2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 978-1-7281-1179-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/EMBC46164.2021.9629856

Towards Stroke Biomarkers on Fundus Retinal Imaging: A Comparison Between Vasculature Embeddings and General Purpose Convolutional Neural Networks

Ivan Coronado, Rania Abdelkhaleq, Juntao Yan, Sergio Salazar Marioni, Amanda Jagolino-Cole, Roomasa Channa, Samiksha Pachade, Sunil A. Sheth, Luca Giancardo**

***Abstract*— Fundus Retinal imaging is an easy-to-acquire modality typically used for monitoring eye health. Current evidence indicates that the retina, and its vasculature in particular, is associated with other disease processes making it an ideal candidate for biomarker discovery. The development of these biomarkers has typically relied on predefined measurements, which makes the development process slow. Recently, representation learning algorithms such as general purpose convolutional neural networks or vasculature embeddings have been proposed as an approach to learn imaging biomarkers directly from the data, hence greatly speeding up their discovery. In this work, we compare and contrast different state-of-the-art retina biomarker discovery methods to identify signs of past stroke in the retinas of a curated patient cohort of 2,472 subjects from the UK Biobank dataset. We investigate two convolutional neural networks previously used in retina biomarker discovery and directly trained on the stroke outcome, and an extension of the vasculature embedding approach which infers its feature representation from the vasculature and combines the information of retinal images from both eyes.**

In our experiments, we show that the pipeline based on vasculature embeddings has comparable or better performance than other methods with a much more compact feature representation and ease of training.

***Clinical Relevance*—This study compares and contrasts three retinal biomarker discovery strategies, using a curated dataset of subject evidence, for the analysis of the retina as a proxy in the assessment of clinical outcomes, such as stroke risk.**

I. Introduction

Annually, 15 million people worldwide suffer from stroke; of these, 10 million die or are left permanently disabled [1]. In the United States, a stroke occurs every 40 seconds, and it is the primary cause of long-term disability. In addition to clinical assessment, the primary diagnostic modality to assess stroke is non-contrast head computed tomography (CT); additional modalities might be acquired for further assessment.

Fundus retinal imaging is a non-invasive visualization of microcirculation and has been indicated as a potential marker of cerebrovascular diseases including stroke [2]. Close embryologic connection between the retina and the brain suggest that neurological disease processes have close homology with retinal tissue changes [3]. Retinal artery or retinal vein occlusions and stroke are acute manifestations of vascular disease in the retina and the brain. Simultaneous blood clots in retinal and brain circulations are common, most likely due to a shared cardiac or large vessel source [2].

Previous work on measurement of retinal characteristics has often relied on predefined measurements such as fractal analysis, vessel width, vessel tortuosity, artery to vein ratio [4]–[7]. However, these predefined measurements, while demonstrating statistical associations, cannot be easily tuned to specific clinical outcomes, which limits their use as biomarkers especially when studying new disease types. In recent years, representation learning algorithms such as general purpose convolutional neural networks (CNNs) have been proposed as an approach to learn imaging biomarkers directly from the data, drastically increasing the speed of development of disease specific biomarkers. These approaches are not a panacea, as they require a non-negligible amount of training data and at the same time they need to be validated with data that controls for the many confounding factors that might have led to "shortcut learning" [8], a common issue in medical imaging and deep learning, that would lead non-generalizable biomarkers.

Poplin et al. [9] used an Inception-V3 CNN trained on two large public datasets (120,090 images from UK Biobank and 237,233 images from EyePACS) for predicting cardiovascular risk factors using retinal imaging achieving AUC=0.97 in the prediction of gender and AUC=0.70 for major adverse cardiovascular events. Lim et al. [10] focused on a similar problem as ours achieving AUC=0.98, however, as highlighted by the authors themselves, their dataset suffered from significant confounding factors including environment-differentiating features that helped to distinguish between stroke and control images that could have led to "shortcut learning" [8]. Giancardo et al. [11] proposed a CNN-based vasculature embeddings approach. Instead of being trained directly on an outcome, it first generates an intermediate compact vector representation from the vasculature. However, their approach was only tested for image retrieval and diabetic retinopathy classification.

In this work, we present an extension of the feature embedding method that combines information from the left and right eye coupled with a tree-based gradient boosting algorithm. We compare this approach with [9] and [10] on the challenging stroke biomarker discovery problem. To do so, we

*equal contribution.

I. C., J. Y, and L. G. are with the Center for Precision Health, School of Biomedical Informatics, University of Texas Health Science Center at Houston (UTHealth), TX, USA

R. A, S. S. M., A. J. C., and S. S are with the McGovern Medical School, UTHealth, Houston, TX, USA

R. C. is with the Department of Ophthalmology and Visual Sciences, University of Wisconsin-Madison, WI, USA

S. P. is with the Center of Excellence in Signal and Image Processing, Shri Guru Gobind Singhji Institute of Engineering and Technology, Nanded 431606, India

have created a curated cohort of patients with stroke history and suitable controls from the UK Biobank biomedical dataset [12]. The whole dataset went through a quality assessment algorithm and our selection of patients was aimed at controlling confounders and comorbidities present in data. Fundus retinal imaging was obtained for each patient and used to optimize a set of techniques on the prediction of stroke.

Finally, in order to facilitate retinal biomarker research discovery, we will return the precomputed vasculature embeddings described in this paper to the UK Biobank organizers.

II. Methods

A. Study Dataset and Patient Matching

Fundus retinal images were obtained from the UK Biobank, a large-scale biomedical database and research resource that includes fundus images from over 80,000 participants. Based on ICD10 codes, we identified patients diagnosed with stroke, as well as controls [13]. Patients were matched based on age, gender, ethnicity, and cardiovascular risk factors including diabetes, hypertension, obesity. Our selection of stroke and control patients was performed to control for potential confounders in the analysis. A nearest neighbors' algorithm was used to match each stroke patient meeting inclusion criteria with the five closest control patients in terms of confounders and comorbidities.

Images from the identified patients were assessed with an automatic image quality algorithm based on an extension of Elliptical Local Vasculature Density Features (ELVD) [14] to determine inclusion in this study. This algorithm assesses retinal image quality based on the morphology of vascular structure. Vessel segmentations required by the algorithm were obtained through the segmentation algorithm employed to obtain the vasculature embeddings. To extract local vessel density, we employed a rectangular window as opposed to the polar window in the original implementation.

Only patients with imaging regarded as good quality by the ELVD algorithm were considered. The final number of patients who had stroke and suitable controls was 412 and 2060, respectively.

Table 1. Dataset Statistics

	Whole Dataset		Age Restricted	
	Control Group	Stroke Group	Control Group	Stroke Group
Number of Subjects	2,060	412	1,001	199
Mean Age (std)	58.4 (7.5)	58.4 (7.5)	64.7 (2.5)	64.7 (2.5)
Female (%)	45%	46%	45%	47%
Congestive Heart failure (%)	21%	24%	29%	31%
Hypertension (%)	69%	58%	74%	64%
Diabetes (%)	14%	16%	14%	17%
Vascular disease (%)	21%	26%	28%	32%
Chronic Kidney Disease (%)	1%	3%	2%	4%
Myocardial infarction (%)	19%	23%	26%	30%
Smoking (%)	15%	17%	10%	12%
Obesity (%)	7%	9%	6%	11%
Dyslipidemia (%)	36%	39%	42%	43%
Atrial Fibrillation (%)	12%	15%	17%	23%

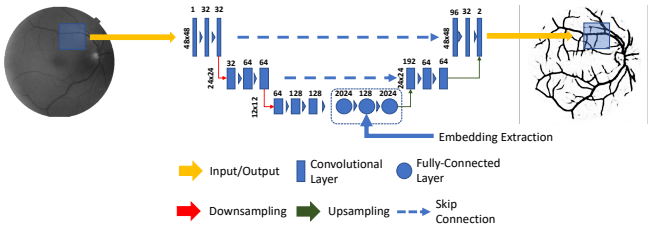

Figure 1. Vasculature Embeddings Learning Algorithm

C. Vasculature Embeddings & LightGBM

A compressed vector representative of retinal vasculature was obtained using an extension of the technique reported by Giancardo et al. [11]. In summary, this approach uses a U-shaped neural network but with the convolutional layers at the last level of abstraction replaced by fully connected layers. The full description of the network is available in the original paper. The algorithm was trained end-to-end to segment retinal vasculature. By constraining the information passing through the fully connected layers, the models learn an implicit compact representation of the vasculature. Embeddings were computed only in areas where vessels were present determined by the output segmentation at the end of the model. Equation 1 shows the combination of image patches to generate the embeddings.

$$A = \left[a_{p=1}, a_{p=2} \cdots a_{p=n}\right], \quad \text{emb} = \begin{bmatrix} f_{med}(A_1) \\ f_{iqr}(A_1) \\ \vdots \\ f_{med}(A_n) \\ f_{iqr}(A_n) \end{bmatrix} \qquad [1]$$

Where n is the number of patches, $a_p \in \mathbb{R}^{128}$ representing the local vessel embedding for a single patch, f_{med} represents the median and f_{iqr} the interquartile range for emb $\in \mathbb{R}^{256}$ As opposed to the original approach, we only included patches in set A that had probability > 0.5 to be a vessel, this allowed us to compute comparable feature embeddings in subjects with various degree of vessels visible. The probability was computed by evaluating the vasculature segmentation automatically generated while computing the embeddings. In addition, the final embedding representing a single patient was obtained by calculating the product of the embeddings for left and right retinas as $\text{emb}_{\text{left}} \odot \text{emb}_{\text{right}} \in \mathbb{R}^{256}$

Fig. 1 graphically shows how the vasculature embeddings computation process works by learning to segment a fundus image as a pretext task. In this study, we pre-trained the embeddings on multiple publicly available datasets that included a manually drawn segmentation: DRIVE (40 images) [15], STARE (400) [16], ARIA (143) [17], CHASEDB1 (28) [18], and (45) HRF [19]. From these datasets, a total of 217 images were used for the first training phase of the embeddings, the remaining images were used to control for overfitting. Note that none of these images were part of the UK Biobank dataset nor did they have any specific information about the stroke status of the subject imaged.

The vasculature embeddings were finally tuned to the stroke outcome using the light gradient boosting method (LGBM) [20]. This algorithm uses gradient boosting decision trees along with efficient sampling of instances and effective

reduction of features. Our implementation was trained with a learning rate equal to 5e-2 and early stopping for 30 rounds if binary cross-entropy loss did not decrease for validation data. The number of leaves for each trainer was experimentally set to 21.

D. Image Preprocessing

To be processed by the models, retinal images were converted to a standard format suitable for computations. First, the images were processed to a single channel by removing the blue and red color channels leaving the green channel only. Then, the circular section of the images corresponding to the retina was identified and cropped. The resulting array was zero-padded and resized to a 512x512 resolution. Textures were normalized using a squared median filter with kernel width of 15 pixels. To match the three-color channels expected by the algorithms, the resulting single channel image was concatenated with itself three times. Signal intensities were normalized to the -128 to 128 range expected by VGG19 and the -1 to 1 range for Inception-v3.

E. Convolutional Neural Networks

To compare our technique, we applied two stroke detection methods using retinal imaging. These algorithms are based on convolutional neural networks (CNNs). One of the methods proposed by Lim et al. uses VGG19 pretrained on ImageNet as base model with a modified output layer [10]. The other method introduced by Poplin et al. has Inception-v3 as base model with architectural modifications for the prediction of cardiovascular risk factors from retinal imaging [9]. Both approaches were trained independently on the fundus imaging dataset available for this study. Convolutional layers were frozen for both models. For each model, a classifier was trained using the "bottleneck" features available at the last layer of the base model. In both cases, the classifier was integrated by two hidden layers of 1024 units each and an output layer with sigmoid activation. Each model ensemble was trained for 100 epochs, using Adam as the optimizer with the AMSGrad variation [21]. The learning rate was set to 1e-5. At the end of each epoch, the models' performance on validation data was evaluated; weights for the version that obtained lowest validation loss were saved.

F. Evaluation

To prove the effectiveness of our technique, we processed the available data with the vasculature embedding gradient boosting method and the CNNs. We evaluated stroke detection by calculating the area under the ROC curve (AuROC) for the predicted probability scores by each method versus ground truth labels. Cross validation was used to evaluate each of the algorithms considered. For the CNNs, 5-fold cross validation was conducted. For LGBM, 50-fold cross validation was used. At each fold the data was split into training, validation, and testing sets. Algorithms and evaluations presented were implemented using Python, and TensorFlow and Microsoft's LightGBM libraries.

Regarding the computation time for the algorithms, training and evaluation for LGBM with the vasculature embeddings took less than a minute. For the CNN models, training time was significantly longer spawning some hours, inference time is in the range of a few milliseconds for one patient.

III. Results

Stroke detection performance based on AuROC for gradient boosting and the CNNs is shown in Figure 2. The AuROC for LGBM is 0.626. For the VGG19 model, AuROC was 0.548. For the Inception-v3 model, AuROC was 0.499. It is observed that stroke detection by gradient boosting is superior to that of the other techniques. Figure 3 shows stroke detection for the age restricted dataset. AuROC for LGBM, VGG19, and Inception-v3 are 0.674, 0.714, and 0.512. For the restricted cohort, VGG19 performed better than the other two methods.

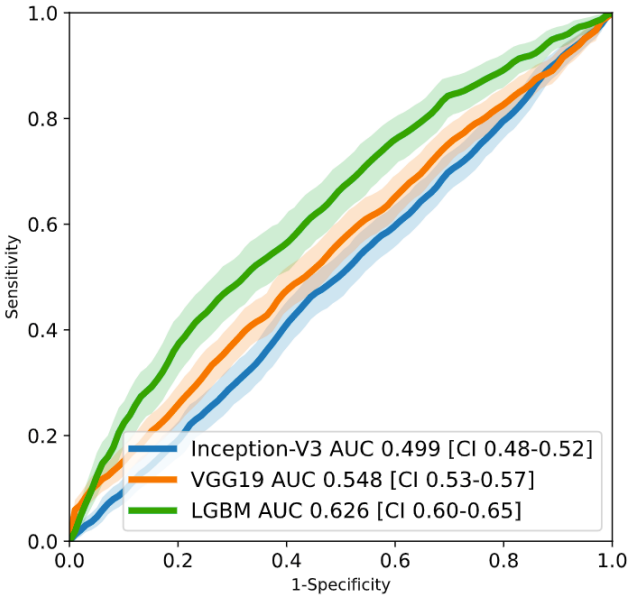

Figure 2. AuROC Curve Complete Dataset

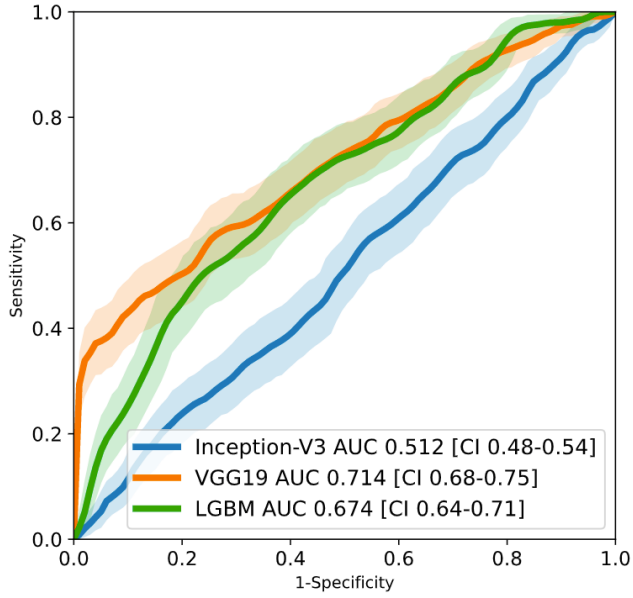

Figure 3. AuROC Curve Age Restricted Dataset

IV. Discussion and Conclusion

This study introduced a technique for detecting stroke through retinal imaging using vasculature embeddings and

gradient boosting. We evaluated this method on a large biomedical dataset and compared it to detection techniques based on convolutional neural networks. Obtained results show that our proposed methodology is fit for the detection of stroke, suggesting that the embeddings used are representative of retinal microvasculature. Compared to the CNN methods considered our methodology is efficient both in compression and computation time. Our compressed representation was a vector of length=256. In the CNN models the last max pooling layer was flattened resulting in vectors of length ~400,000 for Inception-v3 and ~130,000 for VGG19. This indicates the efficiency of our model as fewer parameters are needed to reach comparable performance. An alternative approach to be considered in future work is to replace the CNNs' last max pooling layer for a global average pooling layer. This can reduce the number of parameters in the classification model improving computational efficiency. It was observed that the accuracy in stroke detection reached by Lim et al. did not replicate in our setup, probably due to better patient matching in their approach. The VGG19 had lower performance in the larger dataset evaluated; this might have been caused by heterogeneity in the images part of the dataset. For both datasets, Inception-v3 had poor stroke detection. One likely reason for this failure, is the higher dimensionality of the Inception-v3 vector after the last max pooling layer compared to VGG19. In addition, Inception-v3 is a highly specialized model used in the recognition of real-world images. Therefore, its learned filters might not extend well to other imaging domains when few data are available. While VGG19 was also pre-trained in the same task as Inception-v3, it is a far less specialized model given fewer number of layers. Therefore, the VGG19 filters might generalize better to tasks in different image domains. We expect the general-purpose CNNs' accuracy to improve given a larger retinal imaging dataset specific to stroke patients. However, modality specific models that use proxy tasks as a pre-training step, such as vasculature embeddings, have significant advantages while developing data driven models to identify retina biomarkers with relatively small but curated dataset. In future work, we will explore such approaches for extracting representations of retina vasculature.

ACKNOWLEDGMENT

This work is supported by the Translational Research Institute through NASA Cooperative Agreement NNX16AO69A. This research has been conducted using the UK Biobank resource. Ivan Coronado is supported by a training fellowship from the Gulf Coast Consortia, on the NLM Training Program in Biomedical Informatics & Data Science (T15LM007093).

REFERENCES

[1] R. A. Grysiewicz, K. Thomas, and D. K. Pandey, "Epidemiology of Ischemic and Hemorrhagic Stroke: Incidence, Prevalence, Mortality, and Risk Factors," *Neurologic Clinics*, vol. 26, no. 4. Neurol Clin, pp. 871–895, Nov. 2008, doi: 10.1016/j.ncl.2008.07.003.

[2] H. E. Moss, "Retinal Vascular Changes are a Marker for Cerebral Vascular Diseases," *Current Neurology and Neuroscience Reports*, vol. 15, no. 7. Current Medicine Group LLC 1, Jul. 27, 2015, doi: 10.1007/s11910-015-0561-1.

[3] S. Zafar, J. McCormick, L. Giancardo, S. Saidha, A. Abraham, and R. Channa, "Retinal imaging for neurological diseases: 'a window into the brain,'" *Int. Ophthalmol. Clin.*, vol. 59, no. 1, pp. 137–154, Dec. 2019, doi: 10.1097/IIO.0000000000000261.

[4] E. Trucco, A. Giachetti, L. Ballerini, D. Relan, A. Cavinato, and T. MacGillivray, "Morphometric Measurements of the Retinal Vasculature in Fundus Images with VAMPIRE," *Biomed. Image Underst. Methods Appl.*, pp. 91–111, 2015, [Online]. Available: https://doi.org/10.1002/9781118715321.ch3.

[5] D. Fiorin and A. Ruggeri, "Computerized analysis of narrow-field ROP images for the assessment of vessel caliber and tortuosity," *Proc. EMBS*, pp. 2622– 2625, 2011, [Online]. Available: 10.1109/IEMBS.2011.6090723.

[6] X. Xu *et al.*, "Vessel boundary delineation on fundus images using graph-based approach," *IEEE Trans. Med. Imaging*, vol. 30, no. 6, pp. 1184–1191, 2011, [Online]. Available: 10.1109/TMI.2010.2103566.

[7] M. Z. C. Azemin, D. K. Kumar, T. Y. Wong, R. Kawasaki, P. Mitchell, and J. J. Wang, "Robust methodology for fractal analysis of the retinal vasculature," *IEEE Trans. Med. Imaging*, vol. 30, no. 2, pp. 243–250, 2011, [Online]. Available: 10.1109/TMI.2010.2076322.

[8] R. Geirhos *et al.*, "Shortcut Learning in Deep Neural Networks," *Nat. Mach. Intell.*, vol. 2, no. 11, pp. 665–673, Apr. 2020, doi: 10.1038/s42256-020-00257-z.

[9] R. Poplin *et al.*, "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nat. Biomed. Eng.*, vol. 2, no. 3, pp. 158–164, Mar. 2018, doi: 10.1038/s41551-018-0195-0.

[10] G. Lim *et al.*, "Feature Isolation for Hypothesis Testing in Retinal Imaging: An Ischemic Stroke Prediction Case Study," *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 01, pp. 9510–9515, 2019, doi: 10.1609/aaai.v33i01.33019510.

[11] L. Giancardo, K. Roberts, and Z. Zhao, "Representation learning for retinal vasculature embeddings," in *Cardoso M. et al. (eds) Fetal, Infant and Ophthalmic Medical Image Analysis. OMIA 2017, FIFI 2017. Lecture Notes in Computer Science*, 2017, vol. 10554, pp. 243–250, doi: 10.1007/978-3-319-67561-9_28.

[12] C. Sudlow *et al.*, "UK Biobank: An Open Access Resource for Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age," *PLOS Med.*, vol. 12, no. 3, p. e1001779, Mar. 2015, doi: 10.1371/journal.pmed.1001779.

[13] Y. Kim *et al.*, "Utilization and Availability of Advanced Imaging in Patients With Acute Ischemic Stroke," *Circ. Cardiovasc. Qual. Outcomes*, Apr. 2021, doi: 10.1161/circoutcomes.120.006989.

[14] L. Giancardo, M. D. Abràmoff, E. Chaum, T. P. Karnowski, F. Meriaudeau, and K. W. Tobin, "Elliptical local vessel density: A fast and robust quality metric for retinal images," in *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008, pp. 3534–3537, doi: 10.1109/iembs.2008.4649968.

[15] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imaging*, vol. 23, no. 4, pp. 501–509, 2004, doi: 10.1109/TMI.2004.825627.

[16] A. Hoover, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imaging*, vol. 19, no. 3, pp. 203–210, 2000, doi: 10.1109/42.845178.

[17] D. J. J. Farnell *et al.*, "Enhancement of blood vessels in digital fundus photographs via the application of multiscale line operators," *J. Franklin Inst.*, vol. 345, no. 7, pp. 748–765, Oct. 2008, doi: 10.1016/j.jfranklin.2008.04.009.

[18] M. M. Fraz *et al.*, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2538–2548, 2012, doi: 10.1109/TBME.2012.2205687.

[19] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust vessel segmentation in fundus images," *Int. J. Biomed. Imaging*, vol. 2013, 2013, doi: 10.1155/2013/154860.

[20] G. Ke *et al.*, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, pp. 3149–3157.

[21] S. J. Reddi, S. Kale, and S. Kumar, "On the Convergence of Adam and Beyond," *Int. Conf. Learn. Represent.*, 2018.