum Intensity Projections (MIPs) were used as inputs to deep neural networks for tumor segmentation, followed by Conditional Random Fields (CRFs) for refinement. However, this approach achieved only 70% accuracy, limited by its inability to handle tissue similarities effectively [4].

A dual-path CNN was proposed, where feature maps from two networks were fused, and CRFs were applied to reduce false positives. Although it performed well for thoracic regions, its efficacy for whole-body scans was limited by the influence of surrounding tissues [5].

Spatial CNNs introduced optimizations such as skip connections, feature fusion, and multi-scale processing to enhance segmentation effectiveness. Despite these improvements, the use of 3D CRFs for post-processing resulted in a precision of only 38% [6].

Similarly, a combination of CNNs and region-level set methods achieved coarse tumor segmentation refined using localized region-level set functions. However, this approach was restricted to liver tumors and was unsuitable for whole-body scans [7].

Hybrid architectures combining CNNs with output-level voting-based fusion showed promising results, achieving 86% accuracy for neck tumor regions.

However, the lack of consideration for tissue similarity hindered its applicability to whole-body scans [8]. Another approach extracted cross-modality features from PET and CT images at multiple resolutions, improving segmentation accuracy but failing to account for tumors obscured by similar tissues, limiting its reliability for whole-body scans [9]. Semi-supervised transfer learning methods have also been explored for tumor segmentation. Although these approaches performed well for thoracic regions, accuracy for whole-body scans remained low due to noise from similar surrounding tissues [10]. Two-stage variational deep learning models have combined 3D CNNs with PET/CT fusion techniques for tumor probability mapping and segmentation. While effective for thoracic regions, these methods were not validated for whole-body scans and struggled with tissue similarity issues [11].

Cascaded and ensemble CNNs processed PET/CT scans for tumor segmentation, achieving 68% accuracy. Multi-scale feature fusion using dynamic scale attention mechanisms achieved 73% accuracy but still suffered from high false positives [12].

Methods utilizing 3D U-Net architectures for segmentation have shown regional success but were limited by low precision in handling whole-body data [13][14].

Recent studies have proposed novel deep learning architectures such as hybrid learning feature fusion networks and transfer learning-based models using U-Net variants for segmentation. These methods improved regional segmentation performance but failed to address metabolic tissue similarities, leading to suboptimal accuracy for whole-body scans [15][16]. Additionally, ResNet-based models using PET, CT, and MIP inputs demonstrated detection capabilities but lacked fine-grained evaluations near metabolically similar tissues [17][18].

Lastly, hybrid feature fusion and higher-order radiomics techniques have been explored for refining tumor segments, but their effectiveness for regions with high tissue similarity remains untested [19][20]. Table 1 overviews the survey techniques and findings.

TABLE I: SUMMARY OF TECHNIQUES AND FINDINGS

Author	Method Proposed	Dataset used	Results	Limitations
Jemma et al. [2]	Convolutional Neural Networks to detect tumor.	Non Small Cell Lung Cancer and Non-Hodgkin's Lymphoma	3D Dice score 88.6% on an NHL 93% sensitivity on NSCLC	The false positives were higher as it could not differentiate the similarity of tissues near tumor region.

He et al. [3]	Combination of Deep learning and COD was used for tumor detection.	Non Small Cell Lung Cancer	The method's Dice, Sensitivity, and Precision scores are 0.70, 0.76, and 0.70, respectively.	The accuracy is low at 70% in this method.
Meng et al. [4]	The tumor locations are identified by Three-Dimensional Dual Path multiscale CNN (TDP-CNN). Conditional random fields are used to remove false positives and improve segmentation results.	CT images with hand labelled ground truth segmentations of the liver and liver tumours are part of the Public Liver Tumour Segmentation (LiTS) collection.	For liver tumour segmentation, the method yielded an average distance of 1.07, a Hausdorff distance of 7.69, and a Dice coefficient of 0.689.	Accurately segmenting tumors with blurred boundaries remains challenging
Liu et al. [5]	The Spatial Feature Fusion Convolutional Network (SFFCN) is a new deep learning architecture for tumour segmentation.	LiTS	A Dice Similarity Coefficient (DSC) of 68.9% was attained by Liver Tumour Segmentation.	Boundary blurring and over-segmentation are major problems and the precision is only 38% in this method.
Alirr et al. [6]	A fully convolutional neural network that uses a level set function based on regions.	LiTS and IRCAD	The liver and tumour segmentation dice scores in the IRCAD dataset are 95.2% and 76.1%, respectively, while they are 95.6% and 70% in the LiTS dataset.	Small and camouflaged tumors, especially those with low contrast compared to the surrounding tissue, may still be difficult to detect accurately.
Shiri et al. [7]	U2-Net Architecture and output level voting based fusion technique is added with Deep supervision and residual U-blocks.	Patients with head and neck cancer make up the dataset. PET/CT scans with tumour areas labelled are included in the dataset. There is no mention of the precise dataset information.	With a Dice score of 0.77 ± 0.09, PET performed reasonably well, however CT did not meet expectations, with a Dice score of just 0.38 ± 0.22.	The technique focusses on tumours of the head and neck; it has not yet been investigated if it may be applied to other tumour kinds or anatomical areas.
Li et al. [8]	For head-and-neck tumour segmentation, the Cross-Modal Swin Transformer (SwinCross) integrates cross-modal feature extraction at several resolutions using a cross-modal attention (CMA) module.	HEad and neCK TumOR (HECKTOR 2021)	The highest 5-fold average Dice score 0.732 .	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]	An approach to tumour region segmentation based on deep semi-supervised transfer learning.	Five datasets include 611 18F-fluorodeoxyglucose (18F-FDG) PET/CT scans of patients with lung, melanoma, lymphoma, head and neck, and breast cancer, and 408 prostate-specific membrane antigen (PSMA) PET/CT scans of patients with prostate cancer.	Using PET/CT scan images, the suggested method showed precise tumour segmentation and prognosis for patients with six different cancer types.	As the suggested approach did not account for homologous tissue sounds surrounding tumor locations, whole body scan accuracy would be mediocre.
Li et al. [10]	Two stage variation deep learning model to segment tumor regions.	Public and proprietary datasets containing paired PET/CT scans, with annotated tumor regions.	The average dice similarity index was 0.86 ± 0.05, the sensitivity was 0.86 ± 0.07, the positive predictive value was 0.87 ± 0.10, the volume error was 0.16 ± 0.12, and the classification error was 0.30 ± 0.12.	The issue of tissue similarity influencing tumor detection accuracy has not been taken into account by the approach.
Sibille et al. [11]	Cascaded approach for tumor segmentation using stacked ensemble of 3D CNN.	A large dataset of paired PET/CT scans with annotations for tumor regions, 18F-FDG PET/CT was used.	On 84 stratified test scenarios, a dice=0.68. There was a strong correlation between the manual and automatic Metabolic Tumour Volume (R2 = 0:969, Slope = 0:947). The average inference time was 89.7 seconds.	The accuracy is only 68% in this method. Cascaded and ensembled frameworks increase computational requirements, which may hinder real-time clinical application.

Yuan et al. [12]	Deep learning segmentation network using encoder decoder transformers. scale attention block (SA-Net).	HECKTOR 2020	Average DSC, accuracy, and recall for the model were 0.7318, 0.7851, and 0.7319, respectively.	Using this technique, the accuracy is only 73%.
Dong et al. [13]	Concatenation of PET/CT, and Gaussian noise volumes along with 3D UNet structure.	HECKTOR 2021	In comparison to the previous reference techniques, the suggested 3D diffusion model may provide H&N tumour segmentation masks with increased precision.	The method works only for certain regions in PET/CT scans.
Bourigault et al. [14]	The design of a multimodal 3D H&N tumour segmentation model with complete intra- and inter-skip connectivity.	MICCAI 2021 HECKTOR	A mean Hausdor Distance at 95% (HD95) of 3.28 and an average DSC of 0.753 were obtained using post-processing and ensembling methods.	There was no discussion of over-fitting or the extraction of pertinent features in a field that is currently undergoing research.
Naser et al. [15]	Deep learning architectures with data augmentation a 3DResidualUnet architecture that can segment oropharyngeal tumors with high performance	2021 HECKTOR	A median 95% HD of 3.143 mm and a mean DSC of 0.770 were achieved via independent external validation on the test set.	The study focused on head and neck cancers, and the model's performance on other cancer types or anatomical regions remains unexplored.
Leung et al. [16]	Transfer learning approach along with U-net architecture for tumor segmentation.	18F-FDG and PSMA PET/CT scan	The tumour segmentation task showed median Dice similarity coefficients of 0.81, 0.76, 0.83, and 0.73 for patients with prostate cancer, lung cancer, melanoma, and lymphoma, respectively, and median true-positive rates of 0.75, 0.85, 0.87, and 0.75.	The method worked for head and neck regions, but its performance near metabolic similar tissue regions were not tested.
Froelich et al. [17]	CNN with various hyperparameter optimizations.	PET/CT scans with 18F-FDG in individuals with lymphoma (n = 327) or lung cancer (n = 302)	CNN for anatomic localisation for the organ region is 86.9%, 81.4% for the subregion, and 96.4% for the body part.	All lesions suspected of being malignant are identified and categorised, and the whole-body tumour burden may be measured with some degree of accuracy.
Häggström et al. [18]	Used maximum intensity project, 3D CT and 3D PET images as input to Resnet24 deep learning model to detect tumor.	18F-FDG- PET/CT scans	Using PET MIPs as input, a balanced accuracy of 87–91% was attained, with an AUC of 0·95 in both internal and external test datasets.	Influence of metabolic similar tissues were not tested.
Durand et al. [19]	The total metabolic tumor volume is automatically segmented by using three-dimensional convolutional neural network	FDG-PET/CT scan with diffuse large B cell lymphoma (DLBCL)	The validations set's mean DSC and Jaccard coefficients (± standard deviation) were 0.73 ± 0.20 and 0.68 ± 0.21, respectively. The validation sets of the first cohort showed an underestimation of mean TMTV of −12 mL (2.8%) ± 263 (P = 0.27).	The work did not consider the influence of surrounding similar tissues in tumor segmentation.
Yuan et al. [20]	Hybrid learning feature fusion method with feature fusion.	Diffuse large B-cell lymphoma (DLBCL)	73.03% is the mean dice coefficient and 4.39 mm is the modified Hausdorff distance	The method did not consider the influence of surrounding similar tissues in tumor segmentation.

III. CHALLENGES

Despite significant advancements, several challenges continue to persist, such as In order to compute other metrics, such as MTVwb for measuring tumour levels, an automatic tumour segmentation technique examines PET/CT images and separates tumour sections. Numerous approaches that have been put out for automatic

tumour segmentation were covered in the survey. 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

- There is a challenge to identify the tumor margins on CT images since extra-thoracic tumors have tissue densities similar to the surrounding soft tissue which can camouflage them.Tumors with low contrast or overlapping metabolic activities are still difficult to detect.
- On PET images, metabolically active (FDG-avid) tissues and healthy organs may resemble tumours. The brain, spleen, liver, heart, bone, marrow, lymph nodes, and brown fat are a few examples.

Most recent techniques discussed in the survey did not consider above issues due to which, their accuracy, precision is low.
Also false positives were higher. There is no discussions on over-fitting and extraction of prominent features from images,similar tissues influencing the accuracy of tumor detection is not considered,
Cross modality approach did not consider tumors hidden by similar tissues leads to less accuracy, over segmentation and boundary blurring is also major issue also there is a difficulty in detecting small tumors with low contrast.

IV. METHODOLOGY

Deep learning models may perform semi-automatic and automated segmentation of PET/CT images. However, a number of parameters, such as texture, shape, noise, motion artifacts, and the presence of tumors in patients, have prevented accurate segmentation. It is very crucial to develop a pricier tumor segmentation model which minimizes false positive annotations in PET/CT images especially for scans consisting wide spread lesion metastases like lung cancer.
Traditional deep learning methods require precisely labeled data, and it can be difficult to collect and appropriately annotate data for medical tasks. Self-supervised learning (SSL) is a learning paradigm that has been developed to overcome this problem. To best of my knowledge, using deep learning methods in combination with camouflaged object detection (COD) has not been explored and this approach can increase the accuracy of tumor segmentation.
Thus a prospective solution architecture to address the open issues in detection/segmentation of tumor regions from whole body PET/CT scan is using deep learning methods in combination with COD and classifiers. Methodology for Tumor Segmentation and Classification is as follows.

A. Preprocessing Techniques

- Artifact Removal: Techniques like GANs and denoising autoencoders have been employed to mitigate noise and artifacts.
- Normalization and Standardization: Intensity normalization helps address inter-scanner variability.
- Data Augmentation: Strategies such as flipping, rotation, and intensity perturbation are used to increase dataset diversity.

B. Segmentation Methods

- Convolutional Neural Networks (CNNs): U-Net and its variants dominate segmentation tasks, leveraging skip connections for precise localization.
- 3D Models: 3D U-Net and V-Net process volumetric data directly, preserving spatial context.
- Attention Mechanisms: Attention U-Net and Transformer-based models improve focus on tumor regions.
- Multi-Modal Integration: Methods combining PET and CT data enhance segmentation accuracy by exploiting complementary information.

C. Classification Techniques

Classification involves predicting tumor type, stage, or prognosis.

- Hybrid Architectures: To capture spatial and temporal dependencies Combination of CNNs with recurrent neural networks (RNNs) are used.
- Transfer Learning: For PET/CT applications Pre-trained models on large datasets (e.g., ImageNet) are fine-tuned.
- Self-Supervised Learning: Leveraging unlabeled data for pretraining to overcome limited labeled datasets.
- Ensemble Methods: To improve classification performance an Ensemble methods combine multiple individual models (e.g., Gradient Boosting, Random Forest, etc).

D. Evaluation Metrics

- Segmentation: Dice coefficient, Jaccard index, and Hausdorff distance.
- Classification: Accuracy, precision, recall, F1-score, and area under the ROC curve (AUC).
- Clinical Relevance: Metrics such as MTV and total lesion glycolysis (TLG) are increasingly being used.

V. CONCLUSION

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. Due to this accuracy is reduced and false positives are increased. This survey made a critical analysis of solution in two categories of deep learning and camouflaged object detection to detect tumor from whole body scans.

A comprehensive analysis to identify the prevailing challenges and gaps in the domain under investigation is conducted by the survey, to address the identified issues effectively a prospective solution architecture was conceptualized and designed for mitigating the challenges while aligning with the domain's requirements and constraints a systematic approach is outlined by the proposed architecture.

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] Alirr 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, Alvén 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.