Towards an early diagnostic model of central visual impairments, based on gaze behaviour: Preliminary results

Jesse W. Grootjen[1], Alexandra Sipatchin[2], Siegfried Wahl[2,3], Tonja Machulla[1,4], Lewis Chuang[1,4,5]

[1] LMU Munich, Munich, Germany

[2] Institute for Ophthalmic Research, Tübingen, Germany

[3] Carl Zeiss Vision International GmbH, Aalen, Germany

[4] Dept. of Assistive Technologies, Faculty of Rehabilitation Sciences, TU Dortmund

[5] IfADo-Leibniz Institute for Working Environment and Human Factors TU Dortmund

Background: Central vision loss can occur with age (e.g., Age-related Macular Degeneration, AMD). In view of demographic trends in population aging (Evans et al., 2002), this is a growing concern with a negative impact on the quality of life (Slakt et al., 2005). AMD can impair real-world activities, such as cooking, that requires basic visual abilities that we take for granted, including tracking and discriminating hot and sharp objects. AMD patients often fail to notice early onset symptoms. Instead, they experience a gradual worsening of vision in one or both eyes that might be attributed to other reasons (NEI, 2019). Thus, patients rarely seek explicit ophthalmological testing before the later stages of AMD's progression. There is no known cure for AMD and the best viable remedy is to train patients to fixate with a non-central preferred retinal locus (PRL; Crossland et al., 2005). This training is ideally introduced as early as possible. The current study presents how an early diagnostic model could work based on small samples of eye movement behaviour, which might be derived from wearable eye trackers.

Methods: Our model is based on the eye movement dataset from Barraza-Bernal et al., (2018). Here, five participants with normal vision performed a variety of fine visual discrimination tasks that required visual acuity. They were subjected to a gaze-contingent scotoma that varied in size, which simulated central vision loss. We developed our model to discriminate between visual behaviour during "no scotoma" (class 1) and "early scotoma" (6°; class 2). From each dataset, we removed the first second of data after calibration. We decided on short sampling windows to emulate data-acquisition "in-the-wild". We used a sliding window approach with a window of 1 second and a step-size of 5 msec. Each input sample consisted of normalized screen coordinates and pupil size. For our modelling, we applied a 80:20 participant-wise split. We randomly selected 10% from both the training and validation data. Thus, 54% for our training data and 50% for our test data was labelled as having a scotoma present (class 2). Our proposed model was designed to minimize computing costs for viable mobile or cloud computing devices. Hence, we applied a neural network with two Long-Short-Term-Memory (LSTM) layers. Our first layer is the input layer of 1000 x 3 channels. Followed by an LSTM-layer of 64 nodes, a dropout layer of 0.2, another LSTM-layer of 32 nodes, with as final layer the output. We trained out model with a RMSprop optimizer (learning rate = 0.001), a batch size of 256. We trained our model for 50 epochs. Running the final model on a MacBook Pro 13" 2019 (i7 2.8 Ghz, 16 GB, Intel 655) takes on average 83.1 msecs (SD = 2.57 msecs, n = 100) to evaluate.

Results: Our model achieved an accuracy of 70.47% on the test data. This model has a false positive and negative rate of 0.17 and 0.13 respectively (See Figure 1).

Conclusion:

Preliminary results show that it is viable to develop a model to estimate the likelihood of central vision loss, based exclusively on implicit visual behaviour. Our proposed model is simple and can be further optimized for performance. For instance, we could add a dense layer that would receive an input from all neurons in the previous layer, dense layers provide learning features from all the combinations of the features of the

previous layer. However, this would be at the cost of additional computing power and could result in overfitting. In addition, our sampling currently uses only 10% of the available data to emulate our use-case scenario. We could consider increasing our window sizes and training our model on the complete dataset. However, this would be at the cost of computing times. Determining a viable balance between sufficient training data and computing costs will be an ongoing challenge.

The current work has limitations and more work is required prior to real-world deployment. Foremost, our model is based on a dataset that was collected in a controlled laboratory environment and based on simulated, rather than real, central visual impairment. While this dataset affords us high-quality eye tracking data at high sampling rates (i.e., 1000Hz), real world data based on wearable eye-trackers may suffer from variable illumination, data loss, low sampling rates, and more. One solution could be to schedule data-collection for short periods under optimal conditions, which could be determined for instance from the user's context (e.g., indoors). In this regard, it is promising to note that our model functioned above chance levels based on only 10 minutes of visual behaviour. Besides this, the central visual impairment simulated by Barraza-Bernal et al., (2018) may not accurately reflect the experiences of AMD patients. To begin, they used a black circle with discrete edges even though scotomas tend to be irregular in shape, with diffuse edges. Also, they used young healthy individuals as a proxy who may have exhibit fundamental differences in their visual behaviour, relative to our demographic target (i.e., elderly AMD patients). In spite of this, the current work is promising and offers a complementary or alternative solution to automatic diagnostic systems. Current systems for AMD detection may report accuracy rates above 90%. However, they all operate on the assumption that retinal images, which must be explicitly captured, are available (Burlina, 2017, Treder, 2018). Our proposed system operates on the implicit capture of visual behaviour. This is aligned with current technological trends that are increasingly integrating eye-tracking capabilities in smart wearables, from smartphones to eyewear (Valliappan, 2020). In this, we aim towards a solution wherein diagnosis could operate automatically without the need for explicit testing.

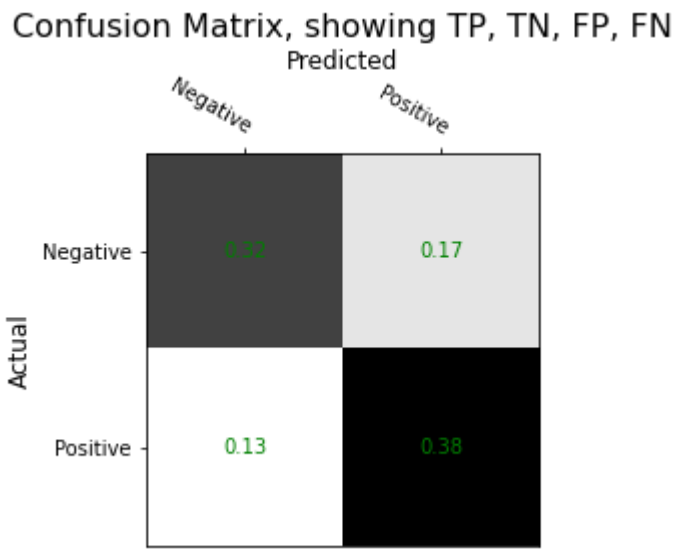

Figure 1: Confusion Matrix showing the ratio of TP (True Positives, bottom right), TN (True Negatives, top left), FP (False Positives, top right), TN (False Negatives, bottom left).

References:

Burlina P, Pacheco KD, Joshi N, Freund DE, Bressler NM. Comparing humans and deep learning performance for grading AMD: a study in using universal deep features and transfer learning for automated AMD analysis. Comput Biol Med. 2017;82:80–6.

Maria J. Barraza-Bernal, Katharina Rifai, Siegfried Wahl; The retinal locus of fixation in simulations of progressing central scotomas. Journal of Vision 2018;18(1):7. doi: https://doi.org/10.1167/18.1.7.

Valliappan, N., Dai, N., Steinberg, E. et al. Accelerating eye movement research via accurate and affordable smartphone eye tracking. Nat Commun 11, 4553 (2020). https://doi.org/10.1038/s41467-020-18360-5

Evans JR, Fletcher AE, Wormald RPL, et al Prevalence of visual impairment in people aged 75 years and older in Britain: results from the MRC trial of assessment and management of older people in the community. British Journal of Ophthalmology 2002;86:795-800.

National Eye Institute, Age-Related Macular Degeneration (AMD) Data and Statistics. July 2019. Retrieved June 2021.

Slakter JS, Stur M. Quality of life in patients with age-related macular degeneration: impact of the condition and benefits of treatment. Surv Ophthalmol. 2005 May-Jun;50(3):263-73. doi: https://doi.org/10.1016/j.survophthal.2005.02.007. PMID: 15850815.

Treder M, Lauermann JL, Eter N. Automated detection of exudative age-related macular degeneration in spectral domain optical coherence tomography using deep learning. Graefes Arch Clin Exp Ophthalmol. 2018;256(2):259–65.