Deep-Learning-Based Cross-Modality Striatum Segmentation for Dopamine Transporter SPECT in Parkinson's Disease

Haiyan Wang, Han Jiang, Gefei Chen, Yu Du, Zhonglin Lu, Zhanli Hu, *Senior Member, IEEE*, and Greta S. P. Mok, *Senior Member, IEEE*

***Abstract*—Striatum segmentation on dopamine transporter (DaT) SPECT is necessary to quantify striatal uptake for Parkinson's disease (PD), but is challenging due to the inferior resolution. This work proposes a cross-modality automatic striatum segmentation, estimating MR-derived striatal contours from clinical SPECT images using the deep learning (DL) methods. ^{123}I-Ioflupane DaT SPECT and T1-weighted MR images from 200 subjects with 152 PD and 48 healthy controls are analyzed from the Parkinson's progression markers initiative database. SPECT and MR images are registered, and four striatal compartment contours are manually segmented from MR images as the label. DL methods including nnU-Net, U-Net, generative adversarial networks, and SPECT thresholding-based method are implemented for comparison. SPECT and MR label pairs are split into train, validation, and test groups (136:24:40). Dice, Hausdorff distance (HD) 95%, and relative volume difference (RVD), striatal binding ratio (SBR) and asymmetry index (ASI) are analyzed. Results show that nnU-Net achieves better Dice (∼0.7), HD 95% (∼1.8), and RVD (∼0.1) as compared to other methods for all striatal compartments and whole striatum. For clinical PD evaluation, nnU-Net also yields strong SBR consistency (mean difference, −0.012) and ASI correlation (Pearson correlation coefficient, 0.81). The proposed DL-based cross-modality striatum segmentation method is feasible for clinical DaT SPECT in PD.**

***Index Terms*—Cross-modality, deep learning (DL), Parkinson's disease (PD), SPECT, striatum segmentation.**

I. Introduction

THE incidence and prevalence of Parkinson's disease (PD) are rapidly increasing worldwide [1], and PD is the second most common neurological disorder [2], [3]. Dopamine transporter (DaT) SPECT imaging is useful for the precise diagnosis and clinical management of PD [4], [5], especially with the use of the U.S. FDA approved ^{123}I-Ioflupane ligand tracer [6]. The striatum uptake of PD patients can be observed noninvasively via DaT SPECT at a single time point, with uptake reduction and asymmetry relate to disease severity [7], while disease progression can be monitored by longitudinal SPECT. PD diagnosis from DaT SPECT is usually based on visual assessment, which is subjective and could be biased by intra- and inter-reader variability [8].

A reliable quantitative analysis on DaT SPECT requires accurate segmentation of the striatal region. However, segmenting individual striatal compartments, i.e., left caudate (LC) and left putamen (LP), right caudate (RC) and right putamen (RP), on SPECT is challenging due to the inferior spatial resolution from current general-purpose scanners. Thresholding-based techniques [9], clustering-based algorithms [10], and region-based active contour methods [11] have been developed to assist SPECT segmentation, while their inability to measure uptake of individual striatal compartments limit their clinical indications [12]. On the other hand, MRI provides high spatial resolution and excellent soft tissue contrast for the striatum segmentation and subsequent quantitative analysis, which is regarded as the ground truth [13], [14]. Although there are existing automatic or semi-automatic MR image segmentation tools, such as FreeSurfer [15] and FMRIB Software Library [16], manual segmentation is still the gold standard for striatal analysis. However, it is very time-consuming and labor-intensive. In addition, integrated SPECT/MR systems are not yet commercially available, so precise registration between the MR and SPECT images is still required if MR is used to guide the SPECT segmentation. Some patients also do not have prior MR scans. Therefore, an automatic cross-modality segmentation method for SPECT is desirable, to obtain the corresponding MR striatum contour maps.

Recently, deep learning (DL)-based segmentation methods have been extensively developed for various medical images [17], and preliminarily applied on simulated DaT SPECT [18], while it is clinically relevant to explore DL methods for cross-modality striatum segmentation on the clinical SPECT data. Existing works using cross-modality image segmentation mainly focus on structural imaging, such as MRI to CT or CT to MRI [19], [20], and some works use CT, MR, and PET images for multimodality image segmentation [21], [22].

Manuscript received 1 April 2024; accepted 4 May 2024. Date of publication 8 May 2024; date of current version 5 September 2024. This work was supported in part by the Collaborative Research Grant from University of Macau under Grant MYRG-CRG2022-00011-ICMS; in part by the Science and Technology Development Fund of Macau under Grant 0016/2023/RIB1; in part by the National Natural Science Foundation of China under Grant 82372038; in part by the Shenzhen Excellent Technological Innovation Talent Training Project of China under Grant RCJC20200714114436080; and in part by the Shenzhen Science and Technology Program of China under Grant JCYJ20220818101804009. *(Corresponding authors: Zhanli Hu; Greta S. P. Mok.)*

Please see the Acknowledgment section of this article for the author affiliations.

This article has supplementary material provided by the authors and color versions of one or more figures available at https://doi.org/10.1109/TRPMS.2024.3398360.

Digital Object Identifier 10.1109/TRPMS.2024.3398360

TABLE I
PATIENT CHARACTERISTICS IN THIS STUDY

	Training data	Validation data	Testing data	Total
Number of subjects	136	24	40	200
Gender (Male/Female)	89/47	17/7	25/15	131/69
Age* (Year)	61.14± 9.39	60.43± 9.78	60.43± 10.49	60.91± 9.67
Diagnosis (Number, Percentage)	PD (104, 76%) HC (32, 24%)	PD (18, 75%) HC (6, 25%)	PD (30, 75%) HC (10, 25%)	PD (152, 76%) HC (48, 24%)

*Data are presented as mean ± standard deviation.

Similar techniques should also be applicable to SPECT. Among the existing DL methods, U-Net is a classic and widely used segmentation network for medical images [23]. It adopts a symmetric encoder–decoder structure and fuses multiscale features through skip connections [24]. In addition, generative adversarial network (GAN) is also gradually widely used in medical image segmentation, such as using conditional GAN (cGAN) for breast mass segmentation [25] and attention-assisted GAN for cerebrovascular segmentation [26]. However, adversarial models often face the challenge of training instability [27]. On the other hand, nnU-Net is an improved and extended network based on the U-Net model [28], which is more stable than the adversarial models. It retains the basic architecture of U-Net, while introducing some key improvements, such as a self-configuring network structure and various built-in preprocessing and post-processing techniques. nnU-Net exhibits stronger generalization performance through a series of improvements as compared to U-Net [29], especially in MR brain image segmentation [30]. However, it has not been applied to SPECT images as well as for cross-modality image segmentation. In this study, we propose an nnU-Net-based cross-modality striatum segmentation method to estimate MR striatum contour maps from clinical DaT SPECT images, and compare it with 1) SPECT thresholding-based segmentation (THR-Seg) method; 2) standard U-Net; 3) cGAN [31]; and 4) squeezed-and-excitation attention-based cGAN (Att-cGAN) [32] models.

II. MATERIALS AND METHODS

A. Patient Data Set

^{123}I-Ioflupane DaT SPECT and T1-weighted MR images from 200 anonymized subjects [152 PD and 48 healthy controls (HCs)] were analyzed from the Parkinson's progression markers initiative (PPMI) database (http://www.ppmi-info.org, [33]). All human subject research procedures and protocols were exempt from review board approval. The patient data were divided for training, validation, and testing (136: 24: 40), while keeping the similar proportion (∼3:1) of the PD and HC in each group (Table I).

The image size of the DaT SPECT data was $91 \times 109 \times 91$ and the voxel size was $2 \times 2 \times 2$ mm^3. The matrix sizes of MR images varied, including nine types, such as $176 \times 240 \times 256$, $256 \times 240 \times 160$, and $256 \times 256 \times 176$, etc., with a voxel size of $1 \times 1 \times 1$ mm^3. The SPECT images were first rigidly co-registered to their corresponding MR images automatically using the ITK-SNAP [34], i.e., using MR as the fixed images and then refined manually. Four individual striatal compartments, i.e., LC, RC, LP, and RP, were manually delineated from the MR images, serving as the gold standard. The registration, resampling, and segmentation processes were all performed by a nuclear medicine physician with ten years of experience [Fig. 1(a)]. To ensure a consistent input image size for the network model and efficiency of training, all MR and SPECT images were converted to a matrix size of $128 \times 128 \times 48$ and a voxel size of $1 \times 1 \times 1$ mm^3, extracting the striatum region for further analysis.

In addition, the activity uptake within caudate and putamen varied in clinical SPECT images for different PD stages. We further stratified the testing dataset according to SPECT visual interpretation assessment scheme [35] (Fig. S1 in the supplementary material): 1) normal category (NC); 2) abnormal category 1 (ANC1); 3) abnormal category 2 (ANC2); and 4) abnormal category 3 (ANC3). Moreover, in the clinical practice, the patient head orientations could be different across different DaT SPECT scans. We investigated this variation by rotating the testing patient datasets from 0° to 180°, with an interval of 10°.

Z-scoring normalization and various data augmentations, such as scaling, rotation, and Gaussian blur were used to preprocess the SPECT brain images for nnU-Net. Based on our ablation experiment, Gaussian blur improved our segmentation model performance (Table S1 in the supplementary material), while mirror flipping misclassified the left or right striatal components, degrading segmentation performance (Table S2 in the supplementary material). Thus, mirror flipping augmentation was not used in this study. Unlike nnU-Net, the baseline networks did not build-in a self-configuring network structure with various pre- and post-processing techniques. Therefore, min-max normalization and four times rotation-based augmentation were implemented for the baseline networks. All processed SPECT images were then paired with MR-based striatal contour maps (containing LC and LP, RC and RP) before inputting to the networks as supervised learning [Fig. 1(b)]. For THR-Seg, the threshold was set to be 67% of the maximum intensity of the SPECT images for each subject [36], [37].

B. Deep Learning Method

A 3-D nnU-Net was used for cross-modality striatum segmentation [Fig. 1(c)]. Similar to a standard 3-D U-Net, the network included a down-sampling encoder followed by an up-sampling decoder. In the down-sampling path, the input image first underwent one layer of two $1 \times 3 \times 3$ convolution operations, and then went through five layers of two $3 \times 3 \times 3$ convolutions. Each convolution was followed by an instance normalization (IN) layer and a leaky rectified linear unit (LReLU) activation function. The up-sampling process was implemented by a transposed convolution. Four layers of

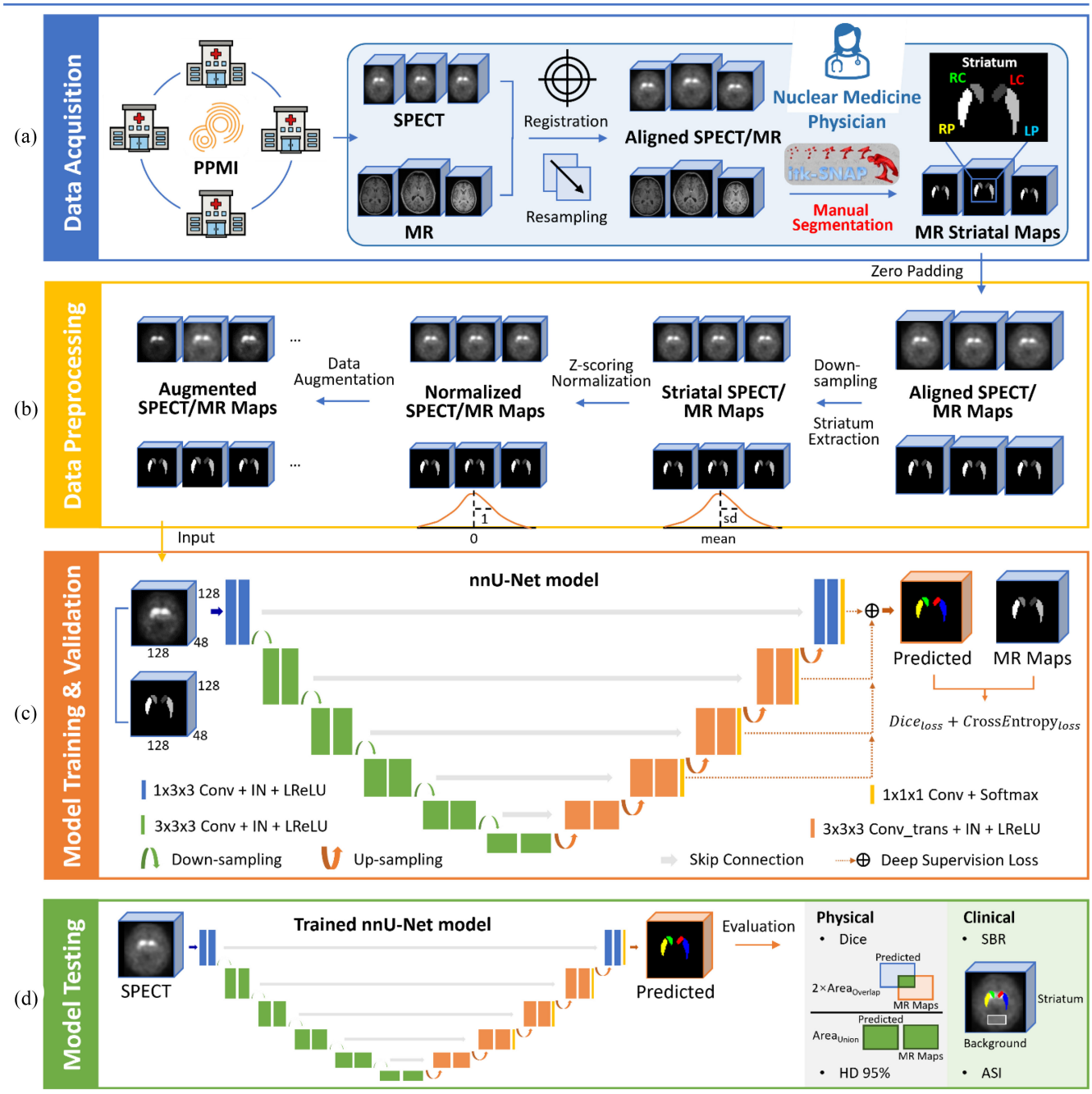

Fig. 1. Workflow and nnU-Net model used in this study. (a) Data acquisition from PPMI database. (b) Data pre-processing of down-sampling, normalization and augmentation. (c) Deep learning model architecture in the training and validation process. (d) Model testing and evaluation metrics.

two $3 \times 3 \times 3$ convolutions were performed first, and one layer of two $1 \times 3 \times 3$ convolutions were followed. Each convolution was also followed by an IN layer and an LReLU activation function, similar to the down-sampling process. An $1 \times 1 \times 1$ convolution and softmax activation function were added in each up-sampling layer (except the first two layers). Their convolution output was mapped to a probability distribution to calculate the deep supervision loss (L) and was used to incorporate feature information from deeper layers to shallow layers for training the network. L integrated Dice loss and cross-entropy loss between the layers output and the corresponding down-sampled MR labels, and it was the weighted sum of the losses output by the decoders, i.e., $L = w_1 \times L_1 + w_2 \times L_2 + w_3 \times L_3 + w_4 \times L_4$, the weights (w) were sequentially halved as the feature map decreases, i.e., $w_1 = 8/15$, $w_2 = w_1/2 = 4/15$, $w_3 = w_2/2 = 2/15$, $w_4 = w_3/2 = 1/15$, and the sum of all weights was 1. The four individual striatal compartments correspond to the four channels of model output, and L was calculated independently for each channel. We implemented the network using the PyTorch which ran on a NVIDIA RTX A6000 GPU. The stochastic gradient descent (SGD) with a Nesterov momentum optimizer was used to optimize the segmentation model with an initial learning rate of 0.01, running up to 1000 epochs. The

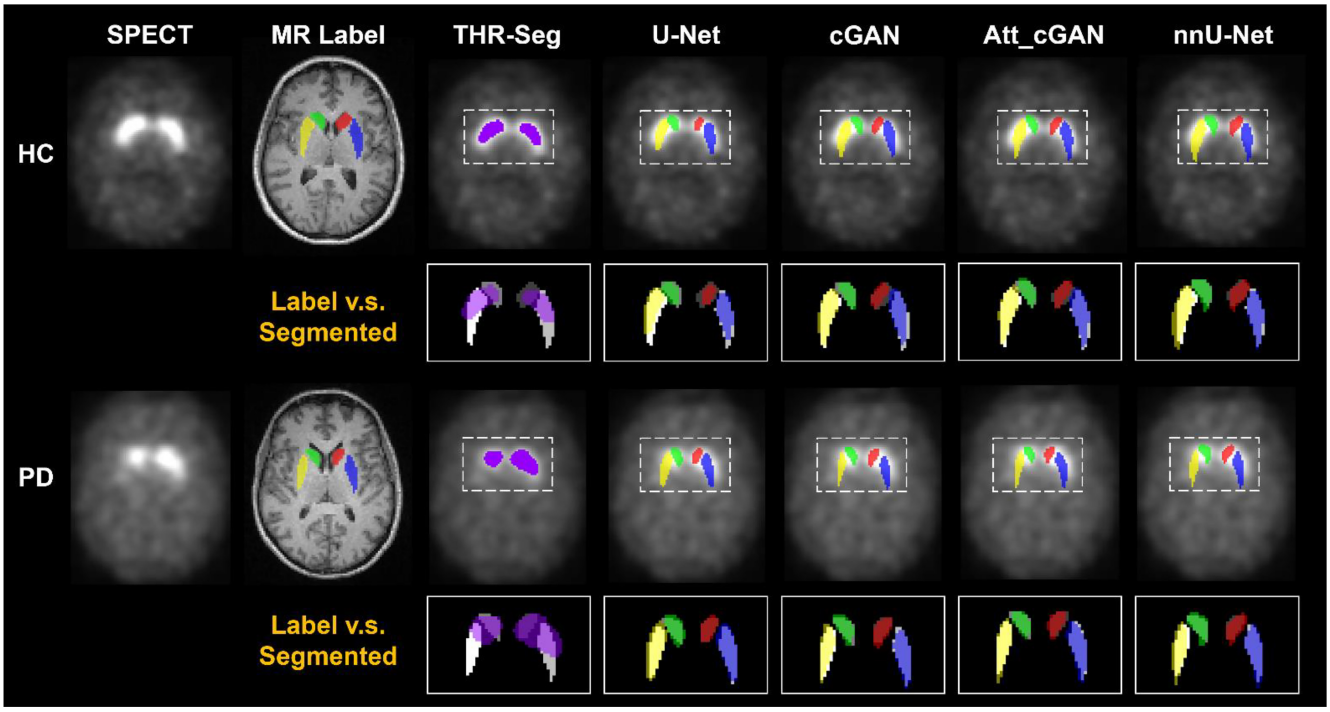

Fig. 2. Comparison of different striatal segmentation methods. Sample images for one HC and one PD subjects. Grey: MR striatal labels. Purple: WS. Green: RC. Yellow: RP. Red: LC. Blue: LP.

experimental environment, basic network parameter settings, training, validation, and testing data of the baseline networks were consistent with those of the nnU-Net model. The model training time was 16.6 h for nnU-Net, 18.5 h for U-Net, 22.5 h for cGAN, and 24 h for Att-cGAN.

C. Data Analysis

Physical metrics, i.e., Dice [38] (1), Hausdorff distance (HD) 95% [39] (2-1) and relative volume difference (RVD) [40] (3) were employed to evaluate the segmentation performance between segmented striatal maps of different algorithms and MR labels

$$\text{Dice} = \frac{2 \times |A \cap B|}{|A| + |B|} \tag{1}$$

$$\text{HD } 95\% = \max(h(A,B), h(B,A)) \times 95\% \tag{2-1}$$

$$h(A,B) = \max_{a \in A} \left\{ \min_{b \in B} \|a - b\| \right\} \tag{2-2}$$

$$h(B,A) = \max_{b \in B} \left\{ \min_{a \in A} \|b - a\| \right\} \tag{2-3}$$

$$\text{RVD} = \frac{|\text{Volume}_B - \text{Volume}_A|}{\text{Volume}_A} \tag{3}$$

where A is the MR label and B is the segmented striatal map. $\|\cdot\|$ means the Euclidean distance between point a and point b. Metrics of the whole striatum (WS) and individual striatal compartments were presented in violin plots.

For the clinical evaluation, striatal binding ratio (SBR) was used to quantify the binding of ^{123}I-Ioflupane in the striatum [14] (4). The striatal asymmetry index (ASI) was used to assess the asymmetry of uptake between the left and right striatum [13] (5)

$$\text{SBR} = \frac{\text{Mean\_Counts}_{\text{ROI}} - \text{Mean\_Counts}_{\text{background}}}{\text{Mean\_Counts}_{\text{background}}} \tag{4}$$

$$\text{ASI} = \left| \frac{\text{SBR}_{\text{left ROI}} - \text{SBR}_{\text{right ROI}}}{\text{SBR}_{\text{left ROI}} + \text{SBR}_{\text{right ROI}}} \right| \times 100\% \tag{5}$$

where the regions-of-interest (ROIs) include the WS and individual striatal compartments. The background area was chosen from a uniform cerebellum region ($20 \times 10 \times n$, $n = 13 - 18$, depending on the size of the striatum), excluding ventricular regions [Fig. 1(d)]. The SBR differences, 95% confidence interval (CI) and mean absolute difference (MAD) between the striatal compartments from different segmentation methods were compared with the MR labels, presented in Bland-Altman plots. A paired two-tailed student t-test was used to assess the statistical difference between different segmentation methods and MR labels, with a p-value of <0.05 indicated statistical significance. Meanwhile, the ASI of the segmented striatal maps was also correlated with the MR labels. R^2 and Pearson correlation coefficient (PCC) were obtained in the scatter plots.

III. Result

Sample SPECT images, MR labels and segmentation results from one HC and one PD subjects are shown in Fig. 2. For the THR-Seg method, the uptake in the striatum in the SPECT image can be directly obtained, but the caudate and putamen cannot be separated. For the DL-based methods, four striatal compartments, including LC, RC, LP, and RP, can be well

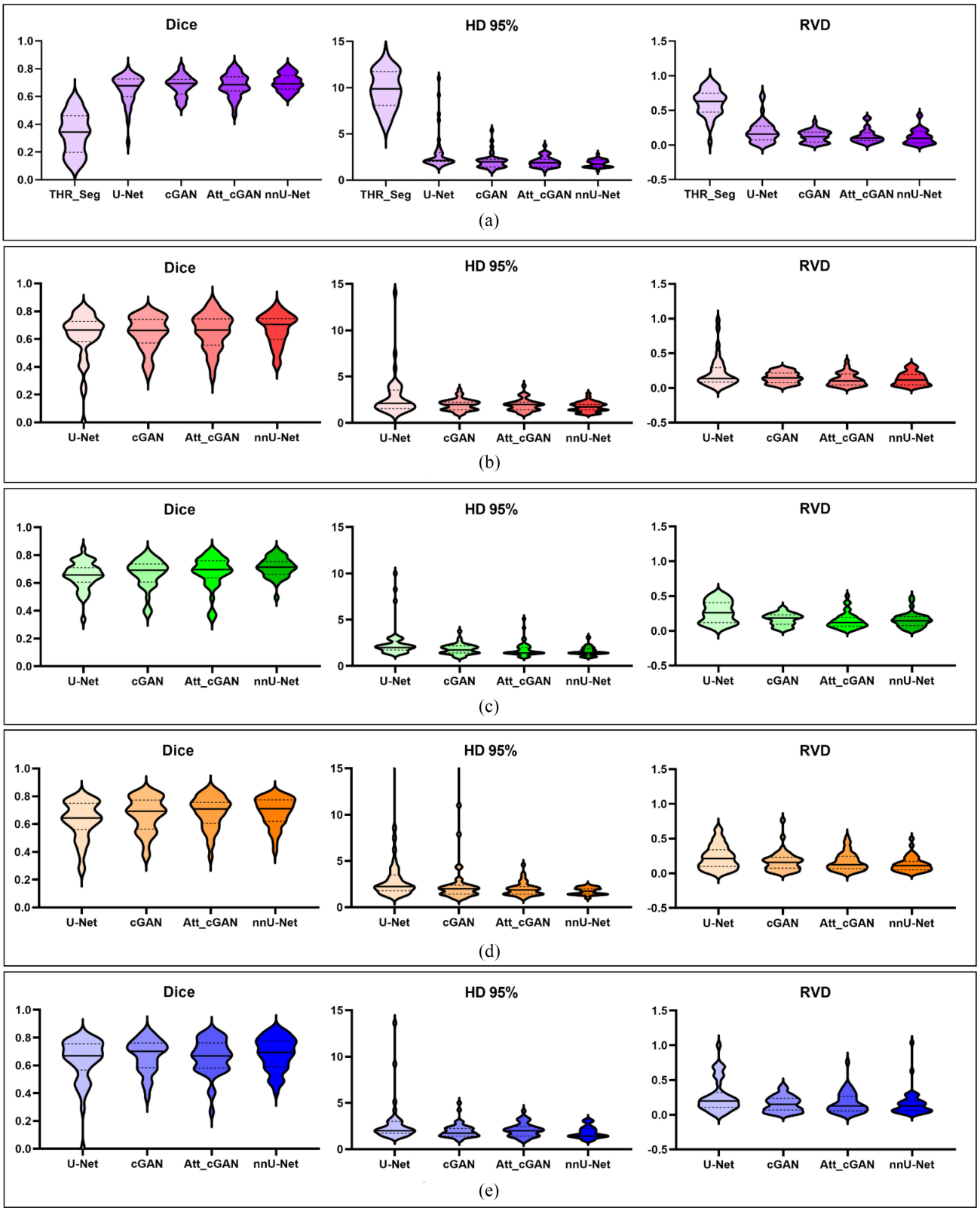

Fig. 3. Violin plots of physical metrics between MR labels and maps segmented by different methods for WS and individual compartments. The lines from top to bottom in the plots represent the first quartile, median, and third quartile. (a) WS. (b) LC. (c) RC. (d) RP. (e) LP.

separated, and the segmented striatum maps are relatively similar to the MR labels for these two patients.

Fig. 3 shows the physical evaluation results of the WS and four striatal individual compartments of 40 tested subjects, including Dice, HD 95%, and RVD, and is presented as a set of violin plots. For the WS, DL-based segmentation methods are obviously superior to the THR-Seg method, with higher Dice, lower HD 95%, and RVD. For each striatal

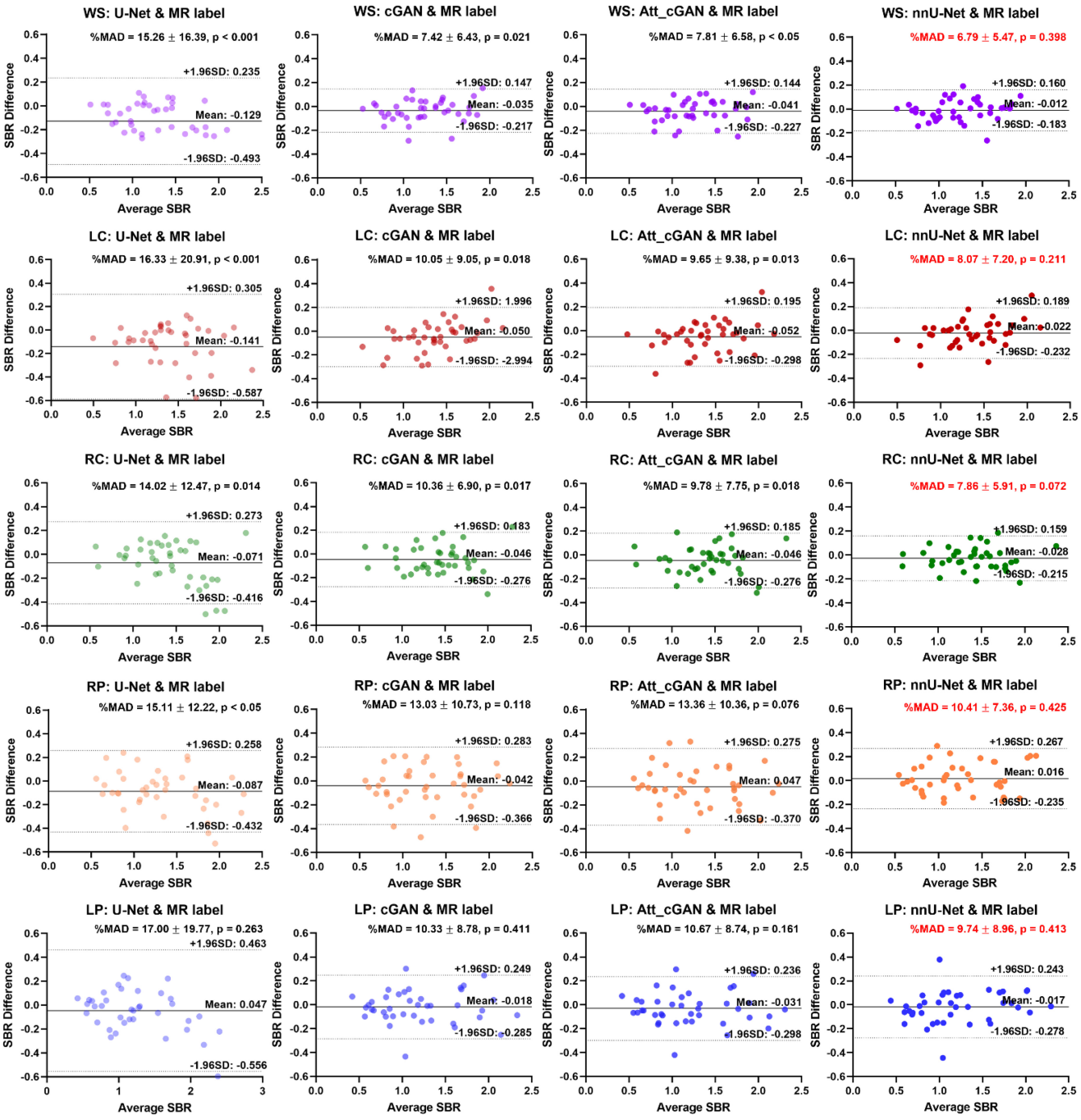

Fig. 4. Bland-Altman plots of SBR between MR labels and maps segmented by different methods for WS and individual compartments. The MAD and p-values are also shown. LC: left caudate. RC: right caudate. RP: right putamen. LP: left putamen.

compartment, it cannot be segmented by THR-Seg, so the corresponding physical indices cannot be obtained. Compared to other networks, nnU-Net achieves better Dice (∼0.7), HD 95% (∼1.8), and RVD (∼0.1).

Fig. 4 analyses the consistency of the clinical indicator, i.e., SBR between the MR labels and maps segmented by different methods for the WS and individual compartments using the Bland-Altman plots. The MAD and p-values between the labels and segmented striatal maps are also presented. For the WS, nnU-Net exhibits better consistency (95% CI of [−0.183, 0.160]) as compared to THR-Seg ([−1.185, 0.109]), U-Net ([−0.493, 0.235]), cGAN ([−0.217, 0.147]), and Att-cGAN ([−0.227, 0.144]). The SBR difference between the THR-Seg and MR label is relatively large (−0.538) and not shown on the plot. Meanwhile, nnU-Net also obtains the lowest %MAD (6.79 ± 5.47), and there is no significant difference ($p = 0.398$) between striata segmented by the nnU-Net and MR labels. The segmentation performance of both the GAN-based methods (cGAN, %MAD, 7.42 ± 6.43 and Att-cGAN, %MAD, 7.81 ± 6.58) is better than the U-Net (%MAD, 15.26 ± 16.39). For the caudate and putamen components, nnU-Net also achieves better consistency and lower MAD than others.

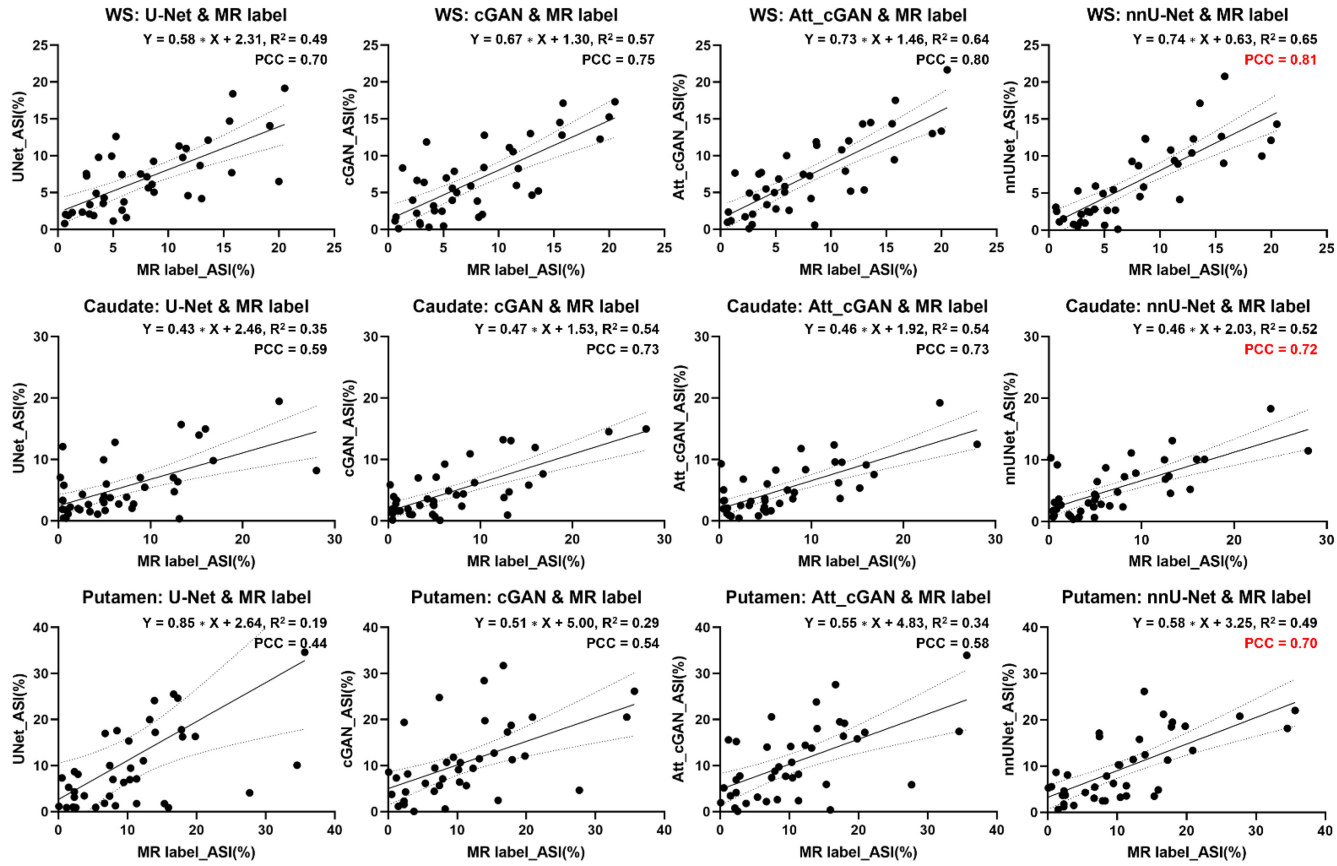

Fig. 5. Scatter plots of ASI between MR labels and maps segmented by different methods for WS and individual compartments. The solid lines represent the best-fit lines and the dotted lines represent the 95% confidence bands of the fitted lines.

Fig. 5 illustrates the correlation assessment of the another clinical indicator, i.e., ASI between MR labels and maps segmented by different methods using scatter plots. Results show stronger correlation between striata segmented by nnU-Net and MR labels with a fitting slop of 0.74 and a PCC of 0.81, than that of Att-cGAN (slope, 0.73 and PCC, 0.80), cGAN (slope, 0.67 and PCC, 0.75), and U-Net (slope, 0.58 and PCC, 0.70). The fitting results between striata segmented by the THR-Seg and MR label is not shown due to its inferior performance (slope, 1.04 and PCC, 0.37). For the individual striatal compartments, the ASI assessment results of nnU-Net are also generally superior (PCC = 0.72 in caudate and 0.70 in putamen) to others.

We find that with the decrease of caudate/putamen uptake, the segmentation performance of the proposed nnU-Net-based model declines (Fig. S2, Table S3 in the supplementary material). For example, for NC, Dice can reach 0.75, but drop to 0.73 for ANC1. For ANC2 and ANC3, Dice is only 0.68. Our results (Figs. S3 and S4 in the supplementary material) also show that with head rotation starting from 50°, the segmented striatum begins to deviate greatly from the MR label, with the Dice begins to drop <0.6; HD 95% exceeds 3; and RVD fluctuate relatively smoothly between 0.1–0.25.

IV. Discussion

This work proposes DL-based cross-modality automatic striatum segmentation methods to separate four individual striatal compartments based on the clinical DaT SPECT data. Selecting MRI striatal contours as labels for the cross-modality segmentation is based on the following two points.

1) The uptake of each striatal compartment can usually not be able to be resolved on SPECT images due to its inferior spatial resolution, making it impossible to conduct quantitative analysis on each compartment which has significant clinical impact [41]. For example, the caudate is involved in movement control and memory formation while the putamen is involved in action planning and execution. The striatal compartment structures can be discriminated on MR images to obtain SBR and ASI of different striatal structures [13], [14].
2) Although SPECT THR-Seg method is also commonly used for DaT SPECT, the segmentation accuracy highly depends on the selected threshold. In addition, SPECT images are more affected by the partial volume effect due to the limited resolution, increasing the difficulty of SPECT-based segmentation. On the contrary, the structural boundary of the striatum in MR images are relatively fixed along the development of the disease. The proposed method would allow for more consistent quantification of the DaT SPECT uptake to monitor the progression of PD. However, comparison of our proposed method with other baseline segmentation methods for longitudinal images is beyond the scope of this study.

In data preprocessing, the rigid registration of MR and SPECT before manual segmentation is important as they are acquired at different time points and under different scanners, and should be appropriate as the head is a rigid body. However, cross modality registration is challenging as MR and SPECT are with different image characteristics and resolutions, which may affect the accuracy of the registration. However, the two images are generally well aligned after the registration based on our visual assessment in this study. To model a more realistic clinical scenario where patients' head positions may be tilted, we test SPECT images with different rotation angles. We find that the performance of the proposed algorithm is relatively stable within a rotation of 50°, which is applicable for most real clinical cases.

From the visual assessment, compared with the standard U-Net, cGAN, and Att-cGAN, the proposed nnU-Net-based method obtains more similar shape of striatal components with MR labels, while individual striatal compartments cannot be separated by the SPECT THR-Seg method. Dice, HD 95%, and RVD indices are used to quantitatively evaluate the spatial similarity and volume difference of segmented striatum. Dice is more sensitive to the internal padding, while HD 95% pays more attention to the boundaries and RVD measures the volumetric accuracy of segmentation results. From Fig. 3, nnU-Net yields more stable and better indices, probably attributed to the optimized network parameter selection according to a specific task. For nnU-Net, we also considered the difference between the volume of each segmented striatal component and the volume of the corresponding MR label in the loss function, but no improved segmentation performance observed, thus it was not incorporated in the final version. On the other hand, the physical metrics of GAN-based methods are generally better than standard U-Net due to the adversarial training characteristic of GAN, consistent with existing studies [42].

In the PD-related clinical analysis, SBR can be used to differentiate PD patients and HC, assess the impact of drug efficacy, and track changes in striatal binding during disease progression [43]. On the other hand, higher ASI values indicate greater asymmetry, which may be associated with disease progression or pathological changes [44]. The striatal map segmented by nnU-Net has a strong SBR consistency (mean SBR difference =-0.012) and ASI correlation (PCC = 0.81) with the MR label. It is also superior to other segmentation methods by comparing SBR difference and 95% CI, as well as ASI fitting slope and PCC. In the correlation assessment of ASI, GAN-based methods still exhibit better performance compared to U-Net, which is consistent with physical evaluation.

The cross-modality segmentation task is relatively challenging. In the model testing stage, only SPECT data is input, and the expected output is the striatal contour maps segmented from the paired MR image, containing four individual compartments. There is a large gap between the input and the output, and almost all of the structural information is obtained through the supervised DL network, resulting a Dice of ∼0.7 for the proposed method. To explore if various activity uptake patterns in the caudate and putamen on SPECT images at different PD stages affect the segmentation performance, we stratify the patient data into NC and PD stages, and find that the segmentation performance decreases with the deterioration of PD. The attributed reason is probably that as PD progresses, the uptake in the putamen gradually decreases (ANC1-2), while even the caudate uptake begins to decrease for ANC3. Therefore, it is more difficult to learn the MR contours from SPECT images. However, the subsequent SBR and ASI obtained by the proposed method still have considerable consistency and correlation with MR labels in clinical analysis, indicating our method should be clinically relevant. The existing automatic striatum segmentation study [18] proposes a tissue-fraction estimation-based DL segmentation method evaluated on simulated SPECT data, with a Dice of ∼0.8. Our results may not be directly compared as their study was based on 580 simulated data with only normal subjects while ours was based on 200 clinical data with both the PD and HC subjects from multiple centers. The SPECT image of PD has less striatal uptake and is much different to the corresponding MR image (Fig. 2). Besides, the anatomical gold standard of the reference study was obtained by automatic segmentation of MR images using the Freesurfer software, while ours was based on manual segmentation. Therefore, our task is expected to be more challenging than the reference study and lower Dice should be justified.

In addition to the striatum, other structures in the brain, such as the globus pallidus and substantia nigra could also be used for PD diagnosis [45]. The globus pallidus can receive neuronal information from the striatum for movement regulation, while the substantia nigra is a major dopamine-producing region in the brain. For the proposed nnU-Net-based model, segmenting the globus pallidus and substantia nigra requires corresponding MR label data for supervised training. The globus pallidus is located on the inner side of the striatum and is smaller than the striatum; while the substantia nigra is located in the midbrain and is smaller than the striatum and globus pallidus. Segmenting these small and adjacent brain structures across modalities is a greater challenge. Advanced network models and image processing technologies to segment these brain structures for PD analysis are warranted for further investigations.

There are other limitations for this study. More dataset will be warrant to verify the effectiveness of the proposed model. Besides, although PPMI contains multicenter and multiscanner data, more external data is still needed for model verification to explore the generalization and robustness of the segmentation model. In addition, the striatum maps segmented by the proposed method still have certain discrepancies with the MR labels, especially at the edge of the striatum. More advanced models are warranted to further improve the segmentation performance.

V. Conclusion

Our proposed DL-based cross-modality automatic striatum segmentation method, is feasible to segment four MR-like individual compartments on clinical DaT SPECT for PD based on evaluation on the PPMI multicenter data. The proposed nnU-Net-based method is superior to standard U-Net, cGAN,

Att-cGAN, and SPECT-based THR-Seg method, both in physical (Dice, HD 95% and RVD) and clinical metrics (SBR and ASI). The proposed DL method is promising for the clinical DaT SPECT segmentation.

ACKNOWLEDGMENT

All authors declare that they have no known conflicts of interest in terms of competing financial interests or personal relationships that could have an influence or are relevant to the work reported in this article.

This work involved human subjects in its research. The authors confirm that all human subject research procedures and protocols are exempt from review board approval.

Haiyan Wang is with the Biomedical Imaging Laboratory, Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, SAR, China, and also with the Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: yc27972@um.edu.mo).

Han Jiang was with the Biomedical Imaging Laboratory, Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, SAR, China. She is now with the PET-CT Center, Fujian Medical University Union Hospital, Fuzhou 350001, China (e-mail: jianghan0123@163.com).

Gefei Chen was with the Biomedical Imaging Laboratory, Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, SAR, China. He is now with the Jiangsu Rayer Medical Technology Co., Ltd., Wuxi 214192, China. (e-mail: seucgf@gmail.com).

Yu Du and Greta S. P. Mok are with the Biomedical Imaging Laboratory, Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, SAR, China, and also with the Center for Cognitive and Brain Sciences, Institute of Collaborative Innovation, University of Macau, Macau, SAR, China (e-mail: yc07905@um.edu.mo; gretamok@um.edu.mo).

Zhonglin Lu was with the Biomedical Imaging Laboratory, Department of Electrical and Computer Engineering, Faculty of Science and Technology, University of Macau, Macau, SAR, China, and also with the Center for Cognitive and Brain Sciences, Institute of Collaborative Innovation, University of Macau, Macau, SAR, China. She is now with the Division of Nuclear Medicine and Molecular Imaging, Department of Radiology, University of Michigan, Ann Arbor, Michigan (e-mail:zhonglil@med.umich.edu).

Zhanli Hu is with the Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China (e-mail: zl.hu@siat.ac.cn).

REFERENCES

[1] B. R. Bloem, M. S. Okun, and C. Klein, "Parkinson's disease," *Lancet*, vol. 397, no. 10291, pp. 2284–2303, 2021.

[2] I. Ozsahin, B. Sekeroglu, P. C. Pwavodi, and G. S. Mok, "High-accuracy automated diagnosis of Parkinson's disease," *Current Med. Imag.*, vol. 16, no. 6, pp. 688–694, 2020.

[3] Z. Zheng, Z. Zhu, C. Zhou, L. Cao, and G. Zhao, "Burden of Parkinson disease in China, 1990-2019: Findings from the 2019 global burden of disease study," *Neuroepidemiology*, vol. 57, no. 1, pp. 51–64, 2023.

[4] H. Sun et al., "Cross-tracer and cross-scanner transfer learning-based attenuation correction for brain SPECT," *IEEE Trans. Radiat. Plasma Med. Sci.*, early access, Mar. 8, 2024, doi: 10.1109/TRPMS.2024.3374207.

[5] K. Tsang and R. Walker, "Dopamine transporter single photon emission computed tomography (DaT-SPECT) use in the diagnosis and clinical management of parkinsonism: An 8-year retrospective study," *J. Neurol.*, vol. 270, no. 5, pp. 2550–2558, 2023.

[6] K. Traynor, "New drug and biological product approvals, 2011," *Amer J. Health-Syst. Pharmacy*, vol. 69, no. 3, pp. 180–182, 2012.

[7] T. Booth, M. Nathan, A. Waldman, A.-M. Quigley, A. Schapira, and J. Buscombe, "The role of functional dopamine-transporter SPECT imaging in parkinsonian syndromes, part 2," *Amer J. Neuroradiol.*, vol. 36, no. 2, pp. 236–244, 2015.

[8] A. Augimeri et al., "CADA—Computer-aided DaTSCAN analysis," *EJNMMI Phys.*, vol. 3, no. 1, pp. 1–13, 2016.

[9] J. Grimes, A. Celler, S. Shcherbinin, H. Piwowarska-Bilska, and B. Birkenfeld, "The accuracy and reproducibility of SPECT target volumes and activities estimated using an iterative adaptive thresholding technique," *Nuclear Med. Commun.*, vol. 33, no. 12, pp. 1254–1266, 2012.

[10] J. Chen et al., "Learning fuzzy clustering for SPECT/CT segmentation via convolutional neural networks," *Med. Phys.*, vol. 48, no. 7, pp. 3860–3877, 2021.

[11] C. Huang et al., "A hybrid active contour segmentation method for myocardial D-SPECT images," *IEEE Access*, vol. 6, pp. 39334–39343, 2018.

[12] M. Otsuka et al., "Differences in the reduced 18F-Dopa uptakes of the caudate and the putamen in Parkinson's disease: Correlations with the three main symptoms," *J. Neurol. Sci.*, vol. 136, nos. 1–2, pp. 169–173, 1996.

[13] Y.-H. Weng et al., "Sensitivity and specificity of 99mTc-TRODAT-1 SPECT imaging in differentiating patients with idiopathic Parkinson's disease from healthy subjects," *J. Nuclear Med.*, vol. 45, no. 3, pp. 393–401, 2004.

[14] B.-H. Yang et al., "Automatic striatal ROI delineation for semi-quantitative analysis of 99m Tc-TRODAT-1 brain SPECT imaging," *J. Med. Biol. Eng.*, vol. 40, pp. 428–439, 2020.

[15] B. Fischl, "FreeSurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[16] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "FSL," *Neuroimage*, vol. 62, no. 2, pp. 782–790, 2012.

[17] P.-H. Conze, G. Andrade-Miranda, V. K. Singh, V. Jaouen, and D. Visvikis, "Current and emerging trends in medical image segmentation with deep learning," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 7, no. 6, pp. 545–569, Jul. 2023.

[18] Z. Liu et al., "A tissue-fraction estimation-based segmentation method for quantitative dopamine transporter SPECT," *Med. Phys.*, vol. 49, no. 8, pp. 5121–5137, 2022.

[19] K. Yao et al., "A novel 3D unsupervised domain adaptation framework for cross-modality medical image segmentation," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 10, pp. 4976–4986, Oct. 2022.

[20] Y. Zhuang et al., "A 3-D anatomy-guided self-training segmentation framework for unpaired cross-modality medical image segmentation," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 8, no. 1, pp. 33–52, Jan. 2024.

[21] Z. Guo, X. Li, H. Huang, N. Guo, and Q. Li, "Deep learning-based image segmentation on multimodal medical imaging," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 2, pp. 162–169, Mar. 2019.

[22] V. V. Valindria et al., "Multi-modal learning from unpaired images: Application to multi-organ segmentation in CT and MRI," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, 2018, pp. 547–556.

[23] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and its variants for medical image segmentation: A review of theory and applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021.

[24] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Interv. (MICCAI)*, Munich, Germany, 2015, pp. 234–241.

[25] V. K. Singh et al., "Conditional generative adversarial and convolutional networks for X-ray breast mass segmentation and shape classification," in *Proc. 21st Int. Conf. Med. Image Comput. Assist. Interv. (MICCAI)*, Granada, Spain, 2018, pp. 833–840.

[26] Y. Chen, D. Jin, B. Guo, and X. Bai, "Attention-assisted adversarial model for cerebrovascular segmentation in 3D TOF-MRA volumes," *IEEE Trans. Med. Imag.*, vol. 41, no. 12, pp. 3520–3532, Dec. 2022.

[27] M. Wiatrak, S. V. Albrecht, and A. Nystrom, "Stabilizing generative adversarial networks: A survey," 2019, *arXiv:1910.00927*.

[28] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nat. Methods*, vol. 18, no. 2, pp. 203–211, 2021.

[29] D. Veiga-Canuto et al., "Independent validation of a deep learning nnU-Net tool for neuroblastoma detection and segmentation in MR images," *Cancers*, vol. 15, no. 5, p. 1622, 2023.

[30] M. Baniasadi et al., "DBSegment: Fast and robust segmentation of deep brain structures considering domain generalization," *Human Brain Map.*, vol. 44, no. 2, pp. 762–778, 2023.

[31] Y. Du et al., "Generative adversarial network-based attenuation correction for 99mTc-TRODAT-1 brain SPECT," *Front. Med.*, vol. 10, Aug. 2023, Art. no. 1171118.

[32] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 7132–7141.
[33] K. Marek et al., "The Parkinson progression marker initiative (PPMI)," *Progr. Neurobiol.*, vol. 95, no. 4, pp. 629–635, 2011.
[34] P. A. Yushkevich, Y. Gao, and G. Gerig, "ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2016, pp. 3342–3345.
[35] D. S. Djang et al., "SNM practice guideline for dopamine transporter imaging with ^{123}I-Ioflupane SPECT 1.0," *J. Nuclear Med.*, vol. 53, no. 1, pp. 154–163, 2012.
[36] R. Prashanth, S. D. Roy, P. K. Mandal, and S. Ghosh, "High-accuracy classification of Parkinson's disease through shape analysis and surface fitting in 123I-Ioflupane SPECT imaging," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 3, pp. 794–802, May 2017.
[37] Y.-H. D. Fang, S.-C. Chiu, C.-S. Lu, T.-C. Yen, and Y.-H. Weng, "Fully automated quantification of the striatal uptake ratio of [(99m)Tc]-TRODAT with SPECT imaging: Evaluation of the diagnostic performance in Parkinson's disease and the temporal regression of striatal tracer uptake," *BioMed. Res. Int.*, vol. 2015, no. 461625, pp. 1–11, 2015.
[38] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool," *BMC Med. Imag.*, vol. 15, no. 1, pp. 1–28, 2015.
[39] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, Sep. 1993.
[40] Y. Xu et al., "PA-ResSeg: A phase attention residual network for liver tumor segmentation from multiphase CT images," *Med. Phys.*, vol. 48, no. 7, pp. 3752–3766, 2021.
[41] M. Oh et al., "Subregional patterns of preferential striatal dopamine transporter loss differ in Parkinson disease, progressive supranuclear palsy, and multiple-system atrophy," *J. Nuclear Med.*, vol. 53, no. 3, pp. 399–406, 2012.
[42] V. K. Singh et al., "Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network," *Exp. Syst. Appl.*, vol. 139, Jan. 2020, Art. no. 112855.
[43] E. Maly, J. Leverenz, and F. DiFilippo, "123I-Ioflupane SPECT striatal binding ratio predicts progression of Parkinsonism in a Dementia with Lewy bodies cohort (P11-6.004)," *Neurology*, vol. 139, no. 112855, pp. 1–14, 2020.
[44] A. McNeill et al., "Dopaminergic neuronal imaging in genetic Parkinson's disease: Insights into pathogenesis," *PLoS ONE*, vol. 8, no. 7, 2013, Art. no. e69190.
[45] G. Wu, Y.-J. Shen, M.-H. Huang, Z. Xing, Y. Liu, and J. Chen, "Proton MR spectroscopy for monitoring pathologic changes in the substantia nigra and globus pallidus in Parkinson disease," *Amer J. Roentgenol.*, vol. 206, no. 2, pp. 385–389, 2016.