Assessing Eye Tracking for Continuous Central Field Loss Monitoring

Jesse W. Grootjen
LMU Munich
Munich, Germany
jesse.grootjen@ifi.lmu.de

Alexandra Sipatchin
Institute for Ophthalmic Research
Tübingen, Germany
jesse.grootjen@ifi.lmu.de

Siegfried Wahl
Carl Zeiss Vision International GmbH
Aalen, Germany
siegfried.wahl@uni-tuebingen.de

Tonja Machulla
Chemnitz University of Technology
Chemnitz, Germany
tonja.machulla@phil.tu-chemnitz.de

Lewis L. Chuang
Chemnitz University of Technology
Chemnitz, Germany
lewis.chuang@phil.tu-chemnitz.de

Thomas Kosch
HU Berlin
Berlin, Germany
thomas.kosch@hu-berlin.de

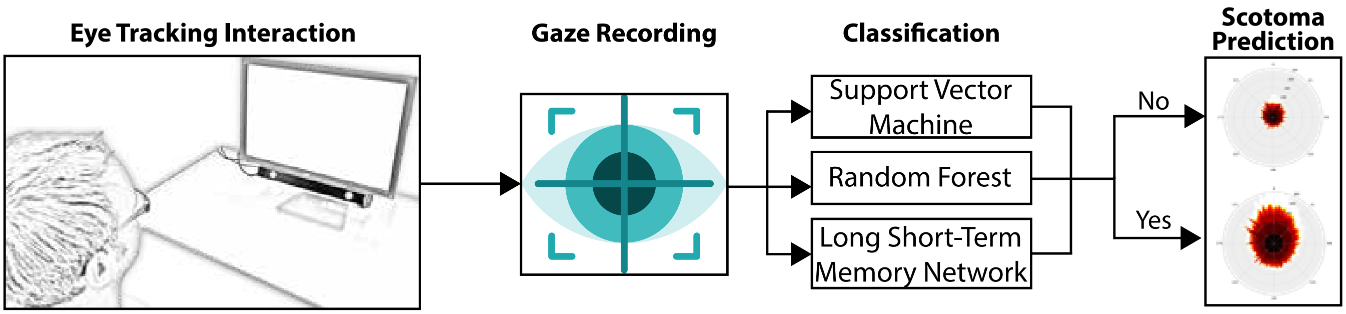


Figure 1: We investigate the computational cost and classification accuracy for predicting emerging CFL using eye-tracking data. We compare the efficiency of support vector machines, random forests, and long short-term memory networks to predict central field loss.

ABSTRACT

Eye tracking is increasingly becoming prevalent for health-related interactive systems. Eye tracking can automatically reveal the presence of Central Field Loss (CFL), a dysfunctional visual behavior requiring time-intensive medical assessments. Since CFL typically results in poor fixation stability and more frequent saccades, this work investigates the use of machine learning to estimate the likelihood of CFL based on eye-movement data. We compared random forests, support vector machines, and long-short-term memory (LSTM) neural networks for their ability to discriminate between the presence or absence of an experimentally-induced CFL. We found that the estimation accuracy increases with larger samples of eye-tracking data. However, the computational costs outweigh any increase in accuracy after classifying window sizes of 1600 msec. Here, traditional machine learning approaches outperform the LSTM neural network. We discuss implications for continuous end-user CFL monitoring and processing power to provide an outlook for gaze-based wearable health devices in human-computer interaction.

MUM '23, December 03–06, 2023, Vienna, Austria

ACM ISBN 979-8-4007-0921-0/23/12…$15.00
https://doi.org/10.1145/3626705.3627776

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**.

KEYWORDS

Central Field Loss, Eye Tracking, Machine Learning, Deep Learning, Cost-Benefit Analysis

ACM Reference Format:
Jesse W. Grootjen, Alexandra Sipatchin, Siegfried Wahl, Tonja Machulla, Lewis L. Chuang, and Thomas Kosch. 2023. Assessing Eye Tracking for Continuous Central Field Loss Monitoring. In *International Conference on Mobile and Ubiquitous Multimedia (MUM '23), December 03–06, 2023, Vienna, Austria.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3626705.3627776

1 INTRODUCTION

Central-field Loss (CFL) can occur as a result of the aging mechanisms affecting the central part of the eye. It occurs when the fovea — a 2° area of high visual acuity in the central visual field — is affected by a CFL. This can be due to age-related macular degeneration [48, 66]. Given that CFL patients are no longer able to see with their fovea, they often learn to fixate with an extrafoveal location near the CFL instead using their foveal vision [68, 73]. This acquired fixation preference is referred to as a preferred retinal locus (PRL). As a result of using a PRL, individuals with CFL move their eyes differently from those with intact vision [74]. Hence, eye movement differences can be measured using eye-tracking, a technique becoming increasingly researched for active interaction and passive sensing [23].

Inspired by previous work in the area of pervasive health-related interactive systems [75], we investigate if machine learning allows for implicit CFL detection using continuous eye movement behavior. Eye tracking is expected to be integrated into pervasive computing devices, which have already moved into mobile devices, including head-mounted displays [2] and smartphones [70]. Already now, eye-tracking is increasingly integrated into user computing devices [11, 32], together with the advances in calibration-less eye-tracking [27, 49, 62]. We envision future end-users devices to implicitly detect and diagnose an emerging CFL in the future (see Figure 1).

However, the analysis of eye movements has not been considered for pervasive CFL diagnosis for two reasons: first, it is more common to directly test for intact vision in clinical settings instead of collecting eye movement behavior and analyzing potentially diagnostic features. Unfortunately, this is a procedure patients seek out only after noticing an emerging CFL during later stages. An emerging CFL is a creeping process that patients may notice too late. Providing an early detection system could support early intervention strategies. Second, specialized training is required to interpret highly variable eye movement behavior, compared to conducting standardized clinical tests for visual acuity. This is laborious as scarcely available trained specialists are required to conduct these clinical tests. Hence, the continuous and automatic assessment of visual impairments has been a focus of previous research [81]. However, the efficiency of detecting CFL using eye movement behavior through artificial intelligence and the trade-off between estimation accuracy and computational costs remain unclear.

This work evaluates two machine learning and one deep learning approach to predict CFL by tracking eye movement behavior, including the cost-benefit between the estimation accuracy and computational costs to facilitate pervasive and continuous CFL detection in the future. Furthermore, we investigate the required costs related to the estimation accuracy for effectively designing and engineering human-machine systems that implicitly diagnose CFL from eye movements through pervasive eye gaze monitoring. Using a dataset collected from five participants experiencing a simulated CFL scenario, we compare the classification efficiency of support vector machines (SVM), random forests (RF), and long-short-term memory (LSTM) networks for different data input lengths. Our results show that we achieve an area under the curve (AUC) precision/recall of .86 using a RF classifier with a single entry of handcrafted features. In contrast, the long-short-term memory network yields an AUC precision/recall of .73 using ten seconds of data with a sample rate of 1000 Hz for existing CFL conditions. We discuss how our results can implicitly detect CFL by integrating eye tracking into general-purpose devices.

CONTRIBUTION STATEMENT

The contribution of this work is threefold:

(1) We investigate if eye tracking is a reliable metric to predict emerging CFL using machine and deep learning in a simulated scenario with five participants.
(2) We evaluate the computational cost required to make predictions, including a CFL accuracy assessment of different classification algorithms.
(3) Finally, we discuss how our results benefit clinical assessment and end-user scenarios for implicit CFL prediction.

2 RELATED WORK

Advancements in specialized hardware and software allow the capture and analysis of physiological measures to predict changes in gaze behavior reliably. It resulted in research on analyzing eye-gaze patterns to understand and support user behavior in various contexts, such as reading comprehension [31], mental health assessment [21, 22], web browsing behavior [10], and novel and playful interactions with computer systems [30] and games [33, 40]. Prior research has used eye gaze along with classical machine learning approaches, such as hidden Markov models and SVMs, to predict user intentions [29, 69]. To the best of our knowledge, there has been limited research performed that relied on machine learning to predict CFL from natural eye movement behavior.

2.1 Central-Field Loss

Healthy vision provides a clear perception of the visual scene, even though high acuity is restricted to only the central field (i.e., the fovea). It is well-established that visual acuity exponentially decreases with increasing eccentricity [60]. Hence, we constantly move our eyes to see things that are relevant to our goals. Selective central field loss impairs our ability to perform everyday tasks. In particular, CFL commonly occurs and worsens with increasing age. For instance, age-related macular degeneration (AMD) and its subvariants are the leading cause of CFL in industrialized countries and affect over 195 million people worldwide [79]. Individuals with CFL tend to get diagnosed when they have experienced progressively worse vision over a long period of time. Many individuals with CFL might be accustomed to their poor vision for such a long time that they do not recognize the benefits of the treatment of CFL. Thus, an early diagnosis of CFL could enable many to learn mitigation strategies and get available treatment. Furthermore, since CFL manifests predominantly in the older population, these individuals tend to attribute it to the aging process. Fletcher et al. [24] found that 56% of their participants were unaware of the CFL, even in those with CFL up to 30° in diameter, which is covering one-sixth of what an individual can see with two eyes. If individuals become aware of their CFL, they need to be clinically diagnosed by a specialist, with procedures taking four to ten hours. This series of examinations is tedious, invasive, and uncomfortable [18, 67].

In line with advances in artificial intelligence, machine learning is increasingly applied to health diagnostics with promising success. Nonetheless, many previous works are impractical for real-world application. Previous work for health diagnostics has focused on work involving visual imaging. An example is the use of optical coherence tomography (OCT), an imaging method used to generate pictures of the back of the eye. This approach has proven to be a viable option in the detection of various visual impairments affecting the central field [37, 64]. However, visual imaging requires an individual person to capture an image of their retina explicitly. Thus, while these machine learning approaches provide health diagnostics, they do not permit a proactive screening in everyday life.

2.2 Gaze Behavior Variations with CFL

Persons with CFL exhibit unstable fixations [36]. For instance, when asked to look at a specific part of a visual target for more than several seconds, the area over which their eyes remain momentarily stationery can be as large as 10 to 20 degrees [19, 54, 56, 77], more than an order of magnitude larger than people with intact fovea and normal vision. The high fixation instability and the implied impaired oculomotor control in people with macular disease have been suggested as the contributing factors to their impaired visual acuity [51], reading ability [19, 55, 57], and face-recognition ability [58].

Previous efforts in studying the fixation characteristics in people with CFL have focused primarily on the location of the Preferred Retinal Locus (PRL) for fixations and fixation stability. The PRL describes the slight compensation by shifting the vision to focus the target of interest on an area outside the CFL. Consequently, previous research provided compelling success in directing the PRL to a saccade target [53]. Nonetheless, Renninger and Ma-Wyatt [53] showed that the scan path of eye movements was not direct for individuals with CFL, even when the target was visible. Instead, it curved towards the target at the end of the saccade and required multiple small saccades to reach the target. In addition to being non-direct, saccades with CFL have characteristics of non-foveating saccades [78]. Such saccades have lower peak velocity and longer duration than regular saccades without CFL.

Given these differences in eye movement, it is plausible that the presence of CFL could be estimated from eye movement behavior, maybe even before the individual perceives it. This raises the question of how much eye movement data is necessary.

The collection of eye movement data presents a potential violation of personal data security. Besides health-related data, which is of interest in the current work, attributes such as age and gender [12, 46] could also be derived from eye movement data. Arguably, it can also discriminate for other privacy-sensitive attributes, including race, sexual preference, BMI, or hormonal cycles [4, 25, 39]. Collecting large amounts of data required for extracting information about the unstable fixations and the number of fixations would result in the loss of privacy of two kinds. First, an individual's identity, as the user's unique gaze pattern, allows for "fingerprinting" as it contains several bio-indicators, as mentioned previously. Second, is the inference of interests, which, when annotated by semantics, the measurement of interest in items displayed on-screen reveals political, sexual, cultural, or other lifestyle preferences [41]. While there are existing methods, such as PrivacEye [65] to preserve privacy during eye tracking, these solutions might exclude the fixation stability or the saccade frequency. This work will estimate the trade-off between computational cost and estimation accuracy of CFL with eye-tracking using small chunks of data within an LSTM-NN model.

2.3 Machine Learning Clinical Diagnosis from Gaze Data

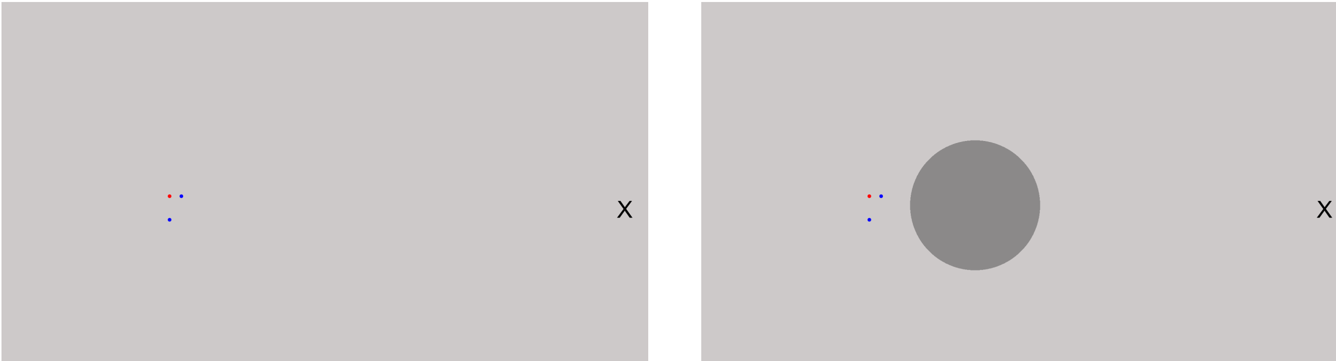
Detecting visual impairments using eye tracking has been investigated in past research [16]. However, most of these works use optical coherence tomography (OCT), an imaging method used to generate pictures of the back of the eye. Smith et al. [63] used a Gaussian Mixture Model in their process of assessing reading performance in patients with Glaucoma using eye tracking. Other use cases outside of visual impairment include but are not limited to, using eye gaze characteristics to train a machine learning model to predict the presence of dyslexia using eye tracking [52].

2.4 Relevant Aspects of Machine Learning Influencing Cost-Benefit

Given their effectiveness, Neural Networks (NNs) are increasingly applied to solving many visual computing tasks. In particular, deep learning-based solutions for gaze estimation have been popular in recent years. For example, Vora et al. [76] compared the performance of several Convolutional Neural Network (CNN) architectures: AlexNet, VGG16, ResNet50, and SqueezeNet in predicting different gaze areas. Especially Long Short-Term Memory Neural Networks (LSTM-NN) [28], a recurrent neural network (RNN) variant, is suited to process non-linear dynamic and spatiotemporal information. The LSTM network retains the cell state in the RNN and adds three gates named the input gate, out gate, and the forget gate. The cell can remember values over arbitrary time intervals, and the three gates regulate the flow of information in and out of the cell. Due to these gates, LSTM can handle long-term dependencies and is more effective for tasks with time-series data, such as eye-tracking data.

2.5 Summary

Previous work informed us that CFL is a slow process that goes unnoticed by the ones affected [79]. Typically, high resolution is confined to the central part of the vision; while there is a decrease in visual acuity, the further away sharpness is considered away from the fovea. As the disease progresses, healthy regions that can be used as PRL increase in eccentricity, affecting, in turn, visual acuity. The visual acuity, therefore, decreases while the PRL eccentricity increases [60], changing how users with CFL focus their gaze on targets of interest. Consequently, CFL leads to unstable fixations and saccades [36] that can be captured via eye tracking and used to detect the development of CFL early. In this context, machine learning has been used to detect CFL automatically with different methods and clinical settings [16, 52, 63]. In contrast to previous work, we evaluate the classification accuracy and computational costs of using Support Vector Machines (SVMs), Random Forests (RFs), and LSTM-NNs to detect CFL using eye tracking. We analyze different feature parameters that maximize the accuracy of predicting CFL and investigate different window sizes (i.e., the duration and hence the amount of data needed) to make reliable predictions about the existence of CFL. Here, we optimize the computational costs of model training and CFL prediction.



(a) (b)

Figure 2: (a): Baseline without simulated CFL. (b): Simulated gaze contingent CFL of 6° of the experiment.

3 METHODS

This section details the steps involved in the proposed approach for determining the cost-benefit trade-off and estimation accuracy of central field loss. First, we describe the data acquisition process. Next, we describe the pre-processing. This is followed by describing the model used to train and validate both the machine learning and deep learning approaches. Finally, we explain how we evaluated the cost associated with the different models and how this relates to variable window sizes for the deep learning approach.

3.1 Dataset and Experimental Description

Our experiment used eye-tracking data from Barraza-Bernal et al. [6]. We decided to use a dataset where participants had a simulated CFL to ensure the reproducibility of our results across participants. This ensures that the CFL does not vary as would with participants affected by CFL (i.e., shape and size of CFL may vary among affected participants). Using a dataset from a study where participants use a simulated visual impairment is commonly used to study phenomena in vision research e.g., Sipatchin et al. [61]. Five volunteers ($\bar{x}$ = 28.8 years) with normal or corrected-to-normal vision participated in the study. Every participant was trained to acquire a peripheral retinal locus of fixation after four training sessions using a simulated gaze-contingent CFL of six degrees, totaling between two to three hours of training per participant [5]. Eye movements were recorded using an SR Research EyeLink 1000 Plus eye tracker with a spatial resolution of 0.01° and a sampling rate of 1000 Hz [6]. The participants were asked to perform a visual discrimination task where the participant was asked to discriminate between the presence of having more red or blue dots on the screen by pressing the arrow up or down on the keyboard (see Figure 2). We used a subset of the data consisting of two CFL simulations: one simulation without CFL (i.e., the baseline) and a simulation with a simulated gaze-contingent CFL of 6°, representing CFL that can go unnoticed [24].

3.2 Machine Learning

3.2.1 Dataset Preprocessing. We identified fixations, saccades, and eye blinks using the EyeLink parsing algorithm and adopted the following cutoffs based on previous related work [71]: saccadic velocity threshold of $30^\circ/s$, a saccadic acceleration threshold of $8000^\circ/s^2$, and motion threshold of 0.1°. Eye movement data below the velocity threshold and acceleration threshold criteria were classified as a fixation. Otherwise, it was labeled a saccade eye movement. For the saccade, we collected the following features: duration, average position of the gaze in X and Y coordinates, and average pupil dilation. For the fixations, we extracted the duration of the saccade, the start and end points in x and y coordinates, the amplitude in degrees, and the peak velocity in degrees per second. For the blinks, we simply extracted the duration of the blink. All cells containing no values (e.g., NaN) were set to an arbitrary number that does not occur (-1) to enable these rows of data to be fed into the models described below.

3.2.2 Model Description. We decided to use a Support Vector Machine (SVM) and Random Forests (RFs) to evaluate the CFL prediction accuracy. CFL leads to a linear gaze deviation. SVMs and RFs are linear statistical models and, hence, suitable to predict CFL. We analyzed the eye movement features mentioned above using hyperparameter optimization for RF and SVM models. SVMs were our first choice as a model as it has been extensively used in classification problems, especially in eye-tracking research. Thus, using SVM as a baseline method allows comparisons with previous and future work. SVM maps its input vectors into a high dimensional feature space through a chosen non-linear mapping and then finds an optimal hyperplane to separate the classes with a maximal margin, which reduces the generalization error [45]. This work uses the SVM implementation in the Python library Scikit-learn [15].

RF is a classifier comprising an ensemble of randomized decision trees, which make a joint decision on the class [14]. Like SVM, RF has been used with good results in various tasks. First, we applied RF to compare SVM with at least one other popular method. In addition to its demonstrated practical usefulness, RF models tend to be robust, which is important as this work could generalize to health diagnostics, generalization capability, and intrinsic feature selection opportunities embedded in decision tree-based methods [20]. Here, we applied the RF and SVM implementation from *scikit-learn*[1] [15].

[1]www.scikit-learn.org/stable

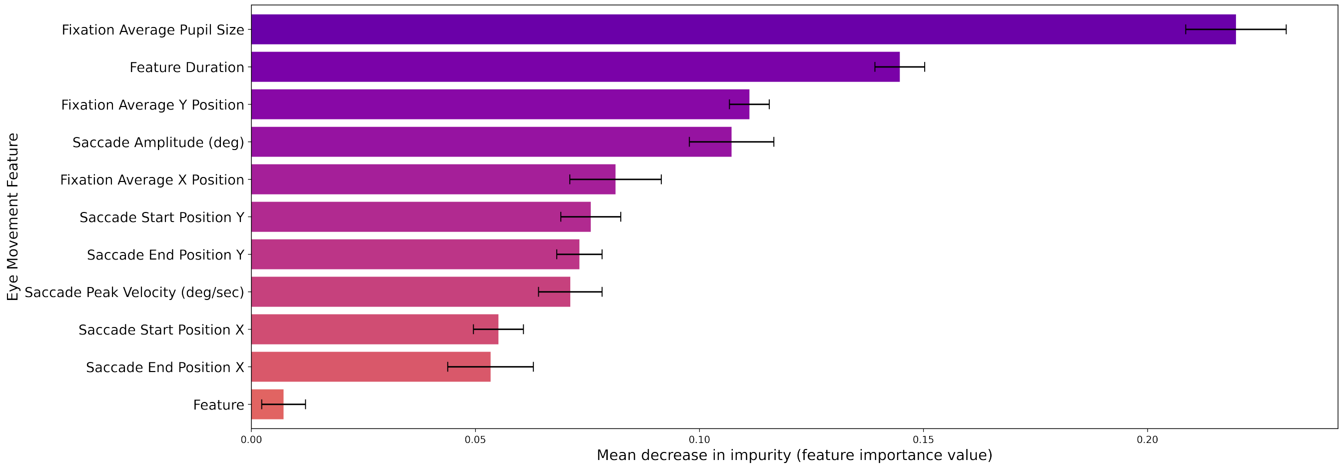


Figure 3: Feature importances measured by the mean decrease of Gini-impurity for the leave-one-participant-out nested cross-validation. The pupil size achieves the highest importance among all features.

An appropriate configuration of the hyperparameters is necessary to produce a model with the best performance for the problem. For this, we used grid search for exploring the hyperparameter space, using *scikit-learn*. We opted to do an exhaustive search of the most suitable kernel type, C (i.e., the regularization parameters), and γ (i.e., the radial basis width) for the SVM.

Similarly, we did an exhaustive search to find the optimal number of trees for the RF (i.e., denoted as $n\_estimators$), the criterion, maximal depth of the trees in the forest, the minimal sample split, and the $max\_features$ parameter. Having a larger number of trees in the forest increases the classification efficiency at the cost of increasing the computational time required to train a model. When splitting a node in the decision tree, the feature used for the split is selected from a random subset of features. The number of features chosen for this subset is determined by the $max\_feature$ parameter. The other RF parameters remained in their default settings.

3.3 Deep Learning

3.3.1 Dataset Preprocessing. For the deep learning models, one second of data was removed from the eye movement data before and after the calibration and validation sequence. The following windows represent the data interval that the model needs to estimate the likelihood of CFL. The window sizes vary from 50 msec to 10,000 msec of data, increasing in steps of 50 msec. The overlap between windows is always exactly half of the window size, e.g., if the window size is 50 msec, the window is moved 25 msec at a time. The window is removed if the window contains invalid values (e.g., because of blinks). Following this, we normalized all windows by moving the series of data points at the start of each window to the coordinate 0,0. Afterward, we turned the window such that the data point following exactly points up. Each window contains the x and y screen gaze coordinates (i.e., the pixel coordinates) and the pupil size value in the sum of the number of pixels inside the detected pupil contour.

3.3.2 Model Description. We use a Long Short-Term Memory (LSTM) network for learning temporal long-term dependencies, especially when predicting time-series which is usually the case when classifying gaze data. The input layer of our model varied in size to accommodate the window size, e.g., when there was a window size of 50 msec, the input layer was 50×3 (x, y, pupil dilation). After the input layer, we used an average pool layer with a pool size of 3. The third layer is our first LSTM layer with 32 nodes. The fourth layer was another average pool layer with a pool size of 3. The fifth layer was a dropout layer with the dropout set to 0.5. The sixth layer is our second LSTM layer with 20 nodes. After which, we have a second dropout layer, again set to 0.5. Followed by the output layer with 2 nodes representing the no-CFL or CFL condition.

We trained the model using an exponential decay learning rate scheduler, with an initial learning date of 0.0001, decay steps of 20,000, and a decay rate of 0.98. We enabled early stopping while monitoring the validation precision-recall with a patience 10. We set the maximum number of epochs to 500 and the batch size of 128. All models were trained with a loss function for binary cross-entropy and precision-recall area under the curve for the metrics, using one GPU per run (Tesla V100), in TensorFlow using the Keras API [17].

3.4 Evaluation

All models were evaluated using a leave-one-participant-out nested cross-validation. All run times of final models are done on a MacBook Pro 13" (i7 2.8 GHz, 16 GB, Intel 655). For the cost function, we take the run time into account that a single entry requires to be evaluated by our models. We denote the cost as the following:

$$Cost = \begin{cases} t, & \text{if } t < w \\ 2t, & \text{if } t \geq w \end{cases} \quad (1)$$

t is the time in msec needed to evaluate a single entry, and w is the duration of the feature or window size in msec, depending on whether we evaluate our approach via machine or deep learning.

Table 1: Accuracy, precision, recall, weighted F_1 scores and AUC PR of the leave-one-participant-out cross-validation for the best models from the hyper-parameters optimization for the feature-based models.

Model	Parameter	Accuracy	Precision	Recall	Weighted F_1-Score	AUC PR
SVM	C: 1000 Kernel: rbf	.730	.736	.722	.726	.821
RF	Criterion: Gini Max Depth: None Max Features: Auto Min Sample Split: 6 Number of Estimators: 900	**.769**	**.758**	.750	**.766**	**.863**
RF	Criterion: Gini Max Depth: None Max Features: Auto Min Sample Split: 9 Number of Estimators: 300	.766	.756	**.752**	.765	.860

The cost function is a linear function of the time required to evaluate an entry. However, if the time required for evaluating the entry is larger than the duration or window size, there will be a penalty of a factor of two for the time.

4 RESULTS

We compared the classification accuracy between traditional machine learning (i.e., SVM, RF) and deep learning (i.e., LSTM), after which we compared these models using the cost-benefit function. We evaluate a general classification model by investigating the classification performance through leave-one-participant-out cross-validation. This means that we use all participant data for training except for one participant, which we use for evaluation. Semantically, this approach learns a model without knowing anything about the person in advance and predicts the presence of CFL independent of individual context and differences. We describe the results in the following.

4.1 Machine Learning

We applied the methods of Section 3.2, which resulted in 11 features and 33,583 entries for these features combined. We then trained and validated the model with 100 cycles for precision and recall and using a leave-one-participant-out cross-validation for the two classification techniques and their combinations of parameters. Table 1 shows an overview of the best results. The column *model* indicates the machine learning method used to produce the model. The column "parameter" column indicates the parameters used to initialize the model. The precision and recall columns contain the average score with their standard deviation.

4.1.1 Support Vector Machine. The SVM showed the best performance by using $C = 1000$ and with a radial basis function kernel, resulting in an accuracy of .73 (precision: .72, recall: .73, weighted F_1 score: .73). We found a single set of parameters that performed best for precision and recall.

4.1.2 Random Forests. The RF classifier outperformed the SVM classifier for optimizing for precision and recall with two sets of parameters as shown in Table 1. We found the difference in the two models on two features, namely, the minimum sample split and the number of estimators. While the best-performing model for precision uses a minimum sample split of 6 and 900 estimators, the best model for recall uses values of 9 and 300, respectively.

Analyzing the most relevant features shows us that the average pupil size accounts for .238 of the Mean Decrease in Impurity (MDI), which can be seen in Figure 3. Followed by the saccadic amplitude, the duration of the feature and the average y position during a fixation. Together, these four features make up over half of the MDI.

4.2 Deep Learning

We applied the methods described in Section 3.3 to train our LSTM models. We trained and validated our models using the leave-one-participant-out cross-validation, resulting in a total of 1000 trained models. Each is trained on a different combination of window sizes and participants. As visualized in Figure 4a, we can see a general increase in mean Area Under the Curve (AUC) for precision and recall as we increase the window size for our LSTM models. This is to be expected.

More interestingly, the steepest AUC benefit is achieved at ca. 1,600 msec. Accuracy increases are minimal and variable after 5 seconds of window size, although increasing the window size to 9950 msec generated accuracy benefits of up to .725. However, increasing the window size also tends to capture more blinks. This resulted in missing values, resulting in sparser data. We visualize the eye-tracking data for windows sizes of 1250 msec in Figure 5. Here we see all the individual data points visualized after processing for participant five for both conditions. The data points scatter more during the CFL simulation conditions than those without CFL.

4.3 Evaluation

We applied the methods from Section 3.4. We ran each model 100 times, evaluating a single data entry. In Figure 4a, we show a linear relation between mean evaluation time and the window size, denoted as $Evaluation\ Time = 0.00415 \times window\ size + 37$. We also see that the penalty function described in Section 3.4 never applies, as the evaluation time stays well below the window size.

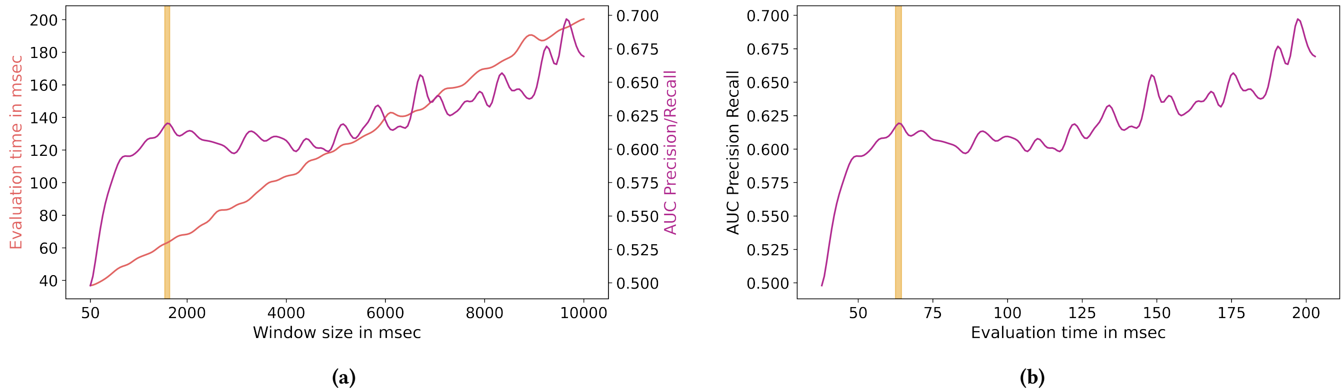


Figure 4: (a): Mean evaluation time plotted on the left-y-axis, mean AUC Precision-Recall plotted on the right-y-axis,the accompanying window size on the x-axis. The yellow line denotes the trade-off between the required window size and classification accuracy. A reasonable trade-off is achieved at ca. 1600 msec, since longer window sizes degrade the classification performance (e.g., due to loss of data through eye blinks). (b): Mean AUC Precision/Recall plotted on the y-axis and the accompanying evaluation time on the x-axis.

When visualizing the cost described by the linear equation above and the benefit from the increase in window size as visualized in Figure 4b we can see that there is a lot of benefit at the start while the evaluation time only increases slightly. After this, there is little gain in the AUC PR while the evaluation time keeps linearly increasing. Given that the information becomes sparser with the increase in window size, we recommend using a window size of 1600 msec for the best cost-benefit trade-off (see Figure 4b).

5 DISCUSSION

While previous research had already shown that classifying the presence of CFL was possible using eye tracking, we investigated what an optimal time frame would be to consider for evaluating the presence or absence of CFL. This work provides a cost-benefit analysis comparing different window sizes against the computational costs for CFL detection.

5.1 Cost-Benefit of CFL Detection

A 6° of CFL is a size many persons will not notice [24]. In contrast, our trained models can recognize 6° of CFL by observing differences in eye movements, although users will not notice this size of CFL. Our findings show that a window size of 1600 msec provides optimal results.

However, several constraints make long-term observations of CFL undesirable. For example, storage constraints and the amount of eye-tracking data that can be temporally or permanently stored on a device are limited. As 10 minutes of eye tracking data results in approximately 35 Megabyte of data, storing information for the entire day on mobile devices would therefore be unrealistic to tailor the model for individual users, even if the current trends of memory storage continues to increase. Further constraints include the processing power of mobile computing devices. As processing time correlates with the power usage required for evaluation, analyzing large amounts of data on mobile devices will reduce the battery life of these devices (e.g., tablets, smartphones, standalone mobile eye trackers). However, this issue can be circumvented by periodically transmitting the collected eye-tracking data to a server for computation. Subsequently, the model can be updated for individual users without laborious processing on mobile devices.

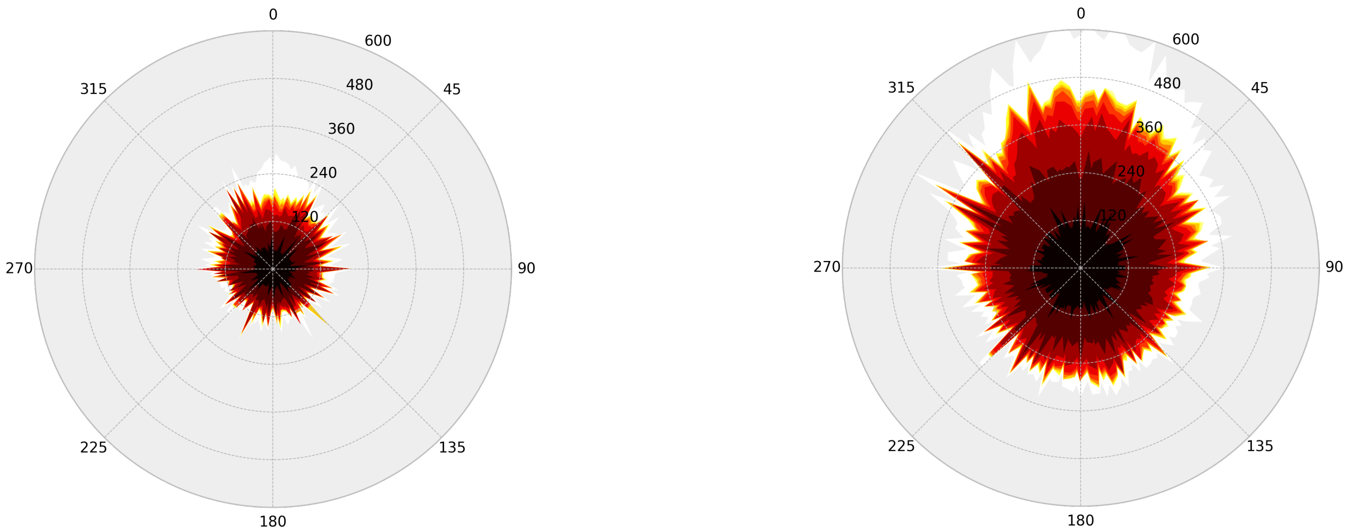
5.2 Generalized CFL Assessment

The leave-one-participant-out cross-validation predicts the presence or absence of CFL independent from the participant with an accuracy of .769 using RF. Consequently, the classification performance can be improved by accumulating crowd-sourced data from multiple or individual participants through non-intrusive ubiquitous eye-tracking. The general model does not require prior knowledge about the user and could be implemented into everyday systems without tailoring the model individually to a specific user. Hence, the general classifier can be used as a seed model to improve the classification accuracy individually. We envision that more research with crowd-sourced eye-tracking data from single users can improve the classification accuracy on an individual level. However, more research is required to consolidate this claim.

Results of feature importance agree with the findings reported by Verghese et al. [74], identifying that individuals with CFL have poor fixation stability and that they have difficulty directing their PRL to a target, which results in more and smaller saccades to reach a target. Furthermore, we identify the pupil size as an important feature (see Figure 3). However, we can also contribute the differences in pupil size as a factor of cognitive workload [7, 34, 72]. Furthermore, the interaction performance with the simulation may be improved when experiencing CFL for the first time, thus exerting more effort during the study [35]. However, there is scarce work on whether this also applies to patients with CFL due to a visual impairment.

5.3 Limitations and Future Work

We acknowledge that our study is prone to several limitations. We used a dataset containing data from a simulated CFL that might not correspond with CFL experienced in the real world. Participants



(a) (b)

Figure 5: Rose plot of eye movements after pre-processing for deep learning for participant 5 for a window size of 1250ms. It shows where eye movements are going towards and how many points are overlapping. Darker is more. (a) Without simulated CFL and (b) with simulated CFL.

who experience CFL for the first time through simulations might have affected eye movements through increased cognitive workload [34]. However, we still decided to use a simulation for our classification to control the CFL severance levels across all participants throughout the study. In future research, we will investigate the classification accuracy and required computational resources with participants affected by CFL with different severity levels.

Furthermore, our results assume the presence of a CFL with 6° for all participants, where the data collection for a few seconds is sufficient to predict CFL reliably. However, this is rarely the case in practice. CFL is a slow process that constantly shifts the gaze focus point by a few degrees. To investigate the evolution of the classification accuracy for different and increasing CFL levels, we will conduct user studies with CFL patients. This follow-up study will provide insights into changes in classification duration and accuracy.

The robustness of our approach heavily relies on the quality of eye-tracking data and the used pre-processing [59]. Thus, reliable in-the-wild CFL classification becomes less reliable with smaller window sizes, and as eye-tracking inherently suffers from blinks this is a challenge as shown by Grootjen et al. [26]. As state-of-the-art mobile eye tracking is currently not far from the sampling frequency we used to create our models [3] (e.g., AdHawks' MindLink with 500 Hz[2]). Previous research states that sampling errors are no practical problem for eye trackers operating above 200 Hz [1]. We state that this is not a cause for concern for our models. However, mobile head-worn eye trackers can exhibit significant errors from movements that occur during speech and facial expressions [47] and is further increased for those with CFL [42, 43]. In future work, we will investigate how eye-tracking data can be reliably obtained in realistic settings. In a real-world study, we will collect data, together with users with and without CFL, to assess the data quality and classification performance.

Furthermore, our findings suggest that individuals with induced CFL could experience higher cognitive workload due to an increased pupil dilation, where this might not be the case for patients who have CFL (see Figure 3), although various other reasons can influence pupil dilation. We will conduct a study to verify our findings with people who have CFL. We additionally observed that as we increased the window size for our models, we gradually got sparser data (e.g., because of blinks). Therefore, we will conduct a study to look into alternative ways to deal with the presence of blinks in eye movement data.

Our models classify between the presence or absence of an induced CFL of 6°. Using normal participants with induced CFL is a valid alternative to AMD patients [13, 44] since extensive literature has shown similar gaze behavior in both study members [8, 9, 38, 50, 80]. Nonetheless, in future work, we will focus on using eye-tracking data from patients with CFL to confirm the results found. Additionally, future studies will look into finding the minimal size of CFL that is detectable using eye tracking by integrating it into daily interactive systems. The earlier we can notify users of the presence of CFL, the earlier the appropriate actions can be taken to counter the progression of visual diseases (e.g., AMD). Consequently, early on, mitigation strategies can be adopted to reduce the decease's impact on the quality of life.

6 CONCLUSION

This paper presents a cost-benefit trade-off between model accuracy and computation to classify the presence or absence of Central Field Loss (CFL). We found that there is a lot of benefit for adding more

[2] www.adhawkmicrosystems.com/adhawk-mindlink

data to the input, and that there is an optimum around 1600 msec for our LSTM-NN model. This finding is unique as other findings focus on the gain acquired with more complex models and using context dependent data. However, using feature-based models, including random forest and support vector machines, outperform the LSTM-NN-based models. Our findings suggest that longer window sizes perform worse due to missing data or noise (e.g., eye blinks). Future work should investigate possible approaches to deal with blinks in continuous eye movement data for deep learning approaches, such as our LSTM-NN-based models.

ACKNOWLEDGMENTS

Jesse W. Grootjen was supported by the German Federal Ministry of Education and Research as part of the project IDeA (grant no. 16SV8119)

REFERENCES

[1] Richard Andersson, Marcus Nyström, and Kenneth Holmqvist. 2010. Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. *Journal of Eye Movement Research* 3, 3 (2010), 1–12. https://doi.org/10.16910/jemr.3.3.6

[2] Samantha Aziz and Oleg Komogortsev. 2022. An Assessment of the Eye Tracking Signal Quality Captured in the HoloLens 2. In *2022 Symposium on Eye Tracking Research and Applications* (Seattle, WA, USA) *(ETRA '22)*. Association for Computing Machinery, New York, NY, USA, Article 5, 6 pages. https://doi.org/10.1145/3517031.3529626

[3] Samantha Aziz, Dillon J Lohr, and Oleg Komogortsev. 2022. SynchronEyes: A Novel, Paired Data Set of Eye Movements Recorded Simultaneously with Remote and Wearable Eye-Tracking Devices. In *2022 Symposium on Eye Tracking Research and Applications* (Seattle, WA, USA) *(ETRA '22)*. Association for Computing Machinery, New York, NY, USA, Article 67, 6 pages. https://doi.org/10.1145/3517031.3532522

[4] Yair Bar-Haim, Talee Ziv, Dominique Lamy, and Richard M Hodes. 2006. Nature and nurture in own-race face processing. *Psychological science* 17, 2 (2006), 159–163. https://doi.org/10.1111/j.1467-9280.2006.01679.x

[5] Maria J Barraza-Bernal, Iliya V Ivanov, Svenja Nill, Katharina Rifai, Susanne Trauzettel-Klosinski, and Siegfried Wahl. 2017. Can positions in the visual field with high attentional capabilities be good candidates for a new preferred retinal locus? *Vision research* 140 (2017), 1–12. https://doi.org/10.1016/j.visres.2017.07.009

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

[7] J Beatty. 1982. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin* 92, 2 (1982), 276–292. https://doi.org/10.1037/0033-2909.91.2.276

[8] James H Bertera. 1988. The effect of simulated scotomas on visual search in normal subjects. *Investigative Ophthalmology & Visual Science* 29, 3 (1988), 470–475.

[9] James H Bertera. 1992. Oculomotor adaptation with virtual reality scotomas. *Simulation* 59, 1 (1992), 37–43.

[10] David Beymer and Daniel M Russell. 2005. WebGazeAnalyzer: a system for capturing and analyzing web reading behavior using eye gaze. In *CHI'05 extended abstracts on Human factors in computing systems* (Portland, OR, USA) *(CHI EA '05)*. Association for Computing Machinery, New York, NY, USA, 1913–1916. https://doi.org/10.1145/1056808.1057055

[11] Alexandre Bissoli, Daniel Lavino-Junior, Mariana Sime, Lucas Encarnação, and Teodiano Bastos-Filho. 2019. A Human–Machine Interface Based on Eye Tracking for Controlling and Monitoring a Smart Home Using the Internet of Things. *Sensors* 19, 4 (2019). https://doi.org/10.3390/s19040859

[12] Ali Borji and Laurent Itti. 2014. Defending Yarbus: Eye movements reveal observers' task. *Journal of vision* 14, 3 (2014), 29–29. https://doi.org/10.1111/j.1467-9280.2006.01679.x

[13] R Brandeis, I Egoz, D Peri, N Sapiens, and J Turetz. 2008. Psychophysical and perceptual performance in a simulated-scotoma model of human eye injury. In *Ophthalmic Technologies XVIII*, Vol. 6844. SPIE, 266–276.

[14] Leo Breiman. 2001. Random forests. *Machine learning* 45 (2001), 5–32. https://doi.org/10.1023/A:1010933404324

[15] Lars Buitinck, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, Peter Prettenhofer, Alexandre Gramfort, Jaques Grobler, Robert Layton, Jake Vanderplas, Arnaud Joly, Brian Holt, and Gaël Varoquaux. 2013. API design for machine learning software: experiences from the scikit-learn project. (2013). https://doi.org/10.48550/ARXIV.1309.0238

[16] Zenghai Chen, Hong Fu, Wai-Lun Lo, and Zheru Chi. 2018. Strabismus Recognition Using Eye-Tracking Data and Convolutional Neural Networks. *Journal of Healthcare Engineering* 2018 (April 2018), 7692198. https://doi.org/10.1155/2018/7692198 Publisher: Hindawi.

[17] François Chollet et al. 2015. Keras. https://keras.io.

[18] HL Cook, PJ Patel, and A Tufail. 2008. Age-related macular degeneration: diagnosis and management. *British medical bulletin* 85, 1 (2008), 127–149. https://doi.org/10.1093/bmb/ldn012

[19] Michael D Crossland, Louise E Culham, and Gary S Rubin. 2004. Fixation stability and reading speed in patients with newly developed macular disease. *Ophthalmic and Physiological Optics* 24, 4 (2004), 327–333. https://doi.org/10.1111/j.1475-1313.2004.00213.x

[20] Adele Cutler, D. Richard Cutler, and John R. Stevens. 2012. *Random Forests*. Springer US, Boston, MA, 157–175. https://doi.org/10.1007/978-1-4419-9326-7_5

[21] Martin Dechant, Sabine Trimpl, Christian Wolff, Andreas Mühlberger, and Youssef Shiban. 2017. Potential of virtual reality as a diagnostic tool for social anxiety: A pilot study. *Computers in Human Behavior* 76 (2017), 128–134.

[22] Martin Johannes Dechant, Julian Frommel, and Regan Lee Mandryk. 2021. The Development of Explicit and Implicit Game-Based Digital Behavioral Markers for the Assessment of Social Anxiety. *Frontiers in Psychology* 12 (2021).

[23] Andrew T Duchowski. 2002. A breadth-first survey of eye-tracking applications. *Behavior Research Methods Instruments and Computers* 34, 4 (2002), 455–470.

[24] Donald C Fletcher, Ronald A Schuchard, and Laura W Renninger. 2012. Patient awareness of binocular central scotoma in age-related macular degeneration. *Optometry and Vision Science* 89, 9 (2012), 1395–1398. https://doi.org/10.1097/OPX.0b013e318264cc77

[25] Reiko Graham, Alison Hoover, Natalie A Ceballos, and Oleg Komogortsev. 2011. Body mass index moderates gaze orienting biases and pupil diameter to high and low calorie food images. *Appetite* 56, 3 (2011), 577–586. https://doi.org/10.1016/j.appet.2011.01.029

[26] Jesse W Grootjen, Henrike Weingärtner, and Sven Mayer. 2023. Highlighting the Challenges of Blinks in Eye Tracking for Interactive Systems. In *Proceedings of the 2023 Symposium on Eye Tracking Research and Applications*. 1–7.

[27] Dan Witzner Hansen and Arthur E.C. Pece. 2005. Eye tracking in the wild. *Computer Vision and Image Understanding* 98, 1 (2005), 155–181. https://doi.org/10.1016/j.cviu.2004.07.013 Special Issue on Eye Detection and Tracking.

[28] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Computation* 9, 8 (Nov. 1997), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[29] Chien-Ming Huang, Sean Andrist, Allison Sauppé, and Bilge Mutlu. 2015. Using gaze patterns to predict task intent in collaboration. *Frontiers in psychology* 6 (2015), 1049. https://doi.org/10.3389/fpsyg.2015.01049

[30] Robert JK Jacob. 1991. The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Transactions on Information Systems (TOIS)* 9, 2 (1991), 152–169. https://doi.org/10.1145/123078.128728

[31] Marcel A Just and Patricia A Carpenter. 1980. A theory of reading: From eye fixations to comprehension. *Psychological review* 87, 4 (1980), 329. https://doi.org/10.1037/0033-295X.87.4.329

[32] Muhammad Qasim Khan and Sukhan Lee. 2019. Gaze and Eye Tracking: Techniques and Applications in ADAS. *Sensors (Basel, Switzerland)* 19, 24 (Dec. 2019), 5540. https://doi.org/10.3390/s19245540 Publisher: MDPI.

[33] Martin Kocur, Martin Johannes Dechant, Michael Lankes, Christian Wolff, and Regan Mandryk. 2020. Eye Caramba: Gaze-based Assistance for Virtual Reality Aiming and Throwing Tasks in Games. In *ACM Symposium on Eye Tracking Research and Applications*. 1–6.

[34] Thomas Kosch, Jakob Karolus, Johannes Zagermann, Harald Reiterer, Albrecht Schmidt, and Paweł W. Woźniak. 2023. A Survey on Measuring Cognitive Workload in Human-Computer Interaction. *ACM Comput. Surv.* 55, 13s, Article 283 (jul 2023), 39 pages. https://doi.org/10.1145/3582272

[35] Thomas Kosch, Robin Welsch, Lewis Chuang, and Albrecht Schmidt. 2023. The Placebo Effect of Artificial Intelligence in Human–Computer Interaction. *ACM Trans. Comput.-Hum. Interact.* 29, 6, Article 56 (jan 2023), 32 pages. https://doi.org/10.1145/3529225

[36] Girish Kumar and Susana TL Chung. 2014. Characteristics of fixational eye movements in people with macular disease. *Investigative ophthalmology & visual science* 55, 8 (2014), 5125–5133. https://doi.org/10.1167/iovs.14-14608

[37] Soichiro Kuwayama, Yuji Ayatsuka, Daisuke Yanagisono, Takaki Uta, Hideaki Usui, Aki Kato, Noriaki Takase, Yuichiro Ogura, and Tsutomu Yasukawa. 2019. Automated Detection of Macular Diseases by Optical Coherence Tomography and Artificial Intelligence Machine Learning of Optical Coherence Tomography Images. *Journal of Ophthalmology* 2019 (April 2019), 6319581. https://doi.org/10.1155/2019/6319581 Publisher: Hindawi.

[38] MiYoung Kwon, Anirvan S Nandy, and Bosco S Tjan. 2013. Rapid and persistent adaptability of human oculomotor control in response to simulated central vision loss. *Current Biology* 23, 17 (2013), 1663–1669.

[39] Bruno Laeng and Liv Falkenberg. 2007. Women's pupillary responses to sexually significant others during the hormonal cycle. *Hormones and behavior* 52, 4 (2007), 520–530. https://doi.org/10.1016/j.yhbeh.2007.07.013

[40] Michael Lankes, Joshua Newn, Bernhard Maurer, Eduardo Velloso, Martin Dechant, and Hans Gellersen. 2018. EyePlay revisited: Past, present and future challenges for eye-based interaction in games. In *Proceedings of the 2018 annual symposium on computer-human interaction in play companion extended abstracts*. 689–693.

[41] Daniel J. Liebling and Sören Preibusch. 2014. Privacy Considerations for a Pervasive Eye Tracking World. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (Seattle, Washington) *(UbiComp '14 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 1169–1177. https://doi.org/10.1145/2638728.2641688

[42] Al Lotze, Kassia Love, Anca Velisar, and Natela M Shanidze. 2022. A low-cost robotic oculomotor simulator for assessing eye tracking accuracy in health and disease. *Behavior Research Methods* (2022), 1–13. https://doi.org/10.3758/s13428-022-01938-w

[43] Kassia Love, Anca Velisar, and Natela Shanidze. 2021. Eye, Robot: Calibration Challenges and Potential Solutions for Wearable Eye Tracking in Individuals with Eccentric Fixation. In *ACM Symposium on Eye Tracking Research and Applications* (Virtual Event, Germany) *(ETRA '21 Adjunct)*. Association for Computing Machinery, New York, NY, USA, Article 16, 3 pages. https://doi.org/10.1145/3450341.3458489

[44] Marcello Maniglia, Kristina M Visscher, and Aaron R Seitz. 2023. Consistency of preferred retinal locus across tasks and participants trained with a simulated scotoma. *Vision Research* 203 (2023), 108158.

[45] LS Moulin, AP Alves Da Silva, MA El-Sharkawi, and Robert J Marks. 2004. Support vector machines for transient stability analysis of large-scale power systems. *IEEE Transactions on Power Systems* 19, 2 (2004), 818–825.

[46] DP Munoz, JR Broughton, JE Goldring, and IT Armstrong. 1998. Age-related performance of human subjects on saccadic eye movement tasks. *Experimental brain research* 121, 4 (1998), 391–400. https://doi.org/10.1007/s002210050473

[47] Diederick C. Niehorster, Thiago Santini, Roy S. Hessels, Ignace T. C. Hooge, Enkelejda Kasneci, and Marcus Nyström. 2020. The impact of slippage on the data quality of head-worn eye trackers. *Behavior Research Methods* 52, 3 (June 2020), 1140–1160. https://doi.org/10.3758/s13428-019-01307-0

[48] A. Petzold and G. T. Plant. 2005. Central and paracentral visual field defects and driving abilities. *International journal of ophthalmology* 219, 4 (2005), 191–201. https://doi.org/10.1159/000085727

[49] Ken Pfeuffer, Melodie Vidal, Jayson Turner, Andreas Bulling, and Hans Gellersen. 2013. Pursuit Calibration: Making Gaze Calibration Less Tedious and More Flexible. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology* (St. Andrews, Scotland, United Kingdom) *(UIST '13)*. Association for Computing Machinery, New York, NY, USA, 261–270. https://doi.org/10.1145/2501988.2501998

[50] Peter E Pidcoe and Paul A Wetzel. 2006. Oculomotor tracking strategy in normal subjects with and without simulated scotoma. *Investigative ophthalmology & visual science* 47, 1 (2006), 169–178.

[51] Jens Reinhard, André Messias, Klaus Dietz, Manfred MacKeben, Raimund Lakmann, Hendrik PN Scholl, Eckart Apfelstedt-Sylla, Bernhard HF Weber, Mathias W Seeliger, Eberhart Zrenner, et al. 2007. Quantifying fixation in patients with Stargardt disease. *Vision research* 47, 15 (2007), 2076–2085. https://doi.org/10.1016/j.visres.2007.04.012

[52] Luz Rello and Miguel Ballesteros. 2015. Detecting Readers with Dyslexia Using Machine Learning with Eye Tracking Measures. In *Proceedings of the 12th International Web for All Conference* (Florence, Italy) *(W4A '15)*. Association for Computing Machinery, New York, NY, USA, Article 16, 8 pages. https://doi.org/10.1145/2745555.2746644

[53] Laura Renninger and Anna Ma-Wyatt. 2011. Recalibration of eye and hand reference frames in age-related macular degeneration. *Journal of Vision* 11, 11 (2011), 954–954. https://doi.org/10.1167/11.11.954

[54] K Rohrschneider, M Becker, FE Kruse, T Fendrich, and HE Völcker. 1995. Stability of fixation: results of fundus-controlled examination using the scanning laser ophthalmoscope. *German journal of ophthalmology* 4, 4 (1995), 197–202.

[55] Gary S Rubin and Mary Feely. 2009. The role of eye movements during reading in patients with age-related macular degeneration (AMD). *Neuro-Ophthalmology* 33, 3 (2009), 120–126. https://doi.org/10.1080/01658100902998732

[56] Ronald A. Schuchard. 2005. Preferred retinal loci and macular scotoma characteristics in patients with age-related macular degeneration. *Canadian Journal of Ophthalmology* 40, 3 (2005), 303–312. https://doi.org/10.1016/S0008-4182(05)80073-0

[57] William Seiple, Patricia Grant, and Janet P Szlyk. 2011. Reading rehabilitation of individuals with AMD: relative effectiveness of training approaches. *Investigative ophthalmology & visual science* 52, 6 (2011), 2938–2944. https://doi.org/10.1167/iovs.10-6137

[58] William Seiple, Richard B Rosen, and Patricia MT Garcia. 2013. Abnormal fixation in individuals with age-related macular degeneration when viewing an image of a face. *Optometry and Vision Science* 90, 1 (2013), 45–56. https://doi.org/10.1097/OPX.0b013e3182794775

[59] Natela M. Shanidze, Zachary Lively, Rachel Lee, and Preeti Verghese. 2022. Saccadic contributions to smooth pursuit in macular degeneration. *Vision Research* 200 (2022), 108102. https://doi.org/10.1016/j.visres.2022.108102

[60] Satoshi Shioiri and Mitsuo Ikeda. 1989. Useful Resolution for Picture Perception as a Function of Eccentricity. *Perception* 18, 3 (1989), 347–361. https://doi.org/10.1068/p180347 arXiv:https://doi.org/10.1068/p180347 PMID: 2798018.

[61] Alexandra Sipatchin, Miguel García García, and Siegfried Wahl. 2022. Assistance for macular degeneration (MD): Different strategies for different augmentations. *Investigative Ophthalmology & Visual Science* 63, 7 (2022), 714–F0442.

[62] John D. Smith, Roel Vertegaal, and Changuk Sohn. 2005. ViewPointer: Lightweight Calibration-Free Eye Tracking for Ubiquitous Handsfree Deixis. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology* (Seattle, WA, USA) *(UIST '05)*. Association for Computing Machinery, New York, NY, USA, 53–61. https://doi.org/10.1145/1095034.1095043

[63] Nicholas D. Smith, Fiona C. Glen, Vera M. Mönter, and David P. Crabb. 2014. Using Eye Tracking to Assess Reading Performance in Patients with Glaucoma: A Within-Person Study. *Journal of Ophthalmology* 2014 (May 2014), 120528. https://doi.org/10.1155/2014/120528 Publisher: Hindawi Publishing Corporation.

[64] Tomoaki Sonobe, Hitoshi Tabuchi, Hideharu Ohsugi, Hiroki Masumoto, Naohumi Ishitobi, Shoji Morita, Hiroki Enno, and Daisuke Nagasato. 2019. Comparison between support vector machine and deep learning, machine-learning technologies for detecting epiretinal membrane using 3D-OCT. *International ophthalmology* 39, 8 (2019), 1871–1877. https://doi.org/10.1007/s10792-018-1016-x

[65] Julian Steil, Marion Koelle, Wilko Heuten, Susanne Boll, and Andreas Bulling. 2019. PrivacEye: Privacy-Preserving Head-Mounted Eye Tracking Using Egocentric Scene Image and Eye Movement Features. In *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications* (Denver, Colorado) *(ETRA '19)*. Association for Computing Machinery, New York, NY, USA, Article 26, 10 pages. https://doi.org/10.1145/3314111.3319913

[66] Janet S Sunness, Carol A Applegate, David Haselwood, and Gary S Rubin. 1996. Fixation patterns and reading rates in eyes with central scotomas from advanced atrophic age-related macular degeneration and Stargardt disease. *Ophthalmology* 103, 9 (1996), 1458–1466. https://doi.org/10.1016/S0161-6420(96)30483-1

[67] Atsunobu Takeda, Judit Z Baffi, Mark E Kleinman, Won Gil Cho, Miho Nozaki, Kiyoshi Yamada, Hiroki Kaneko, Romulo JC Albuquerque, Sami Dridi, Kuniharu Saito, et al. 2009. CCR3 is a target for age-related macular degeneration diagnosis and therapy. *Nature* 460, 7252 (2009), 225–230. https://doi.org/10.1038/nature08151

[68] George T Timberlake, Eli Peli, Edward A Essock, and Reed A Augliere. 1987. Reading with a macular scotoma. II. Retinal locus for scanning text. *Investigative ophthalmology & visual science* 28, 8 (1987), 1268–1274.

[69] Duy Tran, Weihua Sheng, Li Liu, and Meiqin Liu. 2015. A Hidden Markov Model based driver intention prediction system. In *2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*. IEEE, 115–120. https://doi.org/10.1109/CYBER.2015.7287920

[70] Nachiappan Valliappan, Na Dai, Ethan Steinberg, Junfeng He, Kantwon Rogers, Venky Ramachandran, Pingmei Xu, Mina Shojaeizadeh, Li Guo, Kai Kohlhoff, et al. 2020. Accelerating eye movement research via accurate and affordable smartphone eye tracking. *Nature communications* 11, 1 (2020), 4553. https://doi.org/10.1038/s41467-020-18360-5

[71] Stefan Van der Stigchel, Richard AI Bethlehem, Barrie P Klein, Tos TJM Berendschot, Tanja Nijboer, and Serge O Dumoulin. 2013. Macular degeneration affects eye movement behavior during visual search. *Frontiers in psychology* 4 (2013), 579. https://doi.org/10.3389/fpsyg.2013.00579

[72] Pauline van der Wel and Henk van Steenbergen. 2018. Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic bulletin & review* 25, 6 (2018), 2005–2015. https://doi.org/10.3758/s13423-018-1432-y

[73] Cornelis A Verezen, Carel B Hoyng, Carina FM Meulendijks, Jan EE Keunen, and B Jeroen Klevering. 2011. Eccentric gaze direction in patients with central field loss. *Optometry and Vision Science* 88, 10 (2011), 1164–1171. https://doi.org/10.1097/OPX.0b013e31822891e0

[74] Preeti Verghese, Cécile Vullings, and Natela Shanidze. 2021. Eye Movements in Macular Degeneration. *Annual Review of Vision Science* 7 (2021), 773–791. https://doi.org/10.1146/annurev-vision-100119-125555

[75] Mélodie Vidal, Jayson Turner, Andreas Bulling, and Hans Gellersen. 2012. Wearable eye tracking for mental health monitoring. *Computer Communications* 35, 11 (2012), 1306–1311. https://doi.org/10.1016/j.comcom.2011.11.002

[76] Sourabh Vora, Akshay Rangesh, and Mohan Manubhai Trivedi. 2018. Driver gaze zone estimation using convolutional neural networks: A general framework and ablative analysis. *IEEE Transactions on Intelligent Vehicles* 3, 3 (2018), 254–265. https://doi.org/10.1109/TIV.2018.2843120

[77] Janis M White and Harold E Bedell. 1990. The oculomotor reference in humans with bilateral macular disease. *Investigative ophthalmology & visual science* 31, 6 (1990), 1149–1161.

[78] Stephen G Whittaker, James Budd, and RW Cummings. 1988. Eccentric fixation with macular scotoma. *Investigative ophthalmology & visual science* 29, 2 (1988), 268–278.

[79] World Health Organization. 2019. World Report on Vision: Executive Summary. *World Health Organization* (2019). https://www.who.int/news/item/08-10-2019-who-launches-first-world-report-on-vision Last accessed on 2021-01-10.

[80] Haojie Wu, Daniel H Ashmead, Haley Adams, and Bobby Bodenheimer. 2018. Using virtual reality to assess the street crossing behavior of pedestrians with simulated macular degeneration at a roundabout. *Frontiers in ICT* 5 (2018), 27.

[81] Ai Ping Yow, Damon Wong, Huiying Liu, Hongyuan Zhu, Ivy Jing-Wen Ong, Augustinus Laude, and Tock Han Lim. 2017. Automatic visual impairment detection system for age-related eye diseases through gaze analysis. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2450–2453. https://doi.org/10.1109/EMBC.2017.8037352