

Automated PET Tracer Classification Using Deep Learning

Robel K. Gebre, PhD¹, Robert I. Reid¹, PhD, Matthew L. Senjem, MS¹, Val J. Lowe, MD¹, Kejal Kantarci, MD, MS¹, Ronald C. Petersen, MD¹, Jonathan Graff-Radford¹, Clifford R. Jack Jr., MD¹, Prashanthi Vemuri, PhD¹ and

Christopher G. Schwarz, PhD¹

¹Mayo Clinic, Rochester, MN, USA

BACKGROUND

Experts manually classify positron emission tomography (PET) scans in large multi-site databases. PET is particularly challenging because the radiotracer used, the most crucial information, is unknown by the scanner; it must be entered manually at scan time, risking labeling errors. DICOM metadata can also be removed or incorrect.

OBJECTIVE

To develop a deep learning PET classifier that predicts the radiotracer directly from image appearance, independent of DICOM metadata.

METHODS

STUDY DATA

- 10 tracers collected across three cohorts:
 - ADNI, N = 7429
 - SCAN, N = 4869
 - MCSA and Mayo ADRC – Mayo, N = 10558.
- Not all tracers were available in all cohorts (**Table 1**):
 - E.g., Early amyloid perfusion protocols only in ADNI
 - Mayo only had 4 tracers

- ADNI and SCAN were combined to train our models:

MODELING

Two classification tasks were trained:

- 10-class
- 3-class

Preprocessing

- Averaged 10-midslices per view
- Normalized image intensity

METHODS 2

Convolutional neural network

- A **ConvNeXt large** model pretrained on ImageNet.
- These models are a family of pure CNNs models redesigned to mimic a vision transformer

Loss function (cross-entropy):

$$\mathcal{L}_{CE} = -\frac{1}{N} \sum_{i=1}^N \log p_{i,y_i}$$

Data partitioning (ADNI + SCAN)

- We split the data into 3 sets:

- Training, 60%
- Internal validation, 20%
- Internal testing, 20%
- Mayo was an external test set

- To prevent data leakage: participant-level stratified (**Fig. 1**).

- To ensure model stability, data was split 10x using random seeds:

- 10 trained models / 10-class
- 10 trained models / 3-class

Optimization

- AdamW (LR = 3×10^{-4} , batch = 64),
- LR scheduling, and
- Early stopping.

TABLE 1: Tracer Distribution in Study Data

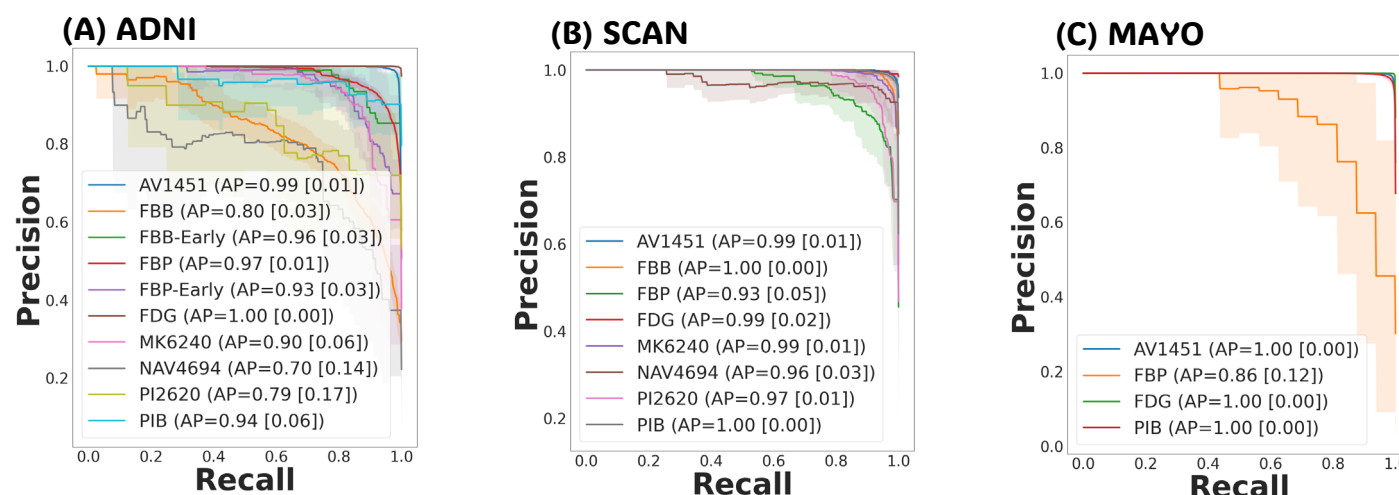
	ADNI	SCAN	Mayo
AV1451	1663	1168	4043
FBB	740	1102	
FBB-Early	53		
FBP	2397	256	16
FBP-Early	117		
FDG	2162	134	2518
MK6240	119	612	
NAV4694	78	168	
PI2620	42	258	
PIB	58	1171	3981

RESULTS

MODEL PERFORMANCE

- 10-class model showed strong performance: Macro-F1 **0.91-0.93**, Average precision (AP) **0.79-0.99**
- 3-class model showed excellent performance: Macro-F1 ~0.99 (not shown in **Fig. 1**).

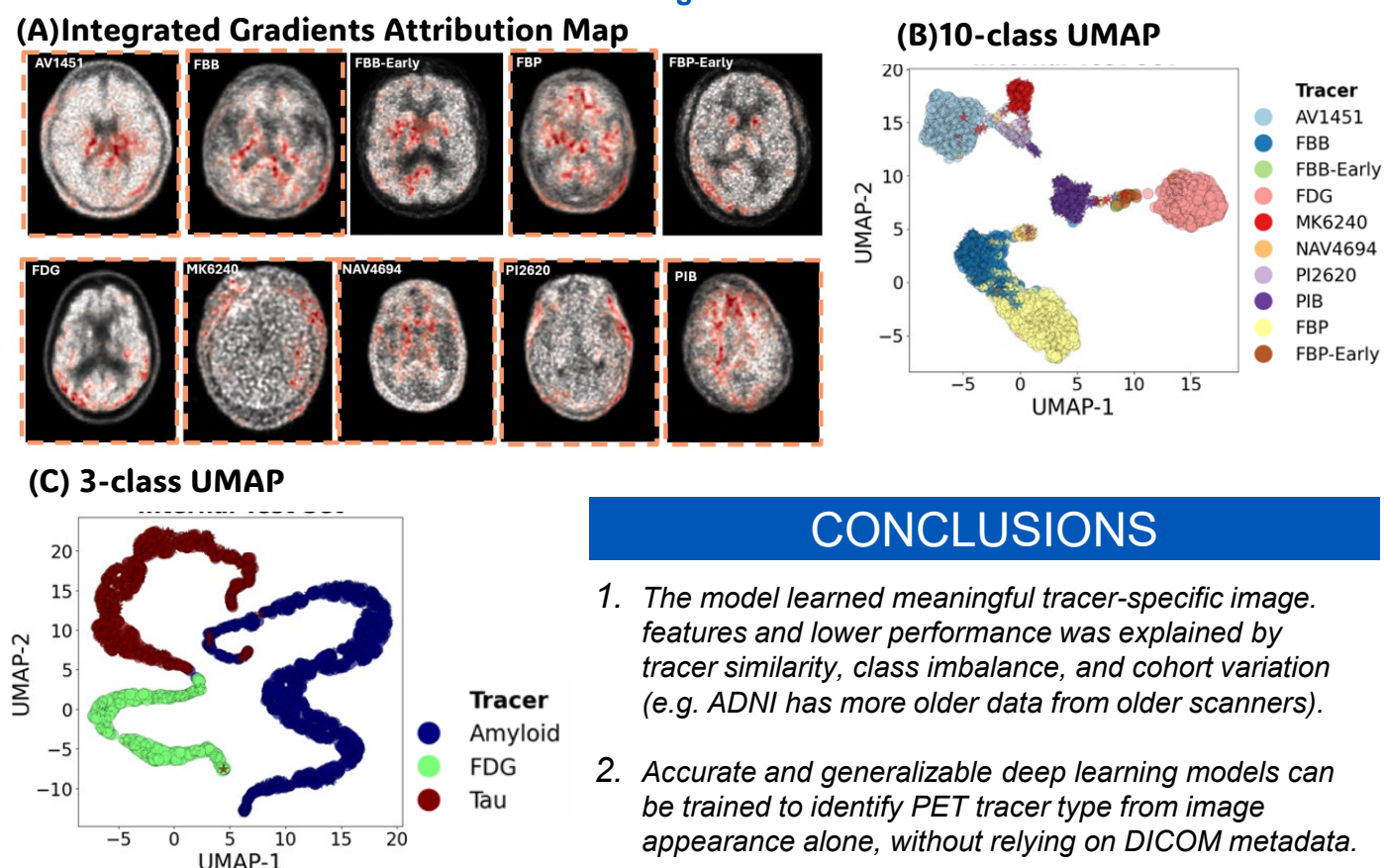
FIGURE 1: Performance of the 10-class model on the 20% internal test set



MODEL EXPLANATION

- Tracer binding properties were captured by the models (**Fig. 2**).

FIGURE 2: Model Attribution & UMAP embedding visualizations on internal test set



CONCLUSIONS

1. The model learned meaningful tracer-specific image features and lower performance was explained by tracer similarity, class imbalance, and cohort variation (e.g. ADNI has more older data from older scanners).
2. Accurate and generalizable deep learning models can be trained to identify PET tracer type from image appearance alone, without relying on DICOM metadata.