LUMF: A Deep Learning Based Accurate and Robust Liver Uptake Measurement Framework for SUV Evaluation in Whole-Body PET/CT Images

1st Yarong Feng#
School of Electronic and Information Engineering
Liaoning Technical University
Xingcheng City, 125105, P. R. China
18093490172@163.com

2nd Yongwei Tang#
School of Electronic and Information Engineering
Liaoning Technical University
Xingcheng City, 125105, P. R. China
2458672708@qq.com

3rd Yuxin Liang
School of Electronic and Information Engineering
Liaoning Technical University
Xingcheng City, 125105, P. R. China
2833391387@qq.com

4th Chao Lv*
Department of General Surgery, Shengjing Hospital
China Medical University
Shenyang city, 110159, P.R.China
clu@cmu.edu.cn

5th Tianyu Shi*
School of information Science and Engineering
Shenyang Ligong University
Shenyang city, 110159, P.R.China
ty.shi@sylu.edu.cn

6th Meng Wang*
School of Electronic and Information Engineering
Liaoning Technical University
Xingcheng City, 125105, P. R. China
wangmeng_dx@lntu.edu.cn

***Abstract*—Accurate localization and segmentation of the liver standard uptake value (SUV) reference region are prerequisites for calculating the liver SUV reference value, and also constitute a critical step in the quantitative analysis of PET images, directly affecting the accuracy and reproducibility of tumor grading and assessment of metabolic activity in tissues and organs. Traditional manual methods and existing semi-automatic methods suffer from low efficiency and high subjectivity. Current fully automatic methods, such as the maximum hyperellipsoid-based liver reference region segmentation (MHE-LRS) method, fit a maximum inscribed hyperellipsoid into the liver to segment the reference region. While efficient and applicable to most standard PET images, these methods lack robustness in complex scenarios. To address the above challenges, this paper proposes, for the first time, a deep learning framework named LUMF for achieving robust liver SUV reference region segmentation and quantitative assessment. This method innovatively combines a deep learning model with a machine learning model. First, a pre-trained TotalSegmentator is used to segment the liver mask from whole-body CT images, based on which paired liver PET/CT images are obtained. These paired images are then used as input to nnU-Net for supervised training. Finally, on the test set, the framework outputs probability predictions and localization information of the liver SUV reference region, which serve as a region-of-interest mask. The MHE-LRS method is subsequently applied to achieve robust segmentation of the liver SUV reference region.**

Keywords—Liver SUV reference region, PET/CT, MHE-LRS, nnU-Net

I. INTRODUCTION

Positron Emission Tomography (PET) and Computed Tomography (CT) are two commonly used methods for acquiring human body images. PET can display tissue metabolic activity, while CT provides precise anatomical localization and morphological information. The combination of both techniques enables simultaneous acquisition of functional and structural data [1]. In PET image analysis, a stable and reliable fully automated SUV reference value calculation method is crucial for accurately assessing metabolic activity in whole-body tumor lesions and monitoring treatment responses [2-3]. SUV serves as a semi-quantitative indicator of normalized radioactive concentration in PET images, facilitating comparisons across different tissues, individuals, and scanning conditions [4]. Traditional liver region localization methods rely on manual delineation by physicians or technicians. This time-consuming and labor-intensive process has become a major bottleneck in large-scale radiomics research and clinical workflows. Furthermore, operator subjectivity introduces significant inter- and intra-observer variability [5-6].

To overcome the limitations of manual delineation, a series of semi-automated or automated methods have been proposed. Early approaches primarily relied on simple image processing techniques, such as automatically identifying the largest spherical Volume of Interest (VOI) within the liver region or employing segmentation algorithms based on fixed thresholds [7-8]. While these methods reduced human bias to some extent, their performance heavily depended on preset parameters and struggled to adapt to complex real-world scenarios. For instance, Bauer et al. [9] proposed a fully automated segmentation method for liver reference regions based on PET images. After setting the threshold to 1.0 within the liver ROI, a morphological closing operation is applied to the binary image to fill internal holes and form connected regions. A three-dimensional distance transform map is then computed to locate the internal point farthest from the liver surface as the center. Finally, using this

# These authors contributed equally
*Corresponding author.

center as a reference, the largest inscribed hyperellipsoid is determined along the X, Y, and Z axes, fitted to the liver, and segmented to yield the SUV reference region. However, this method relies solely on PET images, offering acceptable accuracy but insufficient robustness.

In recent years, breakthroughs in deep learning, particularly convolutional neural networks (CNNs), have opened new avenues for addressing this challenge in medical image segmentation. U-Net and its variants have demonstrated outstanding performance in multi-organ segmentation tasks using CT and MRI data [10,13]. However, directly applying deep learning to PET images for liver reference region segmentation faces core challenges: First, initialization localization is difficult. PET images suffer from low spatial resolution and poor soft tissue contrast, making it challenging for networks to stably anchor the liver within the complex torso background and prone to interference from tissues with high uptake [14]. Second, high-quality annotated data is scarce. Acquiring large-scale, expert-annotated liver SUV reference region PET datasets is challenging, limiting the training and generalization of supervised learning models [15]. Consequently, no studies have yet developed deep learning methods for liver SUV reference region localization and segmentation. This further contributes to the persistent lack of robust, universal solutions for calculating liver SUV reference values.

To address this, this paper proposes a deep learning framework—Liver Uptake Measurement Framework (LUMF)—for locating and segmenting liver SUV reference regions in whole-body PET/CT images, systematically resolving the aforementioned challenges for the first time. We first use Totalsegmentator [16] to segment the liver mask from whole-body CT images. Based on the obtained mask, we further derive corresponding liver PET/CT image pairs. These pairs are then fed into nnUNet [17] for training, ultimately generating probability predictions for liver SUV reference regions. Probability predictions with thresholds θ=0.7 are utilized as localization information within 3D Slicer [18]. Finally, the localization mask is integrated into the Maximum Hyperellipsoid-based Liver Reference Region Segmentation method (MHE-LRS) to achieve precise segmentation and quantitative calculation of liver SUV reference regions. The core contributions of this study are as follows:

- To our knowledge, this marks the first systematic introduction of deep learning into the localization and segmentation of liver SUV reference regions. Rather than simply applying existing segmentation networks, we specifically designed and optimized training strategies tailored to PET image characteristics and clinical requirements.
- By integrating TotalSegmentator with nnUNet, we resolved the challenge of initial localization of liver reference regions within PET images.
- Combining deep learning and machine learning methods, we achieved a highly robust framework capable of accurately segmenting SUV reference regions and calculating SUV reference values.

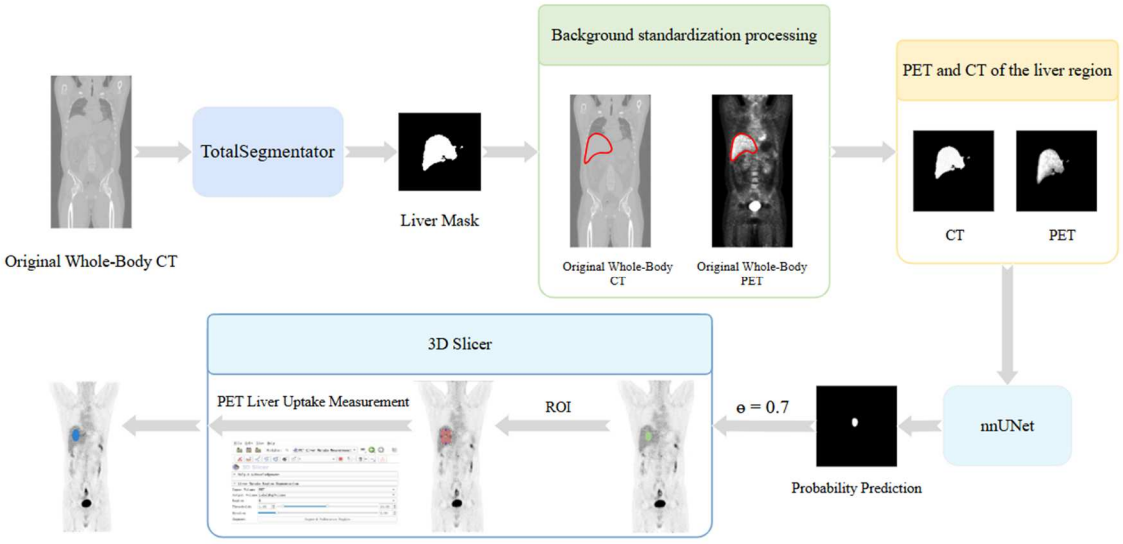

Fig. 1. LUMF structural flowchart, where TotalSegmentor employs the model proposed by Wasserthal et al. [14]. In 3D Slicer, the probability prediction is first displayed as a localization mask over the liver region (green). The ROI function within the Markups module is then used to select the localization mask (red). Subsequently, the SUV reference region (blue) is segmented using the PET Liver Uptake Measurement plugin, which integrates the MHE-LRS method.

II. Related Works

A. Localization of the Liver SUV Reference Region

The liver, serving as the endogenous reference organ for FDG-PET, has seen its SUV reference region delineation methods evolve from entirely manual to automated approaches. While international guidelines clearly establish the principle of placing the VOI in the right hepatic lobe to avoid major vessels and lesions, specific implementation still relies on operator experience, leading to significant variability [2,17]. To reduce this variability, a series of semi-automated or automated methods have been proposed. For instance, automatically

placing a fixed-volume sphere within a predefined liver region partially mitigates operator variability in shape delineation, yet its performance critically depends on the accuracy of the initial liver region [7,20]. Methods based on fixed or adaptive thresholds are simple to implement but highly sensitive to global or local image noise and uneven tracer distribution, with threshold selection lacking universality [21-23]. In recent years, specialized plugins integrated into general medical image analysis platforms represent progress in semi-automated methods. For instance, the PET Liver Uptake Measurement [24] plugin in 3D Slicer employs the MHE-LRS method for liver reference region segmentation. While stable in routine cases with standard anatomical structures, this approach demonstrates insufficient robustness when encountering complex cases involving anatomical deformation, pathological uptake interference, or poor image quality. This often necessitates manual intervention or renders the method entirely ineffective in clinical practice, failing to fundamentally resolve the tension between automation and reliability [25].

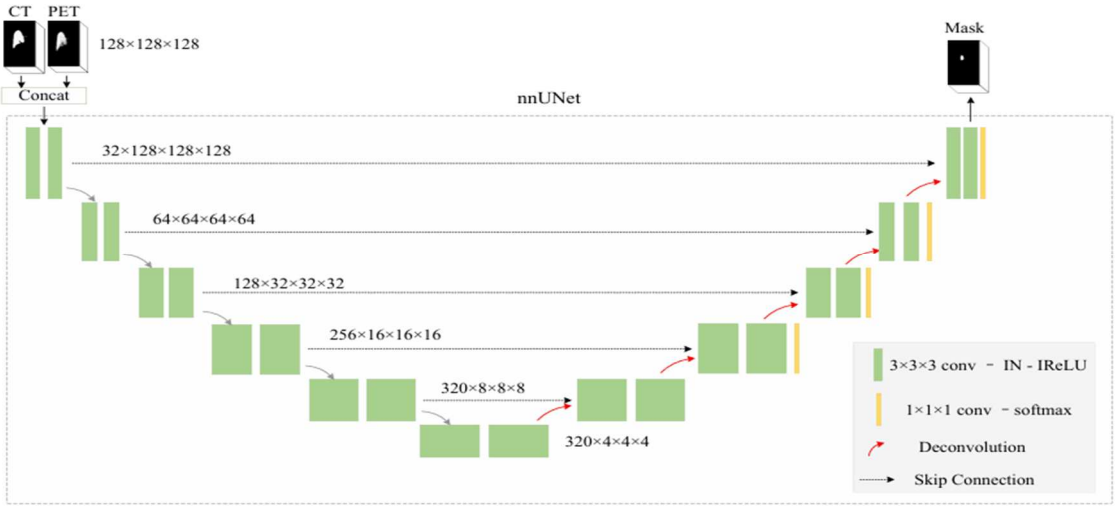

Fig. 2. nnU-Net training architecture diagram

B. *Medical Image Segmentation*

The introduction of U-Net established the central role of encoder-decoder architectures in medical image segmentation [10-12], with subsequent research primarily focused on architectural refinements [26-27]. For instance, Oktay et al. [28] introduced the Attention U-Net model, incorporating attention modules into the classic U-Net architecture. This enables the network to adaptively focus on target organ regions, enhancing segmentation accuracy. Chen et al. [29] proposed a medical image segmentation framework combining Transformers with U-Net, where Transformers encode global context, while U-Net provides local detail information. Multi-level features are fused through skip connections, demonstrating superior performance across multiple medical datasets. However, the emergence of nnU-Net indicates that holistic optimization of data processing, training strategies, and inference workflows is equally crucial compared to designing a single network architecture [17]. It employs a systematic automated pipeline to adaptively determine network depth, data preprocessing, and training strategies based on target dataset characteristics. Achieving leading performance across multiple international segmentation challenges, it demonstrates exceptional generalization capability and reliability, establishing itself as a powerful benchmark framework for medical image segmentation. Nevertheless, no studies have directly transferred these segmentation models to PET/CT images, particularly for liver SUV region of interest segmentation.

III. METHODOLOGY

A. *Overall Network Architecture*

This study proposes a deep learning framework named LUMF (Liver Uptake Measurement Framework) for robust localization and segmentation of liver SUV reference regions in whole-body PET/CT images. As shown in Figure 1, the framework consists of three consecutive stages, integrating deep learning models with machine learning models. First, a pre-trained TotalSegmentator automatically segments the liver mask from whole-body CT images. Based on this mask, liver-specific PET/CT image pairs are obtained. Subsequently, liver region images are input into the nnU-Net model for supervised training to predict the probability distribution of SUV reference regions within the liver. A high-confidence localization mask is generated using a thresholding strategy (θ=0.7). Finally, the localization mask is integrated into the Maximum Hyperellipsoid-based Liver Reference Region Segmentation (MHE-LRS) method to achieve precise segmentation and quantitative calculation of the liver SUV reference region.

B. *Liver Region Extraction Based on TotalSegmentator*

1) Data Preparation: All PET and CT data undergo rigorous spatial registration. Specifically, we utilize the ANTs (Advanced Normalization Tools) registration framework within the 3D Slicer platform to align PET images to their corresponding CT space, achieving spatial consistency between the two modalities at the voxel level. The aligned PET images are represented in CT reference space as body-weight

This work was supported by the National Natural Science Foundation of China (Grant No. 82572327), Science and Technology Plan Joint Program of Liaoning Province of China (Natural Science Foundation - General Program, 2024-MSLH-596)、Science and Technology Program of Liaoning Province of China (Provincial Doctoral Research Startup Fund Program, 2024- BS-258) and National Natural Science Foundation of China (Grant No. 62506243).

normalized standardized uptake value (SUVbw), ensuring consistency in both spatial location and metabolic information scale relative to anatomical structures.

Liver Region Extraction and Background Normalization: Based on the registered PET/CT data, this stage achieves precise extraction of the liver region. TotalSegmentator, a general anatomical segmentation model trained on large-scale CT datasets, demonstrates outstanding organ recognition capabilities. By inputting whole-body CT images into the TotalSegmentator model, a binary segmentation mask of the liver is obtained. Based on this mask, background normalization is performed on the original CT and PET images. Specifically, within the masked region, the original CT values (in HU units) and PET values (in SUV units) are fully preserved, ensuring all quantitative information within the liver is retained. Non-liver regions are set to standard background values (CT: -1024 HU; PET: 0 SUV). This strategy maintains the integrity of internal liver tissue structure and metabolic information while effectively eliminating interference signals from non-target areas, providing deep learning models with pure and standardized input data.

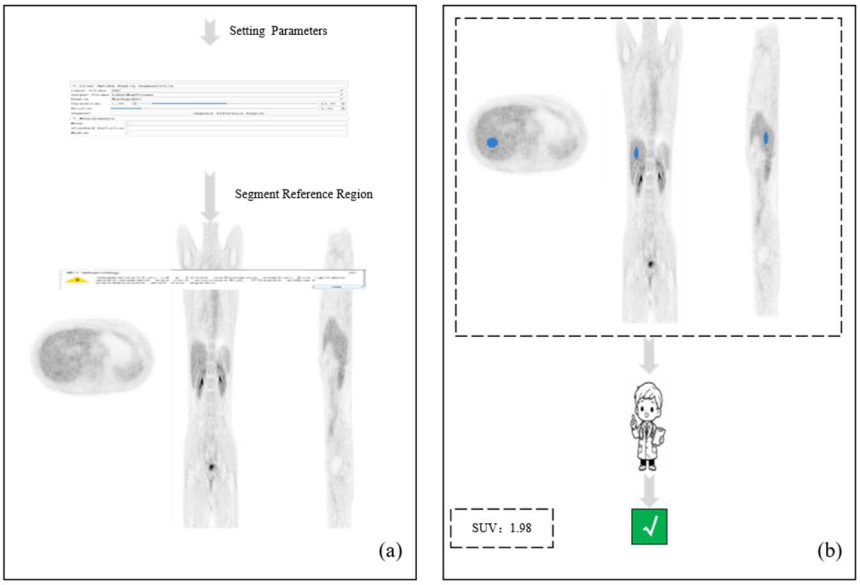

Fig. 3. Qualitative comparison. (a) The PET Liver Uptake Measurement plugin failed to successfully localise and segment the liver SUV reference region. (b) Based on a high-confidence localization mask, the PET Liver Uptake Measurement plugin successfully localized and segmented the liver SUV reference region.

C. *Deep Learning-Driven High-Confidence Localization Module for Liver SUV Reference Regions*

1) Probabilistic SUV Region of Interest Prediction Based on nnU-Net: After obtaining standardized liver PET/CT images, we trained the nnU-Net v2 framework using its 3d_fullres configuration to accurately identify SUV regions of interest within the liver. During training, as shown in Figure 2, liver PET and CT images served as bimodal inputs. These were fed into the standard 3D nnU-Net architecture in a multi-channel format and trained according to the nnU-Net adaptive configuration. After training completion, the model demonstrating optimal performance on the validation set was selected for test set inference. For each test case, a 3D probability map matching the input dimensions was generated, where each voxel value represented the predicted confidence level of belonging to the SUV reference region. This probability map provided reliable information for subsequent generation of high-confidence localization masks.

2) High-Confidence Localization Mask Generation: To extract reliable localization information of the liver SUV reference region from the three-dimensional probability map output by nnU-Net, this study performs thresholding on the probability map to filter out low-confidence voxels and generate a binary localization mask for reference region. The selection of the threshold was based on the analysis of multiple preliminary experimental results shown in Table II. A confidence threshold of $\theta = 0.7$ was ultimately determined to comprehensively balancing localization success rate and regional integrity. Specifically, thresholding is applied to each voxel in the probability map, and only voxels with confidence above and equal to this threshold are retained, thereby generating a binary high-confidence localization mask. After this processing, the mask can accurately lock the spatial location of the liver SUV reference region, with clear boundaries and complete regional coverage, providing a high-confidence initialization for subsequent automated segmentation and quantitative analysis based on MHE-LRS.

D. *Automated Reference Region Segmentation and Quantitative Analysis Guided by Positioning Masks*

To achieve precise segmentation of the liver SUV reference region and subsequent quantitative analysis, this study exported the high-confidence localization mask in the medical image standard format (NIfTI) [30]. It ensured strict alignment with

the spatially registered PET/CT images within the same coordinate system, thereby avoiding localization errors introduced by spatial shifts. Subsequently, based on the spatial distribution of the positioning mask, the outer spatial boundaries were manually defined using the ROI function within the Markups module of 3D Slicer to construct a three-dimensional region of interest [31]. This region was constructed with reference to the overall mask boundary, incorporating appropriate spatial redundancy to ensure complete coverage of the SUV reference region.

After constructing the region of interest, the MHE-LRS method is employed to accurately localize and segment the liver reference region on PET images. MHE-LRS utilizes the internal spatial morphological constraints of the liver to define the reference region by fitting a maximum inscribed hyperellipsoid, and performs an optimized search guided by the high-confidence localization mask. Specifically, the high-confidence localization mask serves as a hard constraint during the MHE-LRS stage to restrict the search space of the hyperellipsoid, thereby effectively preventing the search process from deviating from the valid liver region. This method can stably identify the liver reference region in complex cases, significantly improving the robustness and accuracy of segmentation while ensuring the interpretability of the results. Compared with traditional methods that rely on manual initialization, this study constructs the region of interest guided by the high-confidence localization mask, effectively avoiding the problem of unstable initialization under conditions of metabolic abnormalities, anatomical variations, or noise interference, and significantly enhancing the stability and robustness of the final segmentation results.

After obtaining the final segmented SUV reference region, the corresponding segmentation mask is exported. Combined with the registered PET image, statistical analysis is performed on the SUVbw values of all voxels within the region to calculate the mean and variance of SUV values. These metrics are utilized for subsequent Z-score calculations and quantitative outcome analysis.

IV. Experiments

A. Dataset

The PET/CT dataset employed in this study originates from Shengjing Hospital of China in Shenyang, encompassing clinically authentic cases to evaluate the performance of LUMF in liver SUV reference region localization and segmentation. The dataset is stored in the Digital Imaging and Communications in Medicine (DICOM) format (.dcm) [32]. Each case comprehensively includes PET modality images, CT modality images, and corresponding clinical diagnostic reports, forming a tripartite data system of 'PET-CT-diagnostic report'. This provides the experiment with comprehensive anatomical structure information, metabolic information, and imaging findings.

B. Experimental Settings

To evaluate the performance of LUMF in liver SUV reference region segmentation, this study adopted the 3d_fullres configuration of the nnU-Net v2 framework. The training and validation set comprised 48 PET/CT cases, divided into an 8:2 ratio to fully evaluate the model's generalization capability across varying data distributions. An additional 18 independent cases formed the test set to validate performance on unseen data. Both model training and inference were executed on an NVIDIA RTX 4090 GPU to ensure computational efficiency and training stability. Model performance was primarily quantified using the Dice coefficient. Furthermore, to ensure experimental reliability, the training process for each fold of cross-validation strictly adhered to the default training strategy of nnU-Net, encompassing data augmentation, learning rate adjustment, and loss function design.

TABLE I. Statistics of SUV values and Z-scores for reference liver regions in the test set based on the LUMF method

Patient_id	SUV_{test}	Z-score	Patient_id	SUV_{test}	Z-score
1	2.00	-0.19	10	2.42	0.81
2	2.37	0.69	11	1.71	-0.88
3	2.24	0.38	12	1.64	-1.05
4	2.50	1.00	13	1.74	-0.81
5	1.51	-1.36	14	1.99	-0.21
6	1.98	-0.24	15	2.03	-0.12
7	2.24	0.38	16	1.99	-0.21
8	2.91	1.98	17	1.98	-0.24
9	2.21	0.31	18	2.27	0.45

C. Comparative Experiments

1) Quantitative Results Analysis: As shown in Table 1, we quantitatively evaluated the MHE-LRS method guided by high-confidence localization masks generated by the LUMF framework using statistical metrics to validate the rationality and reliability of liver metabolic levels. Utilizing a human statistics model constructed from prior work covering 300 PET/CT datasets, the mean liver SUV was determined as μ=2.08 with an overall standard deviation of σ=0.42. For the test set comprising 18 cases, the mean SUV of the liver reference region segmented under high-confidence mask guidance (SUVtest) was employed to calculate the Z-score (1):

$$Z = \frac{SUV_{test} - \mu}{\sigma} \quad (1)$$

Statistical results indicate that $|Z| \leq 2$ [33] for all 18 test cases, demonstrating that the segmented liver metabolic levels fall within the normal human distribution range with no significant

abnormal deviations. This outcome validates the metabolic rationality of MHE-LRS segmentation guided by the LUMF framework and confirms its ability to consistently produce quantitative results conforming to human statistical patterns in complex cases.

2) Qualitative Results Analysis: This study conducts a qualitative analysis on complex cases in the test set that exhibit metabolic abnormalities, anatomical variations, or image noise interference. It compares the performance differences between the traditional PET Liver Uptake Measurement workflow and the LUMF framework-guided approach in liver SUV reference region segmentation, and validates the practicality of the proposed method with clinician feedback. The qualitative results show that, under the aforementioned complex conditions, direct application of the traditional workflow often fails to generate an effective SUV reference region mask due to unstable initialization of the reference region. In the test set, when the high-confidence localization mask is not introduced (i.e., without threshold constraints), the original MHE-LRS-based workflow is unable to obtain a usable reference region in these complex cases. Even after multiple manual adjustments of key parameters such as thresholds and morphological erosion, issues such as target region localization offset or segmentation process interruption persist. In contrast, the high-confidence localization mask generated by the LUMF framework reliably constrains the spatial extent of the SUV reference region within the liver. The resulting localization is reasonably positioned, with clear boundaries and complete regional coverage, providing a stable initialization condition for the subsequent MHE-LRS segmentation. Based on this initialization, the MHE-LRS method integrated into the PET Liver Uptake Measurement plugin successfully completes automated segmentation of the reference region in complex cases. As shown in Figure 3, the final SUV reference region is primarily located in metabolically homogeneous areas within the liver, showing high consistency with the anatomical structure and metabolic distribution characteristics of the liver, with no observable under-segmentation or over-segmentation.

Further invitations were extended to radiologists to review segmentation results generated by the LUMF-guided method, conducting a comprehensive assessment based on regional localization accuracy, boundary integrity, and clinical applicability. Evaluation findings demonstrated that all participating clinicians endorsed the segmentation outcomes produced by the LUMF-guided approach. They affirmed that this framework effectively resolves segmentation failures encountered in complex cases within traditional workflows, exhibiting sound clinical usability and practical value.

3) Threshold Sensitivity Analysis: To validate the impact of threshold selection during the generation of the high-confidence localization mask on the final results, this study performed a systematic threshold sensitivity analysis on the liver SUV reference region probability maps output by nnU-Net. Since the probability map reflects the model's confidence that each voxel belongs to the target region, the threshold setting directly determines the spatial extent of the final binary mask and consequently affects the completeness and accuracy of the localized region.

Specifically, we varied the threshold θ from 0.5 to 0.9 and evaluated the results on the test set from the following three aspects: (1) Segmentation success rate, defined as the proportion of cases in which MHE-LRS successfully generates a valid SUV reference region. (2) Region integrity, which assesses whether the generated SUV reference region is spatially connected and free of internal holes. (3) Clinical plausibility, evaluated by radiologists based on three criteria: spatial localization accuracy, appropriateness of region size, and boundary delineation quality, using a 5-point scoring system. As shown in Table II, the choice of threshold θ has a significant impact on the generation of high-confidence SUV reference regions by MHE-LRS. When $\theta = 0.5$, although the segmentation success rate reaches 100%, the inclusion of a large number of low-confidence voxels leads to irregular and overextended boundaries, resulting in poor region integrity. When $\theta = 0.6$, noise is effectively suppressed and the regional structure becomes more stable, with the clinical score increasing to 4.1. At $\theta = 0.7$, the best balance is achieved, producing a continuous region without internal holes and obtaining the highest clinical score (4.8). When $\theta \geq 0.8$, some cases exhibit overly small, fragmented, or even missing masks, leading to a decrease in segmentation success rate and a significant drop in clinical scores. Therefore, $\theta = 0.7$ is selected as the optimal threshold.

TABLE II: Performance comparison of different confidence thresholds (θ) on the test set.

Threshold (θ)	Segmentation success rate (%)	Region integrity	Clinical plausibility score
0.5	100.0	Poor	3.5
0.6	100.0	Moderate	4.1
0.7	100.0	Good	4.8
0.8	88.9	Poor	3.4
0.9	77.8	Very poor	2.8

D. Ablation Study

To validate the efficacy of key design elements within LUMF, ablation experiments were conducted to compare model performance between a whole-body PET/CT model (Whole-body nnU-Net) and a liver-region-specific PET/CT model (Liver nnU-Net). As shown in Table III, the Liver nnU-Net achieved an average Dice coefficient of 0.500, surpassing the Whole-body nnU-Net's 0.418, while exhibiting greater stability across folds. This indicates that training solely on liver-region data eliminates interference from extraneous anatomical structures in whole-body images, enabling the model to focus more intently on liver SUV features. Consequently, segmentation accuracy and robustness are enhanced, validating the rationale and efficacy of employing liver PET/CT as input for LUMF under localization guidance.

TABLE III. Comparison of whole-body nnU-Net and liver nnU-Net performance.

Model \ Fold	1	2	3	4	5	Mean
Whole-body nnU-Net	**0.526**	0.416	0.287	0.461	**0.536**	0.418
Liver nnU-Net	0.488	**0.500**	**0.440**	**0.550**	0.521	**0.500**

V. Discussion

This study addresses the challenge of locating and segmenting the liver SUV reference region within whole-body PET images in complex clinical cases. It proposes and validates the efficacy and robustness of the LUMF deep learning framework. Experimental results demonstrate that LUMF effectively compensates for the shortcomings of traditional liver reference region segmentation methods based on maximum inscribed hyperellipsoid fitting during the automated initialization stage by providing high-confidence spatial localization information. Within the test dataset, conventional methods demonstrated suboptimal performance in liver SUV reference region localization and segmentation. Key contributing factors include: firstly, abnormal liver morphology in certain patients—such as irregular volumes, localized lesions, or residual hepatic tissue post-surgical resection—rendering the method's reliance on standard anatomical priors ineffective. Secondly, metabolic expression in PET images may be affected by localized hypermetabolic foci, noise, or fatty infiltration, leading to uneven SUV distribution. Furthermore, abnormalities in whole-body metabolic distribution may further compromise the stability of reference region localization. The combined effect of these factors renders traditional methods unreliable for segmentation under conditions of metabolic abnormality or anatomical variation. In contrast, LUMF employs deep learning to extract local metabolic features while incorporating spatial constraints, enabling stable target region identification even in complex scenarios. Its segmentation results demonstrate favorable acceptability and interpretability in clinical evaluations.

The ablation experiments in this paper further validate the rationality of the method design. Compared to the Whole-body nnU-Net trained directly on whole-body PET/CT data, the Liver nnU-Net — trained solely on liver-region PET/CT — demonstrated superior performance in both Dice coefficient and fold-to-fold stability. This indicates that, for localized and metabolism-specific tasks such as liver SUV reference region segmentation, eliminating irrelevant anatomical interference and guiding the model to focus on the target region holds significant importance.

Although LUMF demonstrates robust performance and clinical utility in complex cases, this study retains certain limitations. Within the current workflow, the localization results generated by LUMF necessitate manual definition of regions of interest (ROIs) using the Markups module in 3D Slicer to guide subsequent segmentation processes. This has yet to achieve fully end-to-end automation, thereby impacting overall workflow efficiency to some extent. Future research will focus on developing automated ROI generation mechanisms based on LUMF outputs, coupled with cascading interfaces to downstream segmentation methods. This aims to reduce manual intervention and achieve greater process automation. Additionally, the generalisability of the method will be validated on larger datasets, while incorporating uncertainty modelling or adaptive thresholding strategies to enhance stability and reliability in extremely complex cases.

VI. Conclusion

This study proposes the LUMF framework, which integrates deep learning and machine learning techniques to achieve automated localization and segmentation of liver SUV reference regions. The specific workflow comprises: First, utilizing TotalSegmentator to automatically extract liver masks from CT images, thereby constructing standardized liver PET/CT image pairs; Subsequently, nnU-Net generates probabilistic predictions for SUV reference regions, with thresholding yielding high-confidence localization information; Finally, the localization mask is integrated into the MHE-LRS method to achieve precise segmentation and quantitative computation of the liver SUV reference region. Experimental results demonstrate that LUMF reliably generates liver SUV regions of interest in complex cases. Its segmentation outcomes align with normal human metabolic distribution, gain clinician validation, exhibit strong interpretability, and achieve superior segmentation accuracy and stability compared to models trained directly on whole-body PET/CT data. Overall, this study provides a robust technical solution for liver region of interest localization and segmentation, demonstrating clear clinical applicability.

References

[1] W. Zhao, Z. Huang, S. Tang et al., "MMCA-Net: A multimodal cross attention transformer network for nasopharyngeal carcinoma tumor segmentation based on a total-body PET/CT

[2] R. Boellaard, R. Delgado-Bolton, W. J. G. Oyen, et al., "FDG PET/CT: EANM procedure guidelines for tumour imaging: version 2.0," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 42, no. 2, pp. 328–354, Apr. 2015.

[3] R. L. Wahl, H. Jacene, Y. Kasamon, and M. A. Lodge, "From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors," *Journal of Nuclear Medicine*, vol. 50, no. Suppl 1, pp. 122S–150S, May 2009.

[4] M. C. Adams, T. G. Turkington, J. M. Wilson et al., "A systematic review of the factors affecting accuracy of SUV measurements," American Journal of Roentgenology, vol. 195, no. 2, pp. 310–320, 2010.

[5] B. Zhao, Y. Tan, W. Y. Tsai, L. H. Schwartz, and R. A. Gatenby, "Reproducibility of radiomics for deciphering tumor phenotype with imaging," *Scientific Reports*, vol. 6, no. 1, p. 23428, Mar. 2016.

[6] V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, and M. B. Schabath, "Radiomics: the process and the challenges," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1234–1248, Nov. 2012.

[7] M. Werner-Wasik, A. D. Nelson, W. Choi, Y. Xiao, and J. D. Bradley, "What is the best way to contour lung tumors on PET scans? Multiobserver validation of a gradient-based method using a NSCLC digital PET phantom," *International Journal of Radiation Oncology* Biology* Physics* vol. 82, no. 3, pp. 1164–1171, Mar. 2012.

[8] B. Foster, U. Bagci, A. Mansoor, Z. Xu, and D. J. Mollura, "A review on segmentation of positron emission tomography images," *Computers in Biology and Medicine*, vol. 50, pp. 76–96, Jan. 2014.

[9] C. Bauer, S. Sun, W. Sun et al., "Automated measurement of uptake in cerebellum, liver, and aortic arch in full-body FDG PET/CT scans," Medical Physics, vol. 39, no. 6, pp. 3112–3123, 2012.

[10] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015, pp. 234–241.

[11] M. Wang, H. Jiang, T. Shi, and Y. -d. Yao, "HD-RDS-UNet: Leveraging Spatial-Temporal Correlation Between the Decoder Feature Maps for

Lymphoma Segmentation," IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 3, pp. 1116–1127, 2022.

[12] J. Ni, J. Shi, Q. Zhan, Z. Zhang, and Y. Gu, "An improved U-Net model with multiscale fusion for retinal vessel segmentation," Intelligent Robotics, vol. 5, no. 3, pp. 679–694, 2025.

[13] G. Litjens, T. Kooi, B. E. Bejnordi, et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017.

[14] Z. Guo, X. Li, H. Huang, N. Guo, and Q. Li, "Deep learning-based image segmentation on multimodal medical imaging," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 162–169, Mar. 2019.

[15] B. Zhou, H. Xie, Q. Liu, S. Li, and Z. Liang, "FedFTN: personalized federated learning with deep feature transformation network for multi-institutional low-count PET denoising," *Medical Image Analysis*, vol. 90, p. 102993, Aug. 2023.

[16] J. Wasserthal, H. C. Breit, M. T. Meyer, et al., "TotalSegmentator: robust segmentation of 104 anatomic structures in CT images," *Radiology: Artificial Intelligence*, vol. 5, no. 5, p. e230024, Sep. 2023.

[17] F. Isensee, J. Petersen, A. Klein, and K. H. Maier-Hein, "nnu-net: self-adapting framework for u-net-based medical image segmentation," 2018, arXiv:1809.10486. [Online]. Available: https://arxiv.org/abs/1809.10486

[18] S. Pieper, M. Halle, and R. Kikinis, "3D Slicer," in Proceedings of the 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro, Arlington, VA, USA, 2004, pp. 632–635.

[19] G. Krokos, T. Kotwal, A. Malaih, et al., "Evaluation of manual and automated approaches for segmentation and extraction of quantitative indices from [18F] FDG PET-CT images," *Biomedical Physics & Engineering Express*, vol. 10, no. 2, p. 025007, Jan. 2024.

[20] M. C. Adams, T. G. Turkington, J. M. Wilson, and T. Z. Wong, "A systematic review of the factors affecting accuracy of SUV measurements," *American Journal of Roentgenology*, vol. 195, no. 2, pp. 310–320, Aug. 2010.

[21] P. Cheebsumon, F. H. P. van Velden, M. Yaqub, et al., "Effects of image characteristics on performance of tumor delineation methods: a test–retest assessment," *Journal of Nuclear Medicine*, vol. 52, no. 10, pp. 1550–1558, Oct. 2011.

[22] M. Tamal, "Intensity threshold based solid tumour segmentation method for positron emission tomography (PET) images: a review," *Heliyon*, vol. 6, no. 10, Oct. 2020.

[23] C. S. Constantino, S. Leocádio, F. P. M. Oliveira, and J. S. Cardoso, "Evaluation of semiautomatic and deep learning–based fully automatic segmentation methods on [18F] FDG PET/CT images from patients with lymphoma: influence on tumor characterization," *Journal of Digital Imaging*, vol. 36, no. 4, pp. 1864–1876, Aug. 2023.

[24] S. Chauvie, E. Bertone, F. Bergesio, A. Terulla, D. Botto, and P. Cerello, "Automatic liver detection and standardized uptake value evaluation in whole-body Positron Emission Tomography/Computed Tomography scans," *Comput. Methods Programs Biomed.*, vol. 156, pp. 47–52, 2018. doi: 10.1016/j.cmpb.2017.12.026.

[25] J. Driessen, G. J. C. Zwezerijnen, H. Schöder, et al., "The impact of semiautomatic segmentation methods on metabolic tumor volume, intensity, and dissemination radiomics in 18F-FDG PET scans of patients with classical Hodgkin lymphoma," *Journal of Nuclear Medicine*, vol. 63, no. 9, pp. 1424–1430, Sep. 2022.

[26] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: a nested u-net architecture for medical image segmentation," in Proceedings of the International Workshop on Deep Learning in Medical Image Analysis, Granada, Spain, 2018, pp. 3–11.

[27] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, and Y. Iwamoto, "Unet 3+: a full-scale connected unet for medical image segmentation," in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 1055–1059.

[28] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, and K. Misawa, "Attention u-net: learning where to look for the pancreas," 2018, arXiv:1804.03999. [Online]. Available: https://arxiv.org/abs/1804.03999

[29] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, and Y. Wang, "Transunet: transformers make strong encoders for medical image segmentation," 2021, arXiv:2102.04306. [Online]. Available: https://arxiv.org/abs/2102.04306

[30] R. W. Cox, J. Ashburner, H. Breman, K. Fissell, C. Haselgrove, C. J. Holmes, S. C. Lancaster, D. Rex, S. M. Smith, J. B. Woodward, and S. C. Strother, "A (sort of) new image data format standard: NiFTI-1," in *Proc. 10th Annu. Meet. Organ. Hum. Brain Mapp.*, Budapest, Hungary, Jun. 2004, p. 01.

[31] V. Kumar, Y. Gu, S. Basu, A. Berglund, S. A. Eschrich, M. B. Schabath, K. Forster, H. J. Aerts, A. Dekker, D. Fenstermacher, D. B. Goldgof, L. O. Hall, P. Lambin, Y. Balagurunathan, R. A. Gatenby, and R. J. Gillies, "Radiomics: the process and the challenges," *Magn. Reson. Imag.*, vol. 30, no. 9, pp. 1234–1248, 2012. doi: 10.1016/j.mri.2012.06.010.

[32] M. Larobina and L. Murino, "Medical image file formats," *J. Digit. Imag.*, vol. 27, no. 2, pp. 200–206, 2014. doi: 10.1007/s10278-013-9657-9.

[33] D. Hellwig, T. P. Graeter, D. Ukena, A. Groeschel, G. W. Sybrecht, H. J. Schaefers, and C. M. Kirsch, "18F-FDG PET for mediastinal staging of lung cancer: which SUV threshold makes sense?," J. Nucl. Med., vol. 48, no. 11, pp. 1761–1766, 2007. doi: 10.2967/jnumed.107.044362.