Early prediction of multiple sclerosis using scanning laser ophthalmoscopy (SLO) video sequence data with a Deep Learning (DL) based approach

Joe Xing PhD[1], Ari Green MD[2], Lon Dowell[1], Christy K. Sheehy PhD[1]

1. C. Light Technologies, Inc., San Jose and Boston, United States; 2. University of California San Francisco, San Francisco, United States

INTRODUCTIONS

Multiple Sclerosis (MS) is a chronic immune-mediated inflammatory disease (IMID) of the central nervous system (CNS). Early identification of MS is crucial to delay disease progression and improve patient outcomes. In this work, we utilize scanning laser ophthalmoscope (TSLO) video sequence to predict MS disease. While traditional Machine Learning (ML) methods have demonstrated a strong predictive power [Mauro F. Pinto et al., 2020], we propose the use of a novel data representation of "Retinal Encoding" and a DL based model that has much higher learning capacity in capturing different subtle eye motion patterns.

THE DATA

Our approaches were tested using a 250-subject MS/control database. Patients with Expanded Disability Status Scale (EDSS)< 4 are compared to healthy subjects. Both retinal images and frequency and spatial patterns of the eye motion are combined to construct a hybrid image [Joe Xing et al. 2017], denoted as "retinal coding".

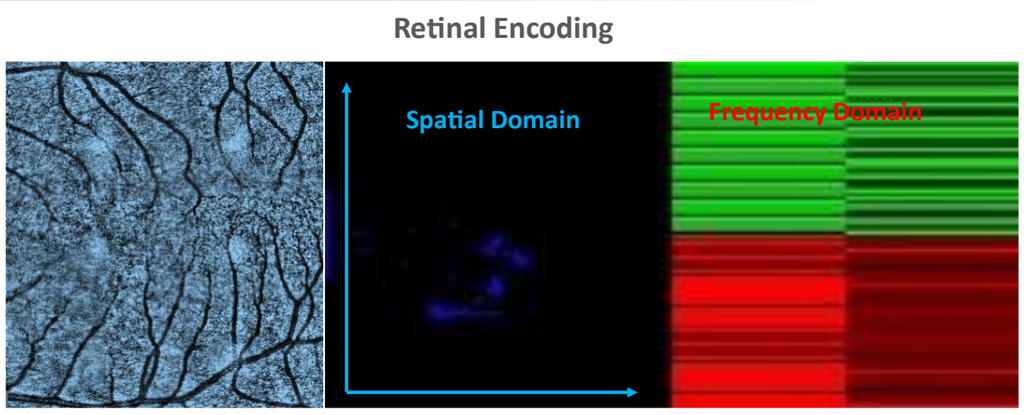

Figure 1 An innovative approach to represent sequential retinal image data, "Retinal Encoding"

Spatial eye motion patterns, temporal frequency information, blink counts, so on and so forth, as well as other explicit and implicit features on the image itself are combined to compose the retinal encoding image.

THE MODEL

The retinal encoding image is directly fed to the DL-based mode consisting of convolutional neural network for image embedding as well as a recurrent neural network to model temporal correlations.

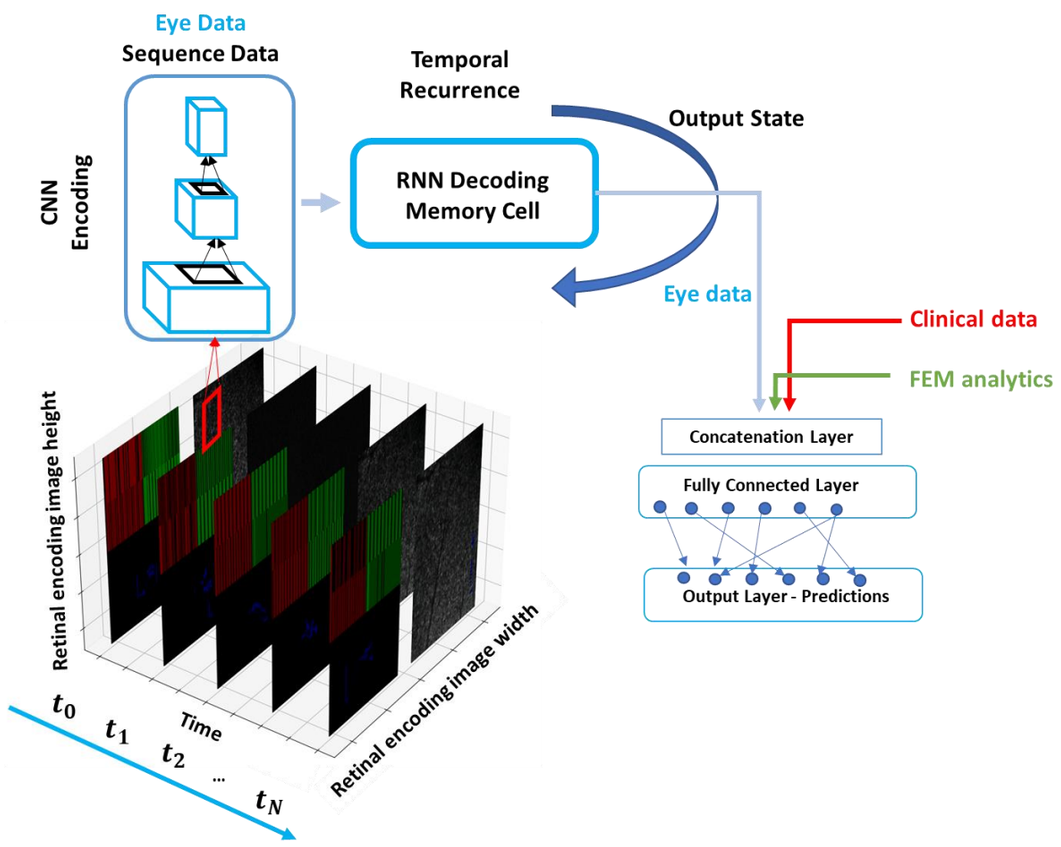

Figure 2 DL based algorithmic system to model SLO video sequences.

Retinal encoding images are fed into convolutional neural network to extract the image embedding vectors. Time-series of image embedding vectors are then input to recurrent neural network to lean the temporal representation. A final concatenation layer is used to incorporate clinical data such as demographics and Fixational Eye Motion analytics such as saccade amplitude, speed, etc., for prediction of the disease.

RESULTS

The trained model is used to predict the early MS disease versus healthy controls on validation dataset. The results are benchmarked using Receiver Operating Characteristics (ROC) and Precision-Recall (PR) curve, metrics. We showcase prediction results when only eye data is used versus both eye and clinical data, such as demographics, Timed 25-Foot Walk (T25-FW), are used.

Validation dataset	Precision	Recall	F-1 Score
Class 0 (Healthy)	0.89	0.90	0.89
Class 1 (Disease)	0.90	0.89	0.89
AUC - ROC curve		0.92	
AUC - PR curve		0.92	

Table 1 Tabulated results of the binary classifications

Class 0 and class 1 indicates health control and MS disease events. True positive, false positive, precision and recall are computed by comparing DL model output with annotation data.

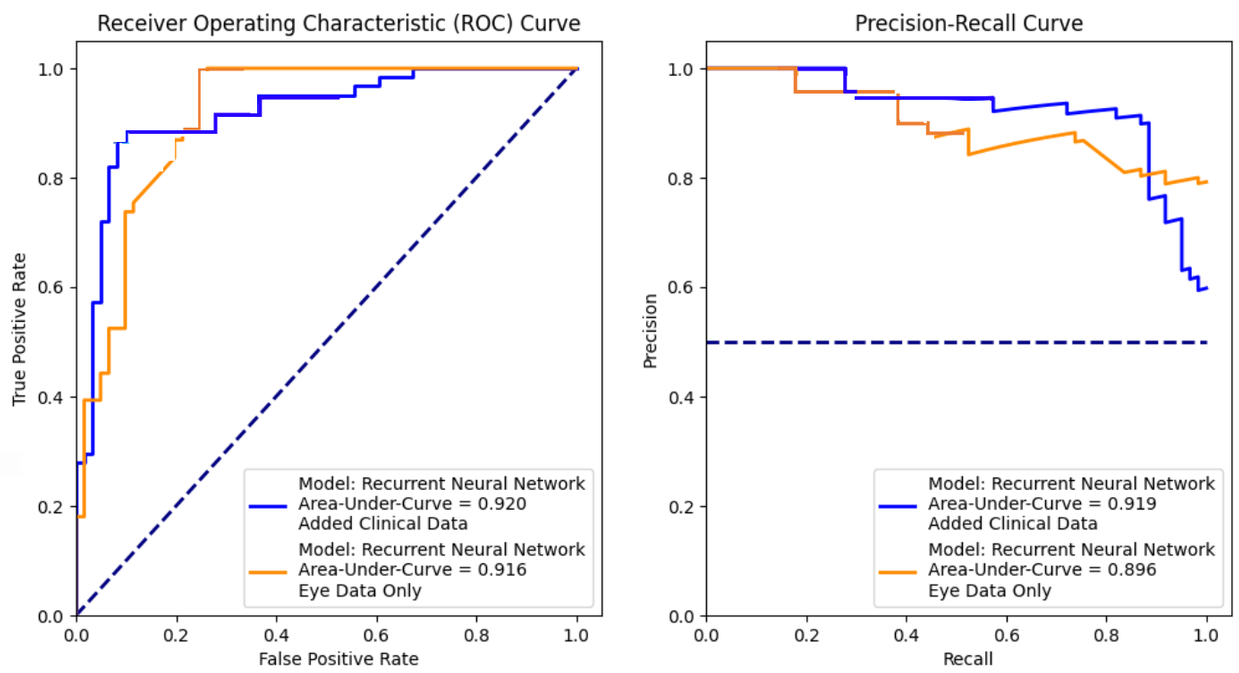

Figure 3 Preliminary results of using DL-based models to predict MS disease

Both the ROC curve for true positive rate versus false positive rate and the PR curve for precision versus recall are examined to evaluate the predictions. The orange curves stand for when the input data only contains eye data, while the blue curves are for using both eye and clinical data as input.

CONCLUSION

Preliminary results were measured using Area-under-Curve (AUC) of the ROC curve, sensitivity, and specificity, as well as an F-1 score. We observe an AUC of 0.920, sensitivity of 0.90, specificity of 0.89 and F-1 score of 0.89 to distinguish MS from controls. The results demonstrate the possibility of predicting early-stage MS and understanding disease's dynamics. Such end-to-end model could be generalizable and trained on other disease states.

ACKNOWLEDGMENTS

The investigating team is appreciative for the support of our study participants.

REFERENCES

Mauro F. Pinto, et. al. 2020, Prediction of disease progression and outcomes in multiple sclerosis with machine learning, Sci Rep 10, 21038 (2020), https://doi.org/10.1038/s41598-020-78212-6.

Joe Xing, et. al. 2017, Modeling of the Latent Embedding of Music using Deep Neural Network, arXiv:1705.05229, 2017 https://arxiv.org/abs/1705.05229.

Conflicts of interest: Joe Xing, C. Light Technologies, Inc. (E)(O)(P); Ari Green, UC San Francisco (F); Lon Dowell , C. Light Technologies, Inc. (E); Christy Sheehy, C. Light Technologies, Inc. (E)(O)(P), UC Berkeley (P), UC San Francisco (P)