

Deep Learning Segmentation Improves Precision of Volume Assessment of Subcortical Structures in early MSA

Kilian Hett¹, Paula Trujillo¹, Yan Yan², Cynthia Wong³, Hakmook Kang², David Stamler³, and Daniel O. Claassen¹

OBJECTIVE

Assessing volumetric changes related to **Multiple system atrophy (MSA)** in the **basal ganglia**.

Here, we assessed **the accuracy in early MSA** patient of three different segmentation techniques:

- FSL FIRST
- Joint Label Fusion (JLF)
- Ensemble deep-learning (AssemblyNet)

We assessed **volumetric difference** between HC, PD, and early MSA, in three basal ganglia structures:

- Putamen
- Caudate
- Pallidum

BACKGROUND

BioMUSE is a natural history study that aims to track the progression of patients with **MSA**, a rapidly progressive parkinsonian disorder that variably presents with parkinsonism, ataxia, and autonomic impairment.

- **Subcortical volume** has been proposed as a biomarker of disease progression.
- Contemporary segmentation algorithms **may not accurately identify subcortical structures in MSA** due to unclear anatomic boundaries in these structures and elevated iron concentration, especially in the putamen and pallidum.

METHODS

We enrolled 21 early MSA patients (age=64.4±9.3) (motor symptoms for less than 3 years), all at baseline: 17 with PD (age=63.2±6.1), and 18 HC (age=65.6±6.8). All underwent 3T brain MRI. Early MSA patients had mild to moderate clinical severity, based on neurologic assessment.

- A neuroradiologist manually delineated **putamen, caudate, and pallidum structures** using T1-weighted scans over 8 patients with MSA (each delineation took over 24 hours).
- **Segmentation accuracy** was estimated using *Sørensen-Dice* coefficient (DSC), *balanced error rate (BER)*, *false positive rate (FPR)*, and *False negative rate (FNR)* on the subset composed of manually delineated early MSA patients.
- **Group difference** was evaluated on the whole cohort using generalized linear model (group, age, and total intracranial volume as co-variates).

CONCLUSION

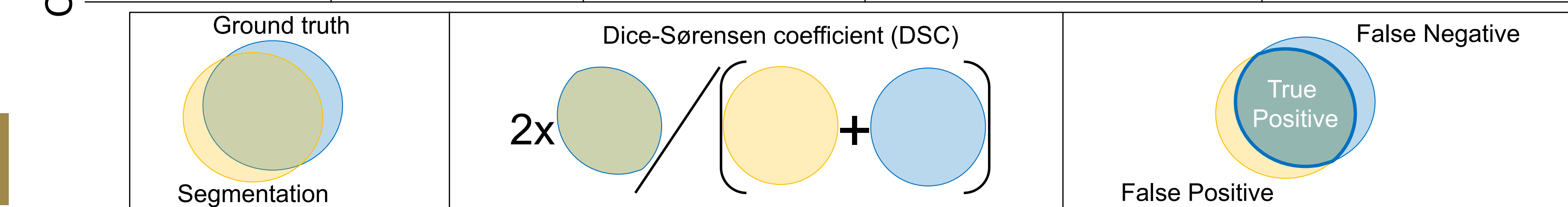
Ensemble **deep-learning segmentation promise** higher benefice compared to FSL and JLF methods (i.e., higher accuracy and lower and more balanced error rate). This improvement was observed in **subcortical structures with notable MSA pathology**.

- The absence of substantial accuracy drop in FSL-FIRST in the caudate – compared to others methods — indicates that MSA-related changes **hinder the precision of automatic segmentation from FSL-FIRST**.
- **In early MSA, significant reductions in putamen and pallidum** volume were observed compared to HC, and in pallidum volume compared to PD.
- This work will allow for a more accurate definition of subcortical structures, an essential step for quantifying changes in MSA, such as tissue atrophy and iron deposition.

RESULTS

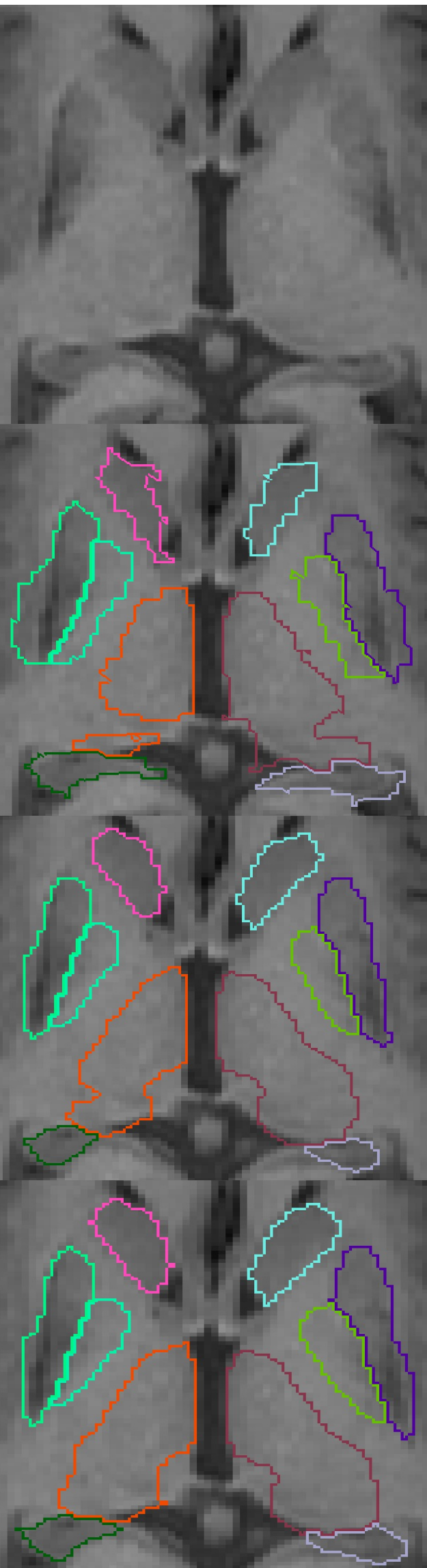
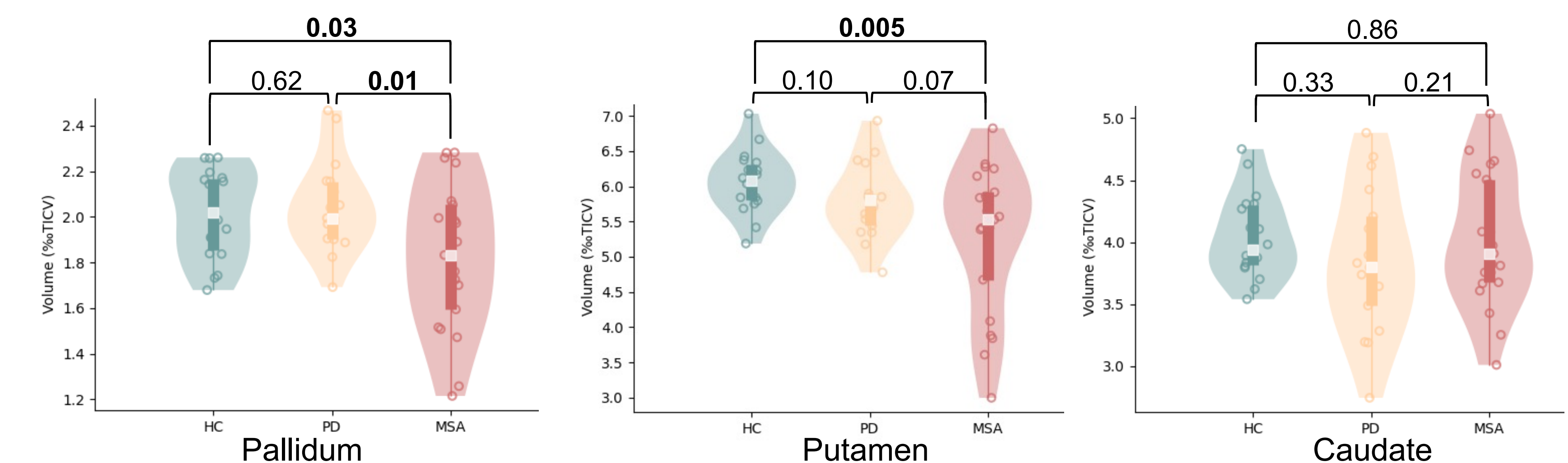
- Evaluation of the segmentation accuracy in MRI from 8 early MSA patients recruited in the BioMUSE study.

| Method | Accuracy DSC (σ) | Balanced Error Rate | False Positive Rate (% Struct. Volume) | False Negative Rate (% Struct. Volume) |
|-----------------|---------------------------|---------------------|--|--|
| Putamen | | | | |
| FSL-FIRST | 82.3 (6.0) | 17.9 (6.9) | 17.7 (13.1) | 18.0 (4.0) |
| JLF | 86.9 (2.0) | 12.5 (1.9) | 8.0 (4.7) | 16.9 (3.5) |
| AssemblyNet | 87.0 (3.0) | 12.5 (3.2) | 10.0 (4.0) | 15.3 (3.0) |
| Pallidum | | | | |
| FSL-FIRST | 76.6 (5.0) | 24.6 (6.2) | 29.8 (11.8) | 19.5 (5.0) |
| JLF | 80.5 (3.0) | 17.4 (2.8) | 6.0 (1.2) | 28.8 (4.0) |
| AssemblyNet | 81.2 (4.0) | 17.0 (3.5) | 11.1 (3.4) | 23.0 (5.0) |
| Caudate | | | | |
| FSL-FIRST | 79.8 (2.0) | 18.7 (1.6) | 11.6 (3.2) | 25.8 (2.7) |
| JLF | 81.5 (2.0) | 16.4 (1.1) | 5.8 (3.7) | 27.1 (4.4) |
| AssemblyNet | 80.5 (1.0) | 17.1 (0.7) | 5.8 (1.7) | 28.5 (2.5) |



Balanced error rate = average of False positive rate and False Negative rate.
Good methods should show high accuracy, low and balanced error rates (FPR and FNR)

- Evaluation of striatal volume difference between HC, PD, and MSA using AssemblyNet using the entire BioMUSE dataset at baseline (i.e., 18 HC, 17 PD, and 21 MSA).



¹ Department of Neurology, Vanderbilt University Medical Center, Nashville, TN, USA

² Department of Biostatistics, Vanderbilt University Medical Center, Nashville, TN, USA

³ Alterity Therapeutics, Newark, CA, USA