Transformer-based retinal nerve fiber (RNFL) segmentation in optical coherence tomography (OCT)

Aditya Nair, MS; Rogerio Guimaraes, BS; Homayoun Bagherinia, PhD; Niranchana Manivannan, PhD

Carl Zeiss Meditec, Inc., Dublin, CA, USA

Poster #2055 - F0044

PURPOSE

- RNFL layer analysis is an important diagnostic tool for glaucoma.
- In this work, we develop an algorithm to predict RNFL layer boundaries using a novel Transformer-based deep-learning architecture to improve the performance (accuracy, runtime and memory) of RNFL segmentation.
- The purpose of this algorithm is to develop a generalized and adaptable algorithm for various scan field of views (FOV).

METHODS

- Training data consists of 645 volumes of 2mm x 6mm x 6mm CIRRUS™ HD-OCT 4000 (ZEISS, Dublin, CA) data.
- We used a knowledge-based algorithm to generate the ground truth segmentation for this problem.
- The ground truth consisted of 2-layer position values each corresponding to every A-scan position as well as 2 confidence of segmentation values.
- We trained a transformer encoder [1,2] that receives an OCT B-scan image patch of 6 mm and outputs the ILM and RNFL segmentations and their confidence values.
- The input image is passed through a linear projection layer to compress the image in the A-scan direction, then sinusoidal positional encoding is added to retain positional information.
- Transformer outputs are passed through fully-connected and sigmoid activation layer to output layer position and confidence values. The network setup is shown in Figure 1.
- We also made use of knowledge-based algorithm's confidence value to weigh the loss function (hence learning from noisy training data in a semi-supervised manner).
- We measured the performance of algorithm on manual segmentation for inferior, superior and central B-scans for held out 35 subjects.

Transformers-based method provides a robust direct OCT image to RNFL layer boundary segmentation

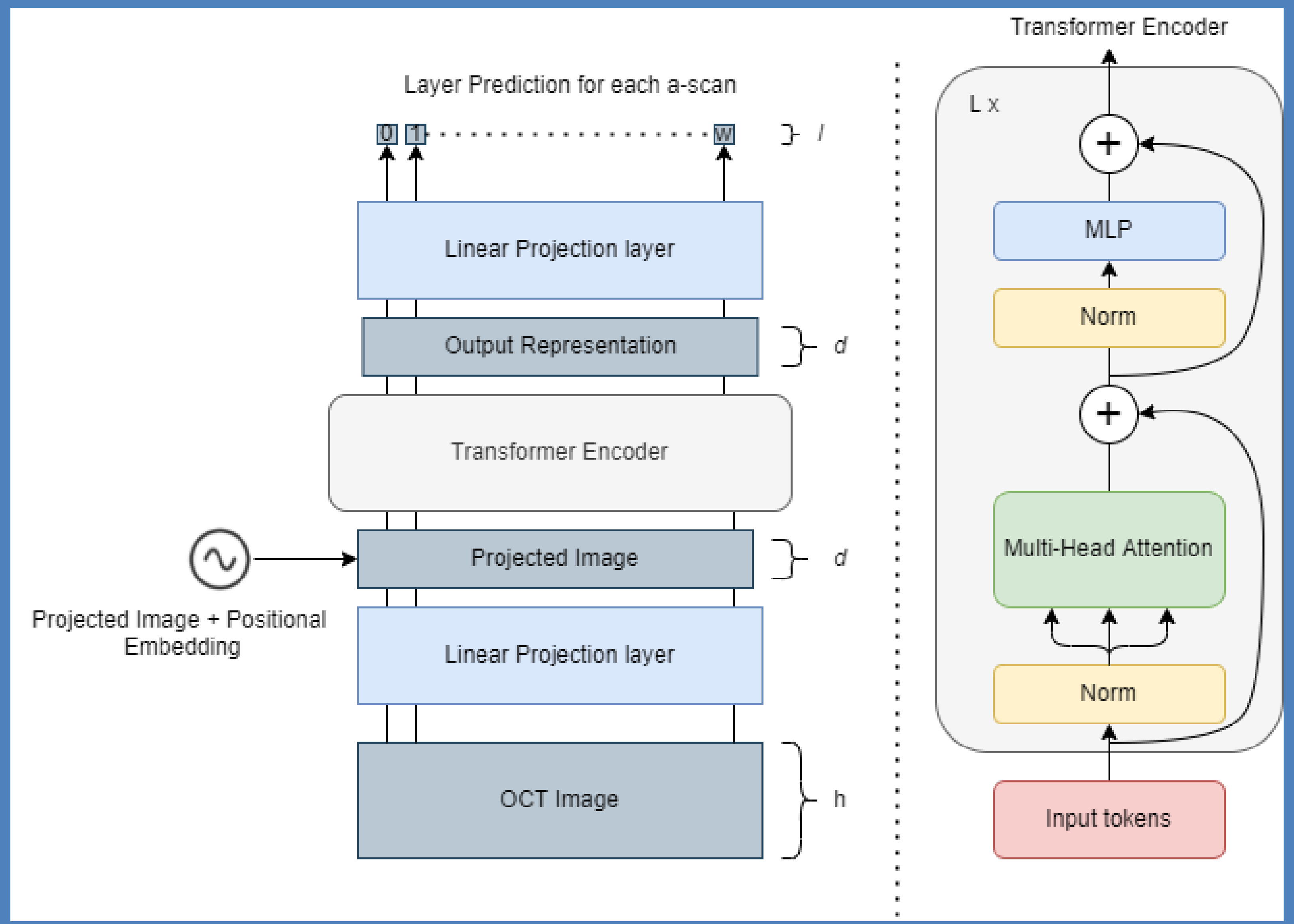

Figure 1: Transformer-based architecture for layer segmentation

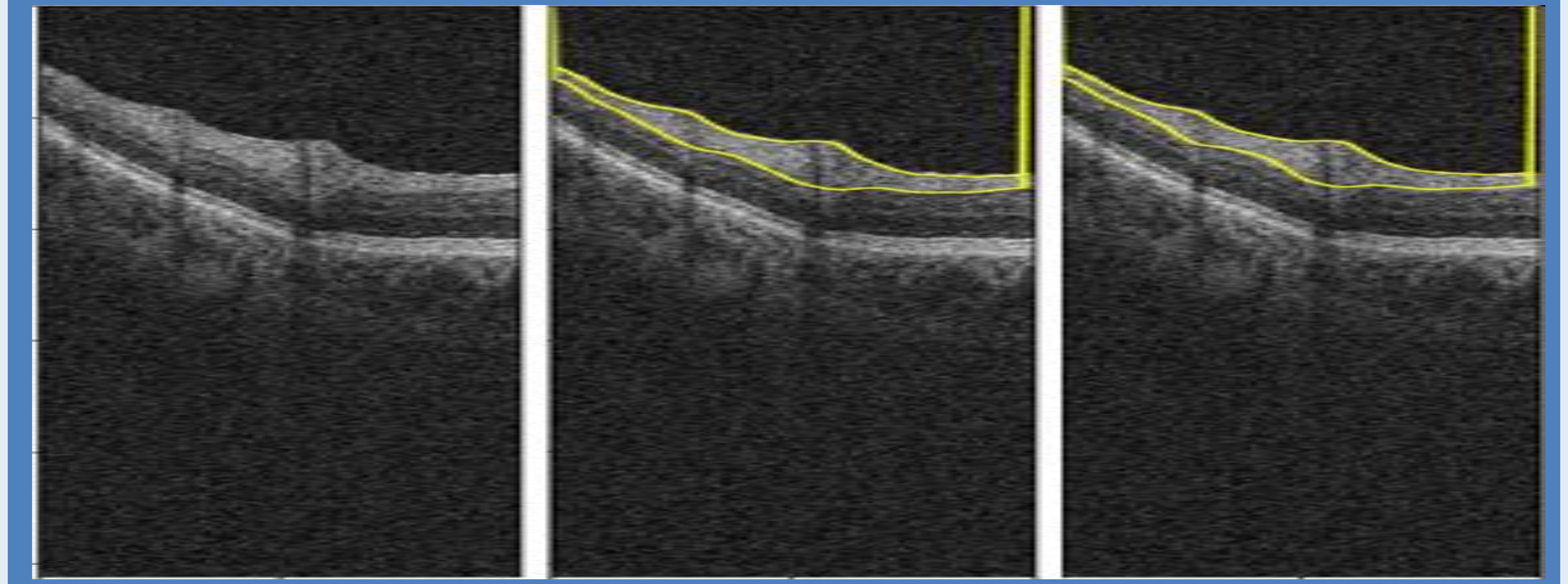

Figure 2: Test OCT B-scan (left), RNFL segmentation using knowledge-based algorithm (middle) and transformer (right)

Email: homayoun.bagherinia@zeiss.com

Disclosures: AN (E), RG (C), HB (E), NM (E) – Carl Zeiss Meditec, Inc.

RESULTS

- The Transformer-based OCT segmentation provided a better agreement with the ground truth (GT) than the knowledge-based algorithm as shown in Figure 3.
- The runtime of this algorithm was 2.12 secs vs the 2.90 secs for the knowledge-based algorithm on Intel® Core i7-9850H CPU@2.60 GHz with a NVIDIA GeForce MX150 GPU.
- The Transformer method also used significantly lower GPU memory than standard methods (< 2 Gigabytes).

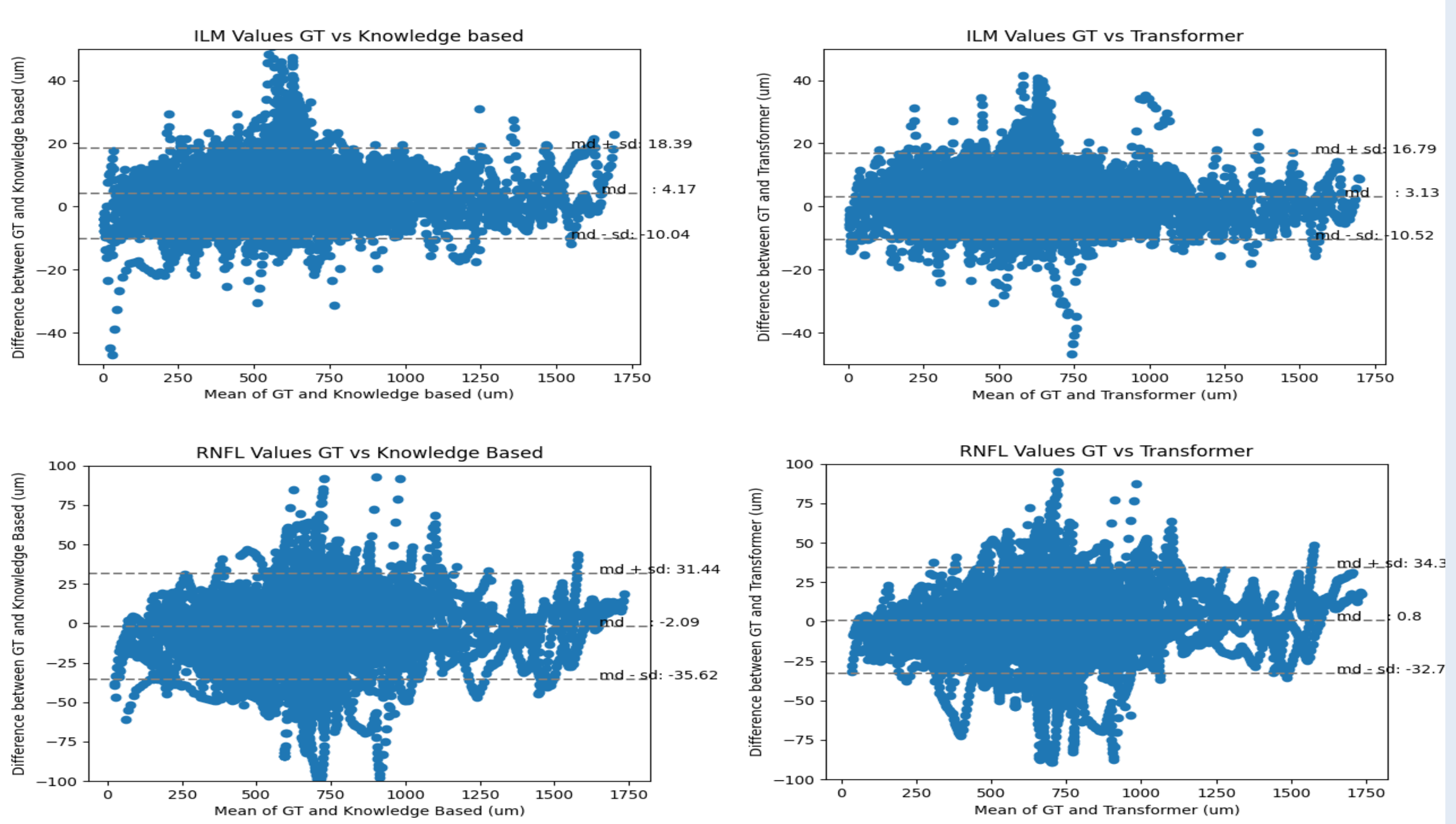

Figure 3: Transformer is closer to manual segmentation (GT) as seen in Bland-Altman plots

CONCLUSIONS

- We developed a Transformer-based deep learning method to segment the RNFL layer in OCT B-scans (Figure 2).
- We showed that our method using Transformer performs better than the knowledge-based approach on speed, efficiency, and accuracy while being adaptable to different scan FOV.

References

[1] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems* 30 (2017)

[2] Devlin, Jacob, et al. "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805* (2018)