Original Article

# Supervised Volumetric Segmentation of White and Gray Matter from Brain Positron Emission Tomography Images Using Magnetic Resonance Labels

**Kiarash Danesh[1], Alireza Karimian[2], Peyman Sharifian[2], Hossein Arabi[3]**

*[1]Department of Medical Radiation Engineering, Faculty of Physics, University of Isfahan, Isfahan, Iran, [2]Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran, [3]Department of Medical Imaging, Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, Geneva, Switzerland*

## Abstract

**Background:** This study evaluates the performance of five deep convolutional neural networks (DCNNs) for supervised segmentation of white matter (WM) and gray matter (GM) in brain positron emission tomography (PET) images using label maps derived from corresponding magnetic resonance (MR) images. The goal is to reduce dependency on hybrid PET/Magnetic resonance imaging (MRI) by extracting anatomical information solely from PET images, thereby providing a potential pathway for MRI-free partial volume correction (PVC). **Methods:** A total of 300 PET images and their corresponding MR-derived labels from the OASIS-3 dataset were used. The five DCNNs, including the UNet, RegUNet, VNet, SegResNet, and HighResNet, were implemented within the Medical Open Network for Artificial Intelligence (MONAI) framework. The networks were evaluated using Dice score and Intersection over Union (IoU) metrics. **Results:** Among the networks, VNet demonstrated superior performance for GM and WM segmentation, achieving Dice scores of 61.18% and 76.23%, and IoU scores of 44.15% and 61.62%, respectively. UNet and VNet showed competitive performance for WM segmentation, with no statistically significant differences between them. **Conclusions:** These findings provide insights into the performance of DCNNs for PET image segmentation, highlighting VNet's capability and emphasizing the potential for optimizing segmentation techniques for PVC applications.

**Keywords:** *Brain segmentation, convolutional neural network, magnetic resonance imaging, positron emission tomography*

**Submitted:** 21-May-2025 **Revised:** 27-Sep-2025 **Accepted:** 22-Oct-2025 **Published:** 12-May-2026

## Introduction

In the past few years, there has been a significant and rapid expansion of artificial intelligence (AI) in various fields such as science, medicine, and industry. AI is now being applied in a wide range of applications, including object classification, language processing, speech recognition, image classification, and many other tasks that just a few years ago were considered beyond human capability.[1-3] In the last 50 years, the primary purpose of AI in the medical field has been focused on solving logical and reasoning problems.[4] However, the recent progress in machine learning and deep learning has led to novel advancements that enable the performance of specific processing tasks. This marks a significant shift in capabilities for image segmentation and interpretation.[5] These extensive processing tasks include, but are not limited to, detection, localization, radiomics feature extraction, and segmentation.[6-9] In medicine, the purpose of utilizing AI and deep learning techniques is to detect, classify, and automatically segment malignant tissue in images obtained from various imaging modalities.[10-12] Segmentation, which encompasses the detection, feature extraction, and classification of images, plays a crucial role in the interpretation of medical images. Accurate interpretation of medical images assists physicians in effectively determining and defining the boundaries of malignant tissue, thereby aiding them in treatment planning.[13-15]

The purpose of volumetric segmentation of brain images is to identify and separate white matter (WM), gray matter (GM), and cerebrospinal fluid from the background regions.[16] The objective of GM and WM segmentation from a positron emission



**How to cite this article:** Danesh K, Karimian A, Sharifian P, Arabi H. Supervised volumetric segmentation of white and gray matter from brain positron emission tomography images using Magnetic Resonance labels. J Med Signals Sens 2026;16:12.

***Address for correspondence:***
*Prof. Alireza Karimian, Department of Biomedical Engineering, Faculty of Engineering, University of Isfahan, Isfahan, Iran.*
*E-mail: karimian@eng.ui.ac.ir, karimianalireza2@gmail.com*

**Access this article online**

**Website:** www.jmssjournal.net

**DOI:** 10.4103/jmss.jmss_49_25

**Quick Response Code:**

tomography (PET) image is illustrated in Figure 1, which shows a sample axial slice of a brain PET image alongside its corresponding magnetic resonance (MR)-derived segmentation mask. Volumetric segmentation of brain images produces detailed three-dimensional (3D) volumes that are invaluable for physicians making critical decisions in clinical settings. These volumes assist in identifying tumors, planning radiotherapy treatments, and predicting the progression of tumors.[17,18] Moreover, considering partial volume effects and high noise characteristics of PET images, two major challenges facing quantitative PET imaging, excellent PET imaging for symptomatic purposes requires a standard dose of radioactive tracer that is infused into the patient's body, increasing the risk of radiation harm. To mitigate such risks, reducing the tracer dose has become a common procedure in radiation imaging, which produces lower-quality PET images.[19-21] These structured volumes can improve the algorithms used in the partial volume correction (PVC) of PET images.[22-24] Nevertheless, manual segmentation of the brain is a costly and time-consuming process that demands advanced skills. This creates a strong need for automated segmentation methods that are both accurate and efficient. Advanced deep convolutional neural networks (DCNNs) have become essential for this purpose, offering precise segmentation by extracting key features from medical images. These DCNNs provide the accuracy needed for reliable and rapid automatic segmentation, meeting the growing demand in the medical field.[25-28]

Segmentation with the help of discriminative/semantic features and deep learning methods can help in the diagnosis of neurodegenerative brain diseases such as Alzheimer's, dementia, and others, since these diseases affect the density of GM in the brain.[29]

## Related works

Various deep learning-based segmentation studies have been carried out across different medical imaging modalities. A selection of significant studies in this area is presented below.

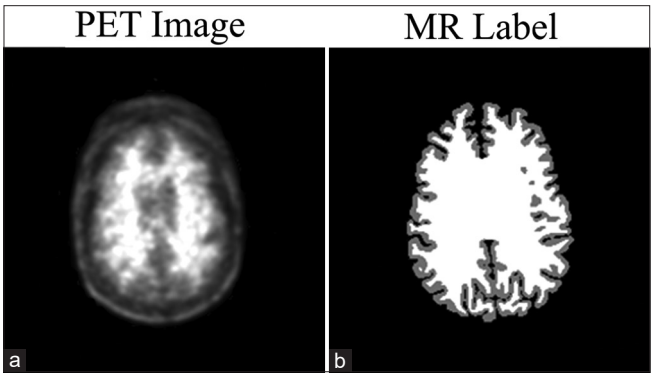


**Figure 1: Illustration of the segmentation task. (a) An axial slice of a florbetapir brain positron emission tomography image. (b) The corresponding magnetic resonance-derived label map showing gray matter (dark gray) and white matter (white). PET – Positron emission tomography; MR – Magnetic resonance**

Bai *et al*. introduced a new approach to segmenting PET-computed tomography (CT) images[30] using a two-stage framework. In the first stage, a guided filter was used to merge PET and CT images, creating a fusion map that highlighted and outlined tumor tissues. The second stage built on this by developing an innovative level set model that incorporated both region and edge information from the fusion map. By combining the strengths of the guided filter and the level set model, the method improved the accuracy of PET-CT image segmentation. Unlike traditional frameworks or deep learning methods, this approach delivered superior results without requiring a large training dataset.

Azimi *et al*. presented an attention-based convolutional neural network (ATB-Net) in[31] designed to perform PVC on brain 18F-fluorodeoxyglucose (FDG) PET images, eliminating the need for anatomical information. The ATB-Net model showed superior performance compared to the standard U-Net, particularly in the iterative Yang PVC method, a region-based approach. The findings highlighted the potential of the attention-based deep learning approach to effectively perform PVC on PET images without the need for magnetic resonance imaging (MRI) data, making it a valuable tool in situations where MRI is unavailable.

Dirks *et al*. outlined a fully automated system in[32] for detecting and segmenting lesions in whole-body 18F-FDG PET/CT images. This method mirrors the manual process used by nuclear medicine specialists and involves two main steps. First, the system segmented all regions with elevated PET tracer uptake using a threshold based on an automatically identified volume of interest (VOI) in the liver. Then, a convolutional neural network (CNN) classified these areas as either physiological uptake or tumor tissue. To enhance interpretability, the system provides intermediate outputs, including the selected liver VOI and regions identified as physiological uptake. When tested on a dataset of 69 patients with malignant melanoma, the system delivered competitive results, surpassing a direct segmentation approach in terms of detection and segmentation accuracy.

Huang *et al*. performed an automated method[33] for segmenting lymphomas in 3D PET and CT images, combining Dempster–Shafer theory with deep learning. The approach included two main components: a deep feature extraction module and an evidential layer. The feature extraction module used an encoder-decoder framework to extract semantic features from 3D images, while the evidential layer utilized these features to calculate a belief function at each voxel, indicating the uncertainty of lymphoma presence. This method, which merged deep feature extraction with evidential segmentation, outperforms the baseline UNet model and three other leading models in tests on a dataset of 173 patients.

In this article, we explore the volumetric segmentation of PET images using label maps derived from MR images using different CNNs. The accuracy of the segmented brain tissue from PET images was evaluated with commonly used segmentation metrics. The findings of this research could significantly reduce the computational demands of medical image segmentation. It also has the potential to eliminate the need for MRI in hybrid PET/MR techniques by extracting analogous anatomical information directly from the PET data itself. This extracted structural information can then be utilized for tasks such as PVC, which traditionally relies on MR-derived tissue maps. In addition, direct PET segmentation can facilitate proper analysis of neurodegenerative diseases and contribute to the development of the PVC algorithms currently in use.[34]

## Methods

### Overall pipeline

A schematic overview of the proposed methodology is presented in Figure 2. The process involves data preprocessing, model training, and evaluation.

### Data preparation and preprocessing

Images from the Open Access Series of Imaging Studies (OASIS-3) database[35] were utilized to create a local dataset consisting of 300 PET images along with their corresponding MR images and labels for the different brain regions. Among the various types of PET images, those acquired through the injection of florbetapir (18F-AV-45) radiopharmaceutical exhibited a favorable signal-to-noise

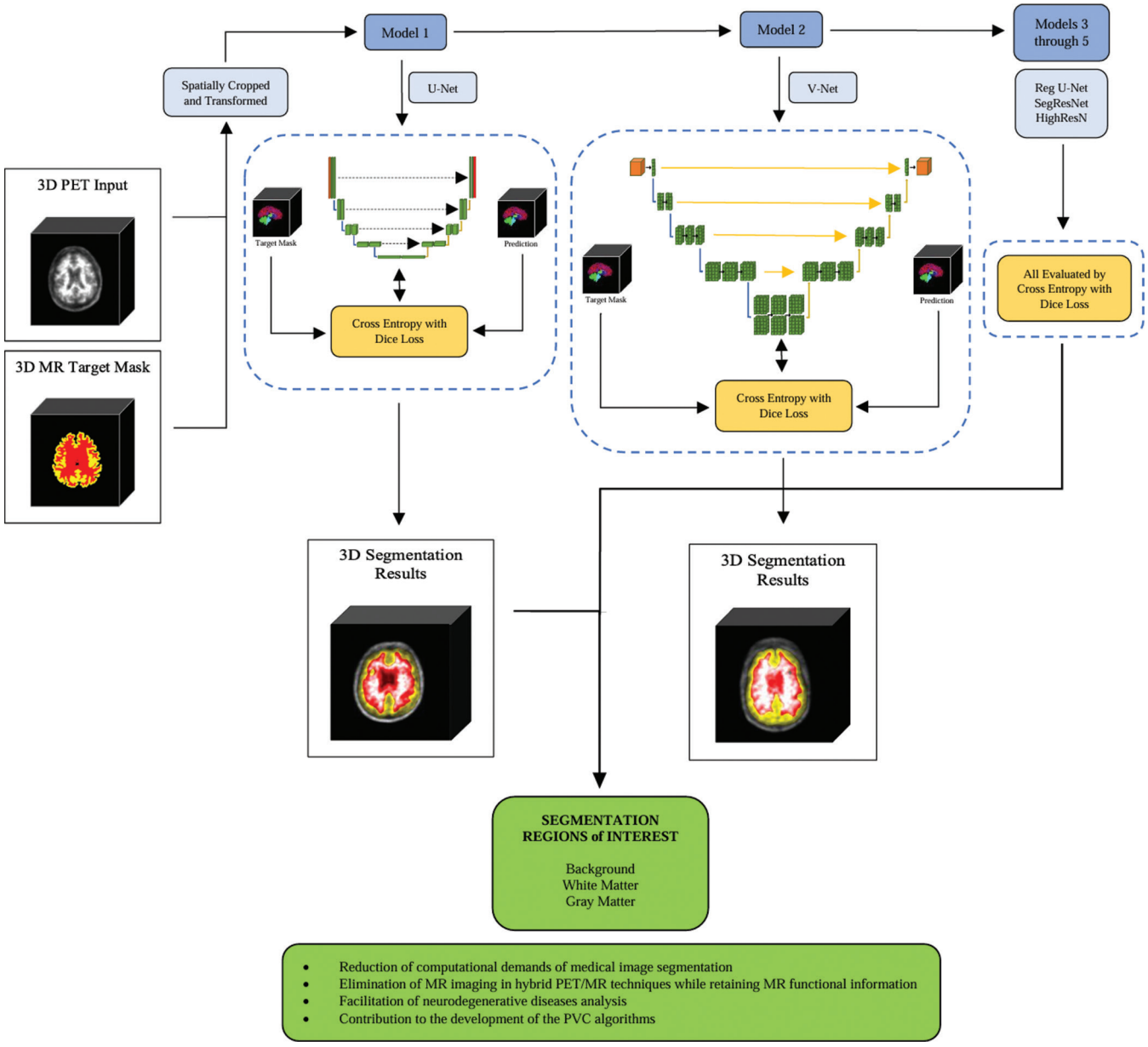


**Figure 2: A block diagram illustrating the main steps of the study pipeline, including data preparation, preprocessing, model training, and evaluation. PET – Positron emission tomography; MR – Magnetic resonance; PVC – Partial volume correction; 3D – Three-dimensional**

ratio, making them a suitable choice for the implementation of the anatomical region segmentation from PET images. The PET images obtained from this dataset had sizes of 256 × 256 × 127 × 26. The last dimension, with a size of 26, represented an additional dynamic time dimension with 26 time frames, wherein, by calculating the weights of each time frame and averaging across all weights, the time dimension was eliminated (a single static image was created from the dynamic PET images with superior signal-to-noise ratio).

Each PET image was paired with two corresponding MR T1-weighted images and their respective T1-weighted label maps. Given that PET and MR images are acquired using different imaging devices, it was necessary to register each corresponding image to ensure spatial alignment in the Talairach coordinates. The process was automatically conducted using the BRAINS general registration algorithm in 3D Slicer software, resulting in image sizes of 256 × 256 × 256 with voxel sizes of one millimeter. The quality of registration was visually verified to avoid any significant errors. Finally, the voxel values of each PET image were normalized between zero and one by identifying the highest voxel value within the image.

Next, the label maps were equally aligned to the PET images, and the number of anatomical regions was reduced to three classes (background and cerebrospinal fluid, GM, and WM). The first class (regions with voxel values set to zero) represented the background (BG) regions, including cerebrospinal fluid, while the second class (regions with voxel values set to one) represented the GM, and the last class (regions with voxel values set to two) represented the WM.

### Deep convolutional neural networks

Five neural networks were trained using the MONAI framework to perform tissue segmentation on PET images based on their corresponding MR-derived label maps. These networks include UNet, RegUNet, and VNet from the UNet family, as well as SegResNet and HighResNet from the ResNet family. The UNet network was prepared following the original work by Ronneberger *et al*.[36] Similarly, the RegUNet, VNet, SegResNet, and HighResNet networks were prepared based on the original works by authors,[37-40] respectively.

The 3D UNet architecture, a volumetric extension of the original two-dimensional (2D) model, consists of a contracting path to capture contextual information and an expansive path for precise localization in 3D semantic segmentation. RegUNet, on the other hand, is a region-aware UNet. RegUNet improves semantic segmentation by identifying the region of interest (RoI) and then extracting the features of the RoIs that will be used in segmentation. Another model of the UNet family, VNet is also designed to segment volumetric medical images like MR and PET. VNet architecture also consists of a contracting path and an expansive path. The difference is that the kernels used are volumetric, and there are different stages in each path, making the model learn residual functions in each stage. Since segmentation involves extracting data that is significantly larger than the usual size of the anatomical structure to be delineated, this configuration would make an important attribute to the model, especially when poorly visible anatomy is being segmented.

SegResNet is a 3D CNN with an encoder-decoder structure, with an asymmetrically large encoder to extract deep image features, and the decoder part reconstructs dense segmentation masks. SegResNet is a CNN architecture with an encoder-decoder structure for deep feature extraction and a dense segmentation mask reconstruction in the encoder and the decoder paths, respectively. The architecture includes an additional branch in the encoder that resembles an autoencoder, serving to reconstruct the original image. Other models of the ResNet family, the HighResNet, consist of 20 convolutional layers, designed for volumetric segmentation. The architecture consists of residual functions as well as dilated convolutions. This enables a complete mapping from the image volume to a dense segmentation at the voxel level.

### Training procedures

The 300 imaging studies were randomly split at the subject level into training (200 subjects), validation (20 subjects), and testing (80 subjects) sets.

To meet the significant hardware requirements of the networks, particularly a powerful GPU and ample memory capacity, all networks were implemented in a Google Colab environment utilizing the NVIDIA A100 GPU with a 40 GB capacity and 83.5 GB RAM.

The prepared pipeline accepted 256 × 256 × 256-sized images, and each model had to segment them into three defined classes: BG with an index of 0, GM with an index of one, and WM with an index of two. The training phase for all networks consisted of 50 epochs with batch sizes of two, except for the HighResNet, which was adjusted to mitigate its demanding hardware requirements. The optimizer of choice in the main pipeline of all networks was the Adam optimizer, used with the rectified linear unit (RELU) activation function. Models were trained using a combined Dice and Cross-Entropy loss function, which is standard for multi-class segmentation tasks as it directly optimizes for the overlap between predicted and ground-truth regions while considering class probabilities. The learning rates, stride sizes, and dropout probability of the networks were adjusted during the optimization phase to achieve the best possible scores in each model. For example, the UNet and RegUNet models had a learning rate set to $1e^{-4}$, while the VNet had a learning rate of $1e^{-2}$, and the HighResNet and SegResNet had learning rates of

$1e^{-3}$. During the training phase, the average loss function, as well as the averages of the dice and Intersection over Union (IoU) metrics, were calculated based on all three defined classes. To achieve better accuracy, during the optimization phase, background regions were ignored, and the metrics were used only on the GM and WM classes.

### Evaluation metrics

Two commonly used metrics in the context of segmentation are the Dice Similarity Coefficient (DSC) and the IoU, also known as the Jaccard index. These metrics are formulated based on the true positive, false positive, true negative, and false negative indices of the confusion matrix. The DSC measures the overlap between the predicted and ground truth masks (Eq. 1), while the IoU calculates the ratio of the intersection to the union of the two masks (Eq. 2):

$$DSC = \frac{2TP}{2TP + FP + FN} \quad \text{Eq.(1)}$$

$$IoU = \frac{TP}{TP + FP + FN} \quad \text{Eq.(2)}$$

During the training and optimization phases, the average loss values were monitored to determine the optimal learning rates for each model while ensuring consistency by defining the best hyperparameters. This was done to ensure consistent outcomes between the prepared images and the models employed. Furthermore, the average loss function was tracked for all models during training to assess model fitting up to the 50$^{th}$ epoch. The results indicated satisfactory fitting conditions with no observed overfitting or underfitting across all five models.

## Results

In the testing phase, the UNet achieved a DSC score of 60.04% and an IoU score of 42.98% for segmenting the GM, as well as a dice score of 76.20% and an IoU score of 61.58% for segmenting the WM. RegUNet demonstrated a DSC score of 56.26% and an IoU score of 39.19% for the segmentation of GM. In addition, it achieved a DSC score of 72.90% and an IoU score of 57.38% for the segmentation of WM.

Among the networks from the UNet family, VNet exhibited the highest performance. It achieved a DSC score of 61.18% and an IoU score of 44.15% for the segmentation of GM. In addition, it attained a DSC score of 76.23% and an IoU score of 61.62% for the segmentation of WM.

The two other networks from the ResNet family achieved lower scores compared to the VNet and UNet. In the testing phase, SegResNet obtained a DSC score of 59.72% and an IoU score of 42.67% for the segmentation of GM. For the segmentation of WM, it achieved a DSC score of 75.39% and an IoU score of 60.54%.

HighResNet, on the other hand, achieved a DSC score of 51.44% and an IoU score of 35.08% for the segmentation of GM during the training phase. For the segmentation of WM, it achieved a DSC score of 64.93% and an IoU score of 49.43%. An overview of the overall results can be found in Table 1. Figure 3 depicts average loss values in comparison to the mean DSC and the mean IoU evaluations in the training phase in the most efficient model, the VNet.

While the accuracy of these models may not yet be sufficient for standalone clinical use, they have demonstrated promising processing times, with all networks completing data processing in under 30 s per epoch on average, resulting in an average training time of 20 min per network.

Furthermore, the statistical performance of GM and WM segmentation for all five networks can be observed in Figures 4 and 5. These figures provide a comparative analysis based on the DSC score, allowing us to evaluate the performance of each network in comparison with each others.

In addition, Figure 6 presents the qualitative segmentation results of all five networks, based on the averaged-score DSC segmentation using VNet as the reference.

Furthermore, Figure 7 showcases the best segmentation output, in comparison to the average segmentation output, and the worst segmentation output achieved by the VNet.

## Discussion

The VNet demonstrated the best overall segmentation performance, achieving a mean DSC score of 68.70%. Its architecture, which is specifically designed for volumetric data, incorporates residual connections within each stage of the

**Table 1: Results of all five networks in the testing phase based on dice similarity coefficient and intersection over union metrics and their corresponding standard deviation**

| | Testing | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | GM | | | | WM | | | |
| | DSC | SD | IoU | SD | DSC | SD | IoU | SD |
| UNet | 0.6004 | 0.0346 | 0.4298 | 0.0348 | 0.7620 | 0.0159 | 0.6158 | 0.0205 |
| RegUNet | 0.5626 | 0.0277 | 0.3919 | 0.0266 | 0.7920 | 0.0150 | 0.5738 | 0.0182 |
| VNet | 0.6118 | 0.0332 | 0.4415 | 0.0339 | 0.7623 | 0.0163 | 0.6162 | 0.0209 |
| SegResNet | 0.5972 | 0.0374 | 0.4267 | 0.0372 | 0.7539 | 0.0203 | 0.6054 | 0.0250 |
| HighResNet | 0.5144 | 0.0924 | 0.3508 | 0.0747 | 0.6493 | 0.1450 | 0.4943 | 0.1292 |

DSC – Dice similarity coefficient; IoU – Intersection over union; GM – Gray matter; WM – White matter; SD – Standard deviation

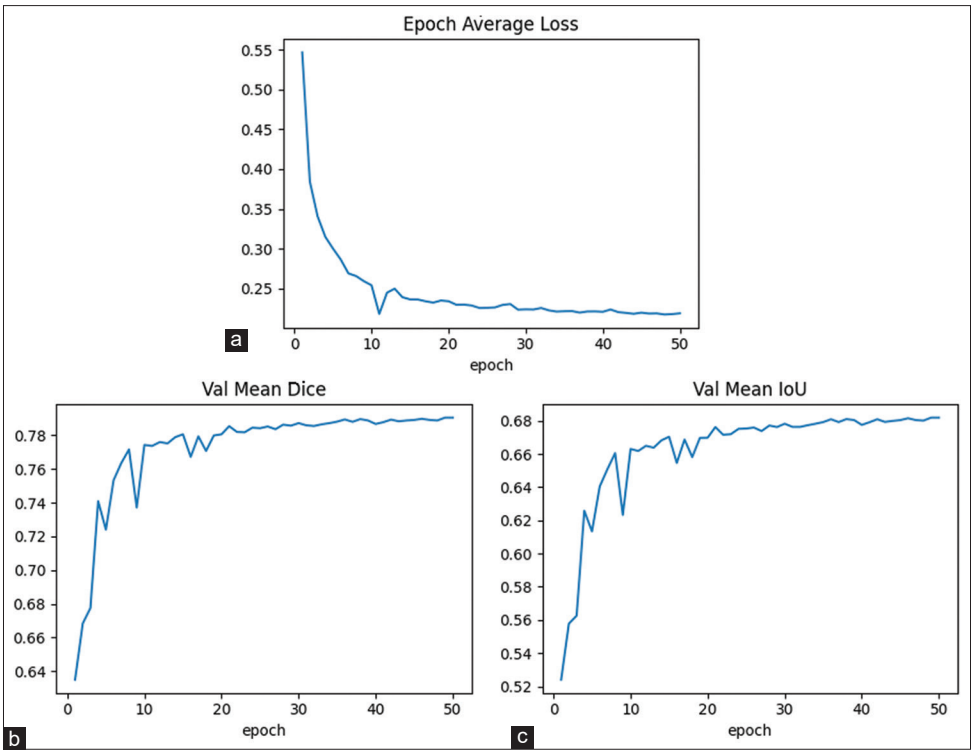


**Figure 3: Average loss values (a) in comparison with the mean dice (b) and mean intersection over union (c) metrics in the training phase of the VNet. IoU – Intersection over union**

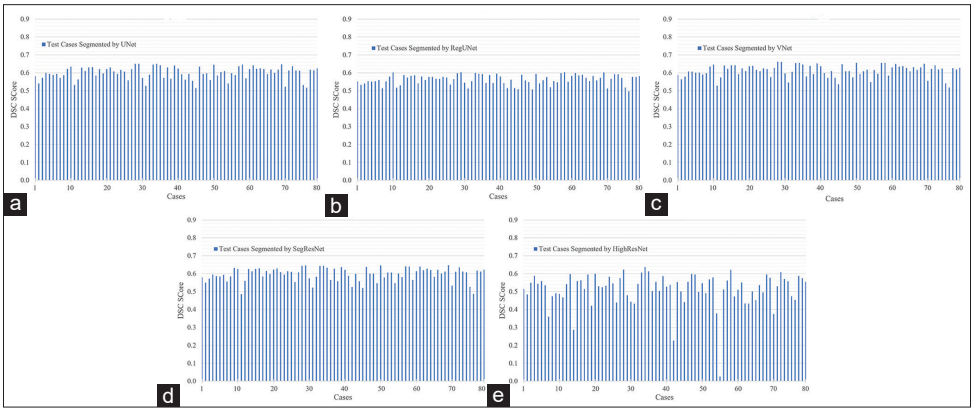


**Figure 4: The statistical performance of all five networks in terms of gray matter segmentation, as evaluated using the dice similarity coefficient metric. The individual networks are labeled as follows: (a) U-Net, (b) RegUNet, (c) V-Net, (d) SegResNet, (e) HighResNet. DSC – Dice similarity coefficient; GM – Gray matter**

encoder and decoder paths. This design likely allows it to more effectively learn complex features from the low-resolution, noisy PET data by mitigating the vanishing gradient problem and facilitating the training of deeper networks. UNet, which shares a similar encoder-decoder structure but without the same level of integrated residual learning, achieved competitive but slightly lower results. This suggests that residual learning is a beneficial characteristic for this specific task.

In contrast, RegUNet, which operates by first defining ROIs, exhibited the lowest performance among the UNet-family models. This approach may be suboptimal for segmenting brain PET images, where the global context of the entire brain volume is crucial for distinguishing GM and WM, and where the initial low resolution makes precise ROI definition challenging.

Regarding the ResNet-family models, SegResNet, which includes an autoencoder branch to reconstruct the input image, achieved higher scores than HighResNet. The reconstruction task in SegResNet may act as a regularizer, forcing the network to learn more robust feature representations that are useful for segmenting lower-resolution images. HighResNet, while employing dilated convolutions to capture multi-scale information, may be more susceptible to the high noise levels in PET data, leading to its lower performance.

In addition, based on the results obtained from Table 1, a statistical analysis was conducted by calculating the *P* values for each network relative to the most effective one, namely, the VNet. These *P* values serve as a measure of significance, enabling us to assess the statistical differences between the

performance of each network. The calculated *P* values are presented in Table 2. The analysis of these *P* values reveals that there are no significant differences between the UNet and VNet models in terms of segmenting the WM, which has a larger volume compared to the GM. This indicates that both the 3D and 2D segmentation networks from the UNet family exhibit competitive performance when it comes to segmenting larger volumes, as opposed to smaller volumes like the GM.

While MR images provide superior anatomical information and soft tissue contrast, their availability is limited in many medical imaging centers, particularly in developing countries. In addition, current PET/MRI techniques are associated with high costs for patients. Therefore, by acquiring the same anatomical information in PET images as the PET/MRI technique, these high costs can be circumvented, and the same anatomical information can be made accessible in clinical and diagnostic situations, especially in the diagnosis and analysis of neurodegenerative brain diseases. Therefore, direct volumetric segmentation of PET images is used in this work, which can improve PVC and further increase the quality of PET images, resulting in better diagnosis and treatment planning. In addition, in the absence of MRI, CT imaging is often used; however, it offers lower soft tissue contrast compared to MRI, resulting in insufficient functional information for segmentation, disease analysis, and PVC. Consequently, direct segmentation of PET images suggests the potential to exclude MR functional information while utilizing this information in clinical situations. In future studies, we plan to evaluate the segmentation response of CT images within the VNet model. In addition, we intend to incorporate MR images with the PET images as inputs to the VNet model to further enhance segmentation accuracy.

**Table 2: Calculated *P* values of different networks relative to the most effective network (the VNet) calculated *P* values in comparison with VNet**

| | GM | | WM | |
|---|---|---|---|---|
| | **DSC** | **IoU** | **DSC** | **IoU** |
| UNet | 0.0353 | 0.0332 | 0.9109 | 0.9082 |
| RegUNet | <0.0001 | <0.0001 | <0.0001 | <0.0001 |
| SegResNet | 0.0102 | 0.0095 | 0.0045 | 0.0037 |
| HighResNet | <0.0001 | <0.0001 | <0.0001 | <0.0001 |

DSC – Dice similarity coefficient; IoU – Intersection over union; GM – Gray matter; WM – White matter

These results represent a promising step toward reducing reliance on MRI, particularly for applications like initial PVC, where approximate anatomical information is sufficient and direct PET segmentation becomes feasible. It can serve as a supportive tool for PVC, guiding Standardized Uptake Value quantification, or narrowing down areas for further investigation, especially when anatomical imaging is unavailable or misaligned. Although not precise enough for standalone diagnostic decisions, it can reduce manual workload and improve workflow efficiency by providing a rough anatomical scaffold, which the short computational time required to achieve these results is in accordance with this claim. Once implemented in practice, technologists in nuclear medicine centers could benefit from rapid processing and accurate anatomical delineation. Moving forward, we aim to further improve segmentation accuracy in future work.

When contextualizing our results, direct comparison with related studies is challenging due to differences in datasets, tracers, and segmentation targets. However, our findings are in the general range of recent works on MRI-free PET analysis. For instance, Azimi *et al*.[31] proposed an attention-based network (ATB-Net) for PVC without anatomical images, reporting performance metrics that, while not directly analogous to our segmentation Dice scores, demonstrate the feasibility of extracting anatomical information from PET data. Our VNet's Dice score of ~ 61% for GM segmentation in a pure segmentation task provides a baseline for this challenging problem. The higher accuracy for WM segmentation (~76%) is consistent with the fact that WM regions are generally larger and have more uniform tracer uptake, making them

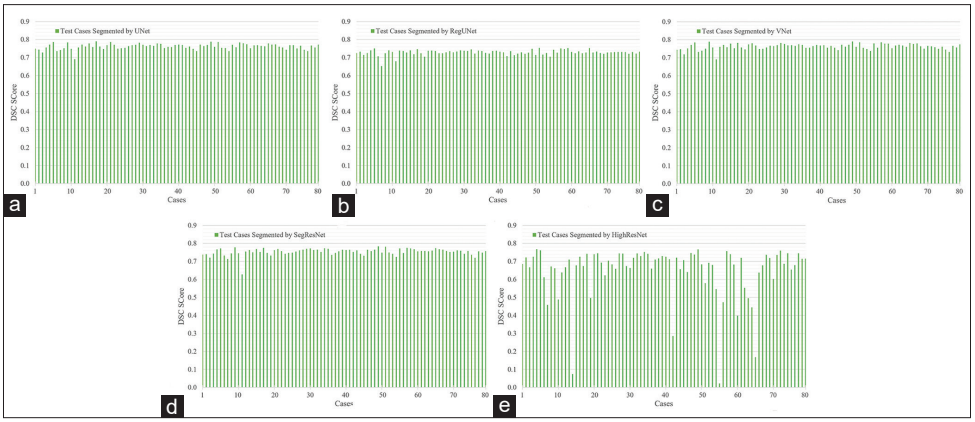


**Figure 5: The statistical performance of all five networks in terms of white matter segmentation, as evaluated using the dice similarity coefficient metric. The individual networks are labeled as follows: (a) U-Net, (b) RegUNet, (c) V-Net, (d) SegResNet, (e) HighResNet. DSC – Dice similarity coefficient; WM – White matter**

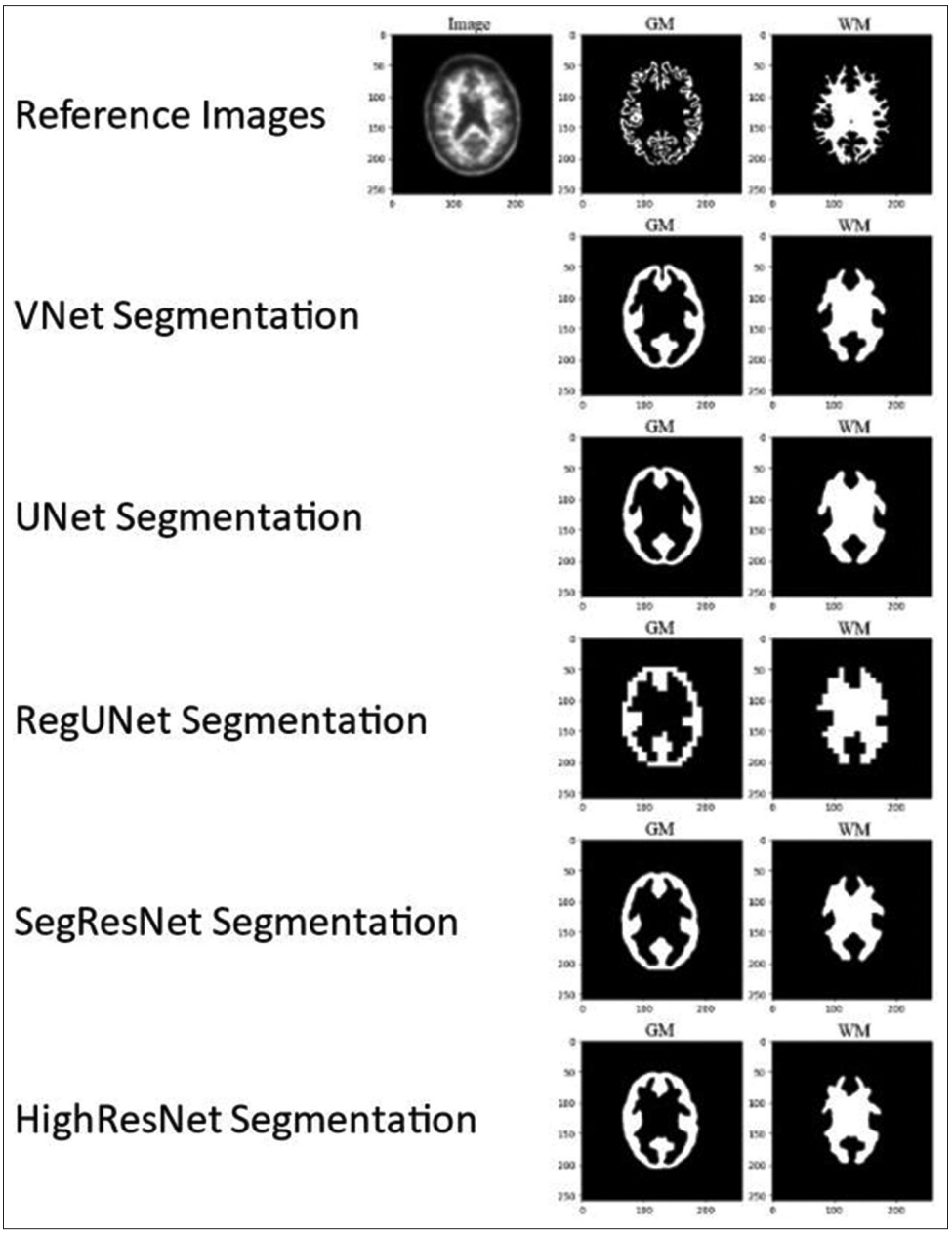


**Figure 6: The qualitative segmentation results of all five networks, based on the averaged-score dice similarity coefficient segmentation using VNet as the reference, are presented. The first row corresponds to the base positron emission tomography (PET) image, and each subsequent row displays the segmentation of the base PET image by VNet, UNet, RegUNet, SegResNet, and HighResNet, respectively. GM – Gray matter; WM – White matter**

easier to segment. In another work, Poitrasson-Rivière *et al*.[41] successfully applied a 3D UNet to segment the left ventricular myocardium in FDG PET scans for cardiac sarcoidosis, improving clinical readability in over 90% of cases. This success relied on standardized imaging protocols, including metabolic suppression to highlight inflammation and consistent LV geometry aligned with perfusion datasets. In contrast, brain FDG PET lacks physiological uptake suppression, shows high baseline cortical activity, and targets variable regions. As a result, ground truth is harder to define, often requiring manual annotation. Thus, despite using the same families of architectures, brain PET segmentation demands more advanced preprocessing and modality-specific adaptations, and achieving similar performance requires more sophisticated preprocessing, multi-modal fusion, or domain-specific adaptations. Furthermore, Kaneko *et al*.[42] presented a method for segmenting the liver directly from FDG PET images using a 3D UNet model, without relying on CT or MRI. The model achieved an average Dice score of 94% and an average IoU of 89%. In contrast, brain PET images lack physiological uptake suppression, anatomical simplicity, and consistency, which can result in weaker contrast and less reliable ground truth. Therefore, achieving accurate segmentation in brain PET is inherently more complex.

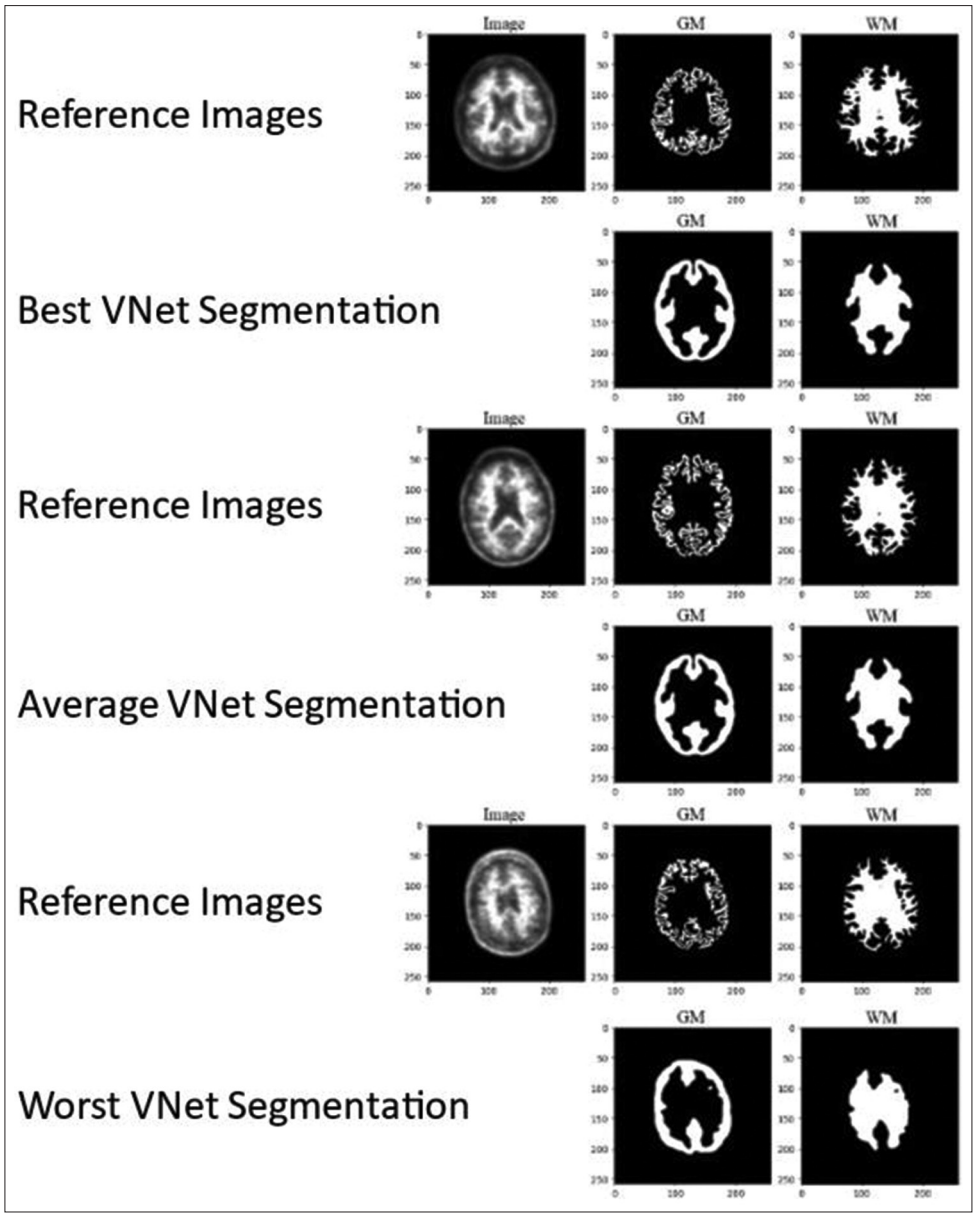


**Figure 7: The qualitative segmentation with the highest score (first two rows), with the average score (second two rows), and the worst score (last two rows) achieved by VNet**

## Limitations

This study has several limitations. First, the achieved segmentation accuracy, particularly for GM (~61% Dice), is moderate and may not be sufficient for direct clinical application without further refinement. Second, the dataset, while public and well-established, was limited to 300 subjects from a single source (OASIS-3), potentially limiting the generalizability of our models to data from different scanners or populations. Third, the use of 3D CNNs, especially with high-resolution volumes (256 × 256 × 256), is computationally expensive, requiring significant GPU memory and training time, which could be a barrier for some researchers. Finally, the segmentation labels were derived from MR images, and any registration errors between PET and MR, despite visual verification, would be propagated into the training labels, potentially limiting the upper bound of achievable performance.

## Conclusion

In this work, we conducted a comparative analysis of five 3D DCNNs for the direct segmentation of gray and WM from brain PET images. The results indicate that VNet, with its volumetric design and residual learning blocks, is the most suitable architecture for this task. The competitive performance of UNet highlights the effectiveness of the standard encoder-decoder structure. Furthermore, the superior performance of SegResNet over HighResNet suggests that auxiliary tasks like image reconstruction can enhance feature learning for challenging low-resolution

data. This study serves as a benchmark for architecture selection in MRI-free PET segmentation. Future work will focus on incorporating advanced techniques like attention mechanisms and leveraging larger, multi-center datasets to improve segmentation accuracy towards clinical utility.

### Ethical approval

The data used in this study were obtained from the publicly available Open Access Series of Imaging Studies (OASIS-3) dataset. Ethical approval and participant consent were obtained by the OASIS-3 investigators, and all procedures were conducted in accordance with relevant guidelines and regulations.

### Funding

This research did not receive any specific grant funding.

### Availability of data and materials

The data utilized in this study are publicly available from the OASIS-3 database.

### Financial support and sponsorship

Nil.

### Conflicts of interest

There are no conflicts of interest.

## References

1. Arabi H, Zaidi H. Applications of artificial intelligence and deep learning in molecular imaging and radiotherapy. Eur J Hybrid Imaging 2020;4:17.
2. Sahiner B, Pezeshk A, Hadjiiski LM, Wang X, Drukker K, Cha KH, *et al.* Deep learning in medical imaging and radiation therapy. Med Phys 2019;46:e1-36.
3. Arabi H, AkhavanAllaf A, Sanaat A, Shiri I, Zaidi H. The promise of artificial intelligence and deep learning in PET and SPECT imaging. Phys Med 2021;83:122-37.
4. Currie G, Hawk KE, Rohren E, Vial A, Klein R. Machine learning and deep learning in medical imaging: Intelligent imaging. J Med Imaging Radiat Sci 2019;50:477-87.
5. Currie G. Intelligent imaging: Anatomy of machine learning and deep learning. J Nucl Med Technol 2019;47:273-81.
6. Aghakhan Olia N, Kamali-Asl A, Hariri Tabrizi S, Geramifar P, Sheikhzadeh P, Farzanefar S, *et al.* Deep learning-based denoising of low-dose SPECT myocardial perfusion images: Quantitative assessment and clinical performance. Eur J Nucl Med Mol Imaging 2022;49:1508-22.
7. Mostafapour S, Gholamiankhah F, Dadgar H, Arabi H, Zaidi H. Feasibility of deep learning-guided attenuation and scatter correction of whole-body 68Ga-PSMA pet studies in the image domain. Clin Nucl Med 2021;46:609-15.
8. Currie G, Rohren E. Intelligent imaging in nuclear medicine: The principles of artificial intelligence, machine learning and deep learning. Semin Nucl Med 2021;51:102-11.
9. Arabi H, Zaidi H. Truncation compensation and metallic dental implant artefact reduction in PET/MRI attenuation correction using deep learning-based object completion. Phys Med Biol 2020;65:195002.
10. Karimzadeh R, Fatemizadeh E, Arabi H, editors. Attention-Based Deep Learning Segmentation: Application to Brain Tumor Delineation. 2021 28th National and 6th International Iranian Conference on Biomedical Engineering (ICBME) IEEE; 2021.
11. Moeskops P, Wolterink JM, van der Velden BH, Gilhuijs KG, Leiner T, Viergever MA, *et al.*, editors. Deep Learning for Multi-task Medical Image Segmentation in Multiple Modalities. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. 2016//; Cham: Springer International Publishing; 2016.
12. Karimzadeh R, Fatemizadeh E, Arabi H. A novel shape-based loss function for machine learning-based seminal organ segmentation in medical imaging. arXiv 2022; 2203.03336.
13. Michael E, Ma H, Li H, Kulwa F, Li J. Breast cancer segmentation methods: Current status and future potentials. Biomed Res Int 2021;2021:9962109.
14. Arabi H, Zaidi H, editors. Inclusive-Exclusive Model Training Framework to Jointly Perform Semantic Segmentation and Uncertainty Map Estimation. 2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). IEEE; 2022.
15. Gholamiankhah F, Mostafapour S, Goushbolagh NA, Shojaerazavi S, Arabi H, Zaidi H, editors. A Novel Unsupervised COVID-19 Lesion Segmentation from CT Images Based-on the Lung Tissue Detection. 2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). IEEE; 2021.
16. Bui TD, Shin J, Moon T. 3D densely convolutional networks for volumetric segmentation. arXiv 2017;1709.03199.
17. Minoshima S, Cross D. Application of artificial intelligence in brain molecular imaging. Ann Nucl Med 2022;36:103-10.
18. Dehghani F, Arabi H, Karimian A. Automated brain tumor segmentation on multi-MR sequences to determine the most efficient sequence using a deep learning method. Comput Intell Electr Eng 2023;14:31-44.
19. Azimi MS, Kamali-Asl A, Ay MR, Arabi H, Zaidi H, editors. A Novel Attention-based Convolutional Neural Network for Joint Denoising and Partial Volume Correction of Low-dose PET Images. 2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). IEEE; 2021.
20. Azimi MS, Kamali-Asl A, Ay MR, Hosseini MS, Khorshidi F, Moafpourian Y, *et al.* Evaluation of deep-learning based partial volume correction of PET images without the use of anatomical information. J Nucl Med 2023;64 Suppl 1:P971.
21. Azimi MS, Kamali-Asl A, Ay MR, Zeraatkar N, Hosseini MS, Arabi H, *et al.* Comparative assessment of attention-based deep learning and non-local mean filtering for joint noise reduction and partial volume correction in low-dose PET imaging. J Nucl Med 2022;63 Suppl 2:2729.
22. Azimi MS, Kamali-Asl A, Ay MR, Zeraatkar N, Arabi H. Deep Learning-Based Partial Volume Correction in Standard and Low-Dose PET-CT Imaging. arXiv 2022; 2207.02553.
23. Azimi MS, Kamali-Asl A, Ay MR, Zeraatkar N, Hosseini MS, Sanaat A, *et al.* Deep learning-based partial volume correction in standard and low-dose positron emission tomography-computed tomography imaging. Quant Imaging Med Surg 2024;14:2146-64.
24. Rousset O, Rahmim A, Alavi A, Zaidi H. Partial volume correction strategies in PET. PET Clin 2007;2:235-49.
25. Arabian H, Karimian A, Rasti R, Arabi H, Zaidi H, editors. Deep Attention-based Seminal Segmentation: A Practical Deep Learning Framework for Accurate Segmentation of the Hippocampus from Magnetic Resonance Images. 2021 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). IEEE; 2021.

26. Ilhan A, Sekeroglu B, Abiyev R. Brain tumor segmentation in MRI images using nonparametric localization and enhancement methods with U-net. Int J Comput Assist Radiol Surg 2022;17:589-600.
27. Sharifian P, Karimian A, Arabi H, Zaidi H, editors. Anatomical Brain Regions Segmentation From MR Images Using Global and Local Deep Learning Networks. 2023 IEEE Nuclear Science Symposium, Medical Imaging Conference and International Symposium on Room-Temperature Semiconductor Detectors (NSS MIC RTSD). IEEE; 2023.
28. Danesh K, Azimi M, Sharifian P, Karimian A, Arabi H, Zaidi H, editors. Anatomical Brain Segmentation from CT Images in Brain PET/CT Imaging for the Purpose of Partial Volume Correction in PET Imaging. 2024 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and Room Temperature Semiconductor Detector Conference (RTSD). IEEE; 2024.
29. Schwarz CG. Uses of human MR and PET imaging in research of neurodegenerative brain diseases. Neurotherapeutics 2021;18:661-72.
30. Bai S, Qiu X, Hu R, Wu Y. A novel level set model initialized with guided filter for automated PET-CT image segmentation. Cogn Robo 2022;2:193-201.
31. Azimi M, Kamali-Asl A, Ay MR, Zeraatkar N, Hosseini MS, Sanaat A, *et al.* Attention-based deep neural network for partial volume correction in brain (18) F-FDG PET imaging. Phys Med 2024;119:103315.
32. Dirks I, Keyaerts M, Neyns B, Vandemeulebroucke J. Computer-aided detection and segmentation of malignant melanoma lesions on whole-body (18) F-FDG PET/CT using an interpretable deep learning approach. Comput Methods Programs Biomed 2022;221:106902.
33. Huang L, Ruan S, Decazes P, Denœux T. Lymphoma segmentation from 3D PET-CT images using a deep evidential network. Int J Approx Reaso 2022;149:39-60.
34. Arabi H, Zaidi H. Recent advances in positron emission tomography/magnetic resonance imaging technology. Magn Reson Imaging Clin N Am 2023;31:503-15.
35. LaMontagne PJ, Benzinger TL, Morris JC, Keefe S, Hornbeck R, Xiong C, *et al*. OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease. medrxiv. 2019:2019-12.
36. Ronneberger O, Fischer P, Brox T, editors. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI. Cham: Springer International Publishing; 2015.
37. Huang X, Chen W, Liu X, Wu H, Wen Z, Shen L, editors. Left and Right Ventricular Segmentation Based on 3D Region-Aware U-Net. 2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS); 2022.
38. Milletari F, Navab N, Ahmadi SA, editors. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. Fourth International Conference on 3D Vision (3DV); 2016.
39. Myronenko A, editor. 3D MRI Brain Tumor Segmentation Using Autoencoder Regularization. Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries; Cham: Springer International Publishing; 2019.
40. Li W, Wang G, Fidon L, Ourselin S, Cardoso MJ, Vercauteren T. On the Compactness, Efficiency, and Representation of 3D Convolutional Networks: Brain Parcellation as a Pretext Task; 2017. p. 348-60.
41. Poitrasson-Rivière A, Vanderver MD, Hagio T, Arida-Moody L, Moody JB, Renaud JM, *et al*. Automated deep learning segmentation of cardiac inflammatory FDG PET. Journal of Nuclear Cardiology. 2024;42:102052.
42. Kaneko Y, Miwa K, Yamao T, Miyaji N, Nishii R, Yamazaki K, *et al.* (18) F-FDG PET-based liver segmentation using deep-learning. Phys Eng Sci Med 2025;48:1415-24.