Automated detection and quantification of neuroendocrine tumors on ^{68}Ga-DOTATATE PET/CT images using a U-net ensemble method

Amy J Weisman[1], Ojaswita Lokre[1], Brayden Schott[2], Victor Santoro-Fernandes[2], Robert Jeraj[1, 2], Timothy G Perk[1], Steve Y Cho[2], Scott B Perlman[2]
[1]AIQ Solutions, [2]University of Wisconsin-Madison

BACKGROUND

^{68}Ga-DOTATATE PET/CT is a promising imaging tool used to detect and monitor disease in patients with advanced gastroenteropancreatic neuroendocrine tumors (GEP-NETs). Patients often present with a high disease burden, sometimes with tens to hundreds of lesions, making comprehensive lesion-wise assessment clinically infeasible.

Here, we implement automated, convolutional neural network-based (CNN) methods for automatic individual lesion detection and disease burden assessment.

MATERIALS AND METHODS

- Baseline & follow-up ^{68}Ga-DOTATATE PET/CT images from 59 patients with GEP-NETs undergoing theranostic ^{177}Lu-DOTATATE (Lutathera) therapy were retrospectively analyzed (116 total scans, 1-7 per patient).
- Two different CNNs (Figure 1) were trained separately using 5-fold cross validation
- Ensembles of the two CNNS were created with the intersection and union of the outputs

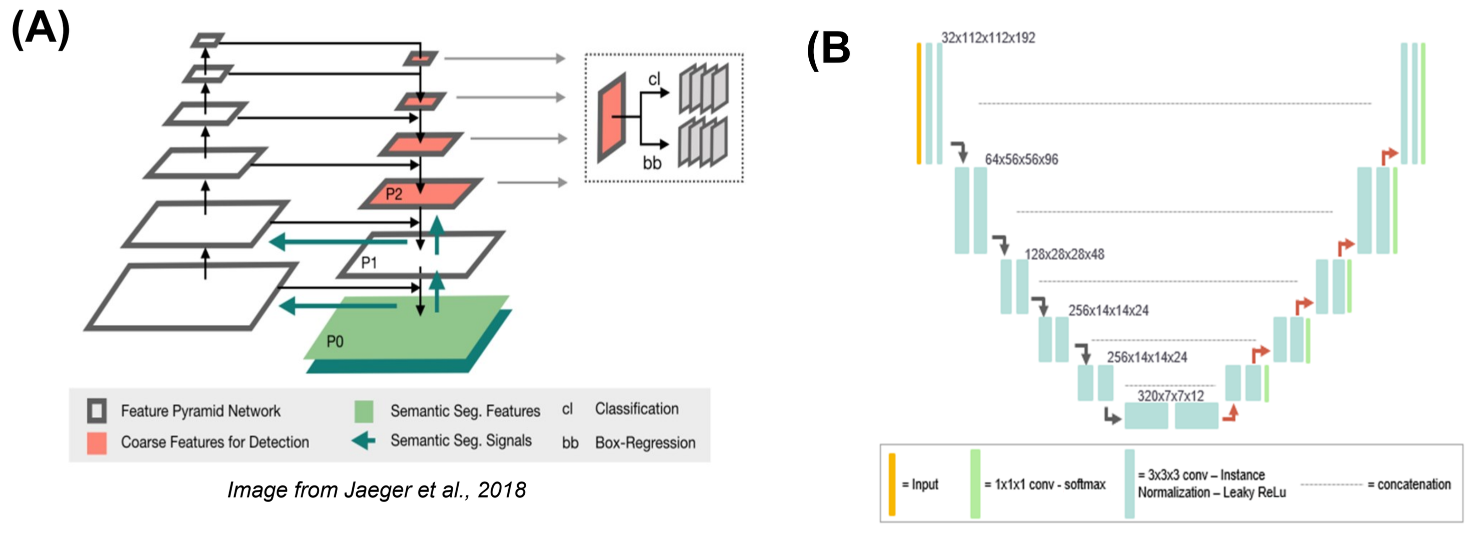

Figure 1: Architecture for the (A) Retina U-net [1] and (B) nn-Unet [2] model used in this study.

Performance Quantification

- Lesion detection performance quantified with lesion detection sensitivity and number of false positives (FPs) per patient
- Quantification performance of SUV_{max}, SUV_{mean}, SUV_{total}, and total volume assessed with Pearson's correlation coefficient (R).

RESULTS

- A total of 2,634 lesions from the 59 baseline PET/CT images were contoured by a radiographer (range: 1-239 lesions per scan).
- Median sensitivity was 87% with 2 FPs/patient for nnU-net, and 92% sensitivity with 5 FPs/patient for retina U-net (**Figure 2**)
- The union ensemble achieved 93% sensitivity with 5 FPs/patient, and intersection achieved 82% sensitivity with 2 FPs/patient (**Figure 2**, examples of intersection method **Figure 3**)

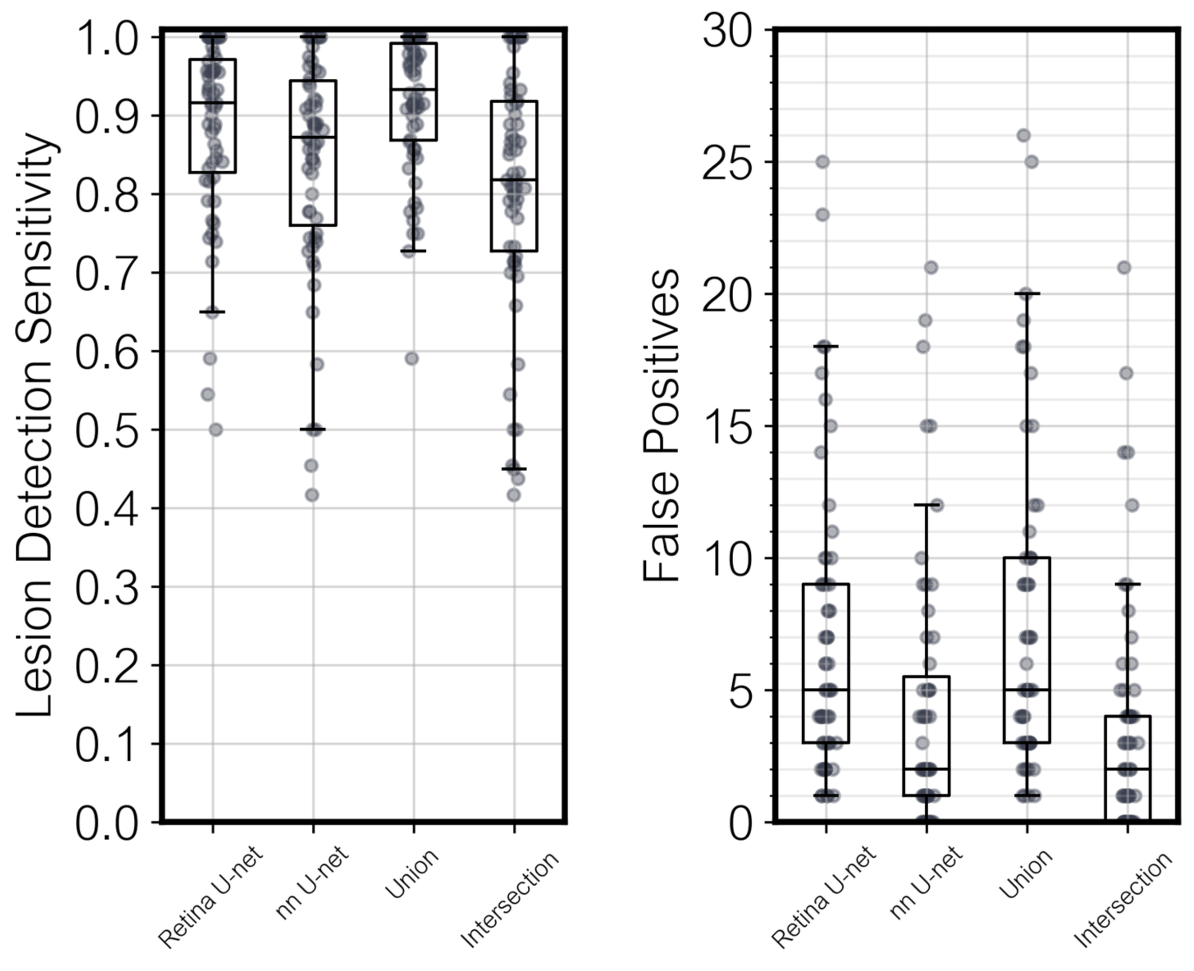

Figure 2: Detection performance of all methods

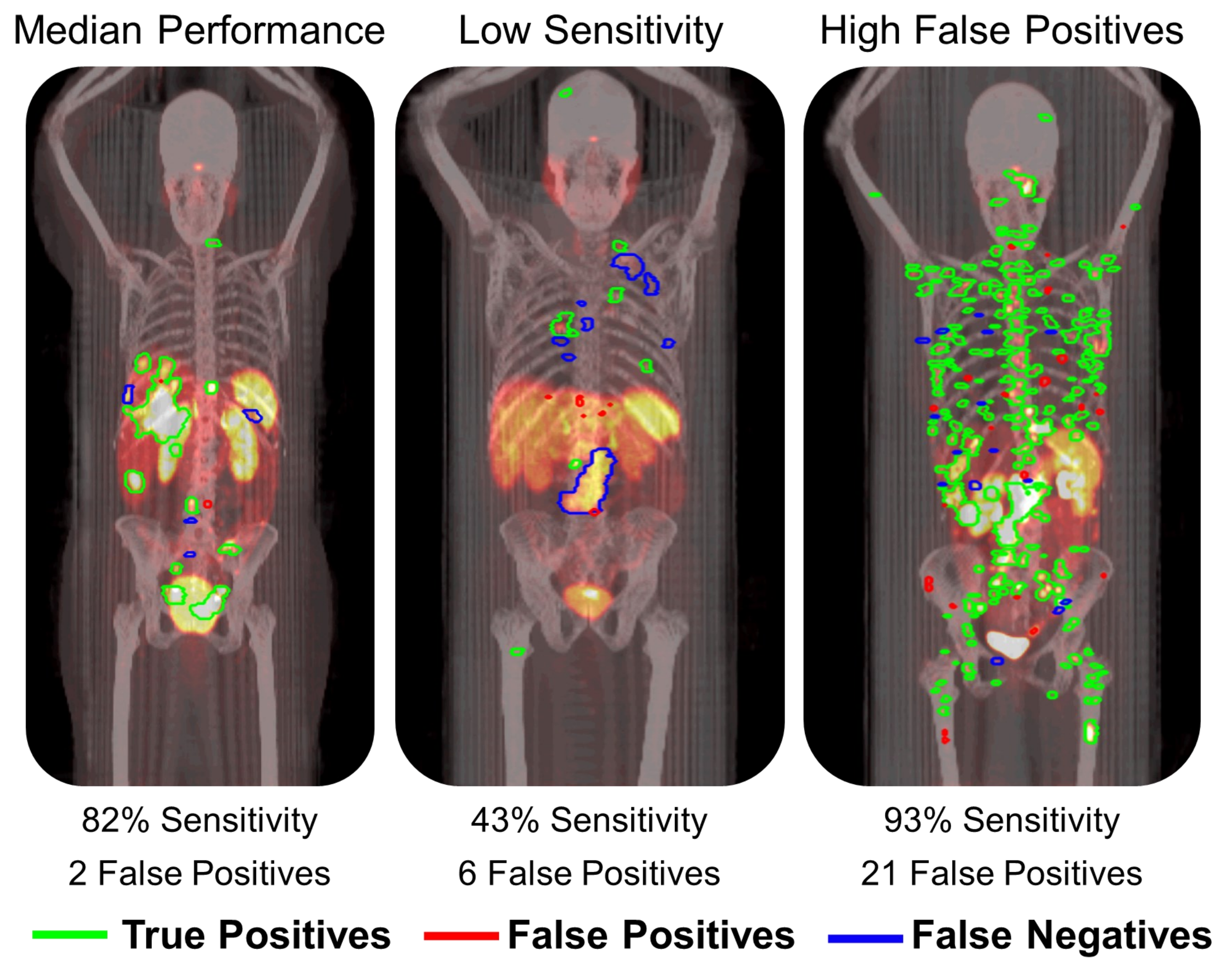

Figure 3: Intersection method performance in 3 patients

- For baseline patient-level quantification, the ensemble intersection method achieved the best overall quantification performance, with Pearson correlation coefficients of R=0.95 for SUV_{mean}, R=0.97 for SUV_{total}, and R=0.92 for total volume (**Figure 4**).
- Patient-level SUV_{max} was correctly captured in 49 of 59 scans

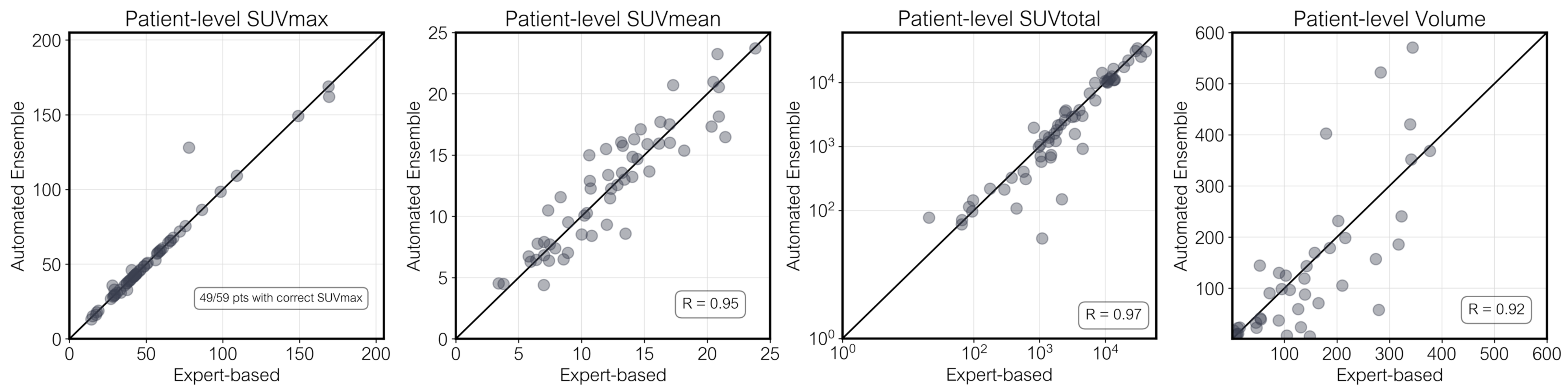

Figure 4: Patient-level quantification performance for the intersection method

KEY FINDINGS

An ensemble of two U-net based CNNs trained for lesion detection achieved excellent performance for quantifying patient-level PET imaging metrics.

Despite a lower sensitivity, **the method with the fewest false positives achieved the best quantification performance**, indicating the majority of missed lesions have low uptake and represent a small fraction of the total disease burden.

REFERENCES

[1] Jaeger, P. F. *et al.* Retina U-Net: Embarrassingly Simple Exploitation of Segmentation Supervision for Medical Object Detection. *Proc. Mach. Learn. Res. NeurIPS 2019* 1–12 (2018).

[2] Isensee, F. *et al.* nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, (2020).

DISCLOSURES

Authors AJW, OJ, and TGP are employed by AIQ Solutions. RJ is a cofounder and board member of AIQ Solutions. SBP is a consultant for AIQ Solutions. Remaining authors have no relevant relationships to disclose.

CONTACT INFORMATION

Amy.Weisman@aiq-solutions.com
@aj_weisman