Evaluating Accuracy of the Tobii Eye Tracker 5

Chapter · January 2022

DOI: 10.1007/978-3-030-98404-5_36

5 authors, including:

Andrew Housholder
Augustana College

Jonathan Reaban
Augustana College

Tauheed Khan Mohd
Eastern Michigan University

Evaluating Accuracy of the Tobii Eye Tracker 5

Andrew Housholder[1], Jonathan Reaban[2], Aira Peregrino[3], Georgia Votta[4], and Tauheed Khan Mohd[5]

Augustana College, Rock Island, IL 61201

Abstract. Eye-tracking sensors are a relatively new technology and currently has use as an accessibility method to allow those with disabilities to use technology with greater independence. This study evaluates the general accuracy and precision of Tobii eye-tracking software and hardware, along with the efficacy of training a neural network to improve both aspects of the eye-tracker itself. With three human testers observing a grid of data points, the measured and known point locations are recorded and analyzed using over 250 data points. The study was conducted over two days, with each participant performing four trials each. In this study, we use basic statistics and a k-means clustering algorithm to examine the data in depth and give insights into the performance of the Tobii-5 eye-tracker. In addition to evaluating performance, this study also attempts to improve the accuracy of the Tobii-5 eye-tracker by using a Multi-Layer Perceptron Regressor to reassign gaze locations to better line up with the expected gaze location. Potential future developments are also discussed

Keywords: Tobii, eye-tracking, Tobii Eye Tracker 5, calibration, python, point-of-gaze

1 Introduction

Eye-tracking as a field of study has been around since the late 1800s when Louis Emile Javal noticed strange behavior in people's eyes while reading across a page; people would spend differing amounts of time on different words instead of a smooth motion across the page. This was the first real use of eye tracking, even if it was relatively simple. Later on, Edmond Huey was able to build a device that identified what words people were pausing on for long periods of time. Huey's research cemented him as an important figure in the early days of the study of reading. Over the years, eye tracking gradually progressed from educational fields like reading into the more lucrative fields of marketing, business, and web design. By the late 1900s, eye-tracking technology became instrumental in modernizing web pages from looking like the print sources such as newspapers that were their predecessors. A study in 1991 specifically examined the use of eye-trackers as a normalized human-computer interaction technique [1]. They suggested that, while eye-tracking at time of writing is not accurate enough to replace or substitute standard mouse/keyboard control, eye-tracking could reasonably be used as a complement with mouse/keyboard in the future. Today, eye-tracking is finding a niche in "in-game" advertising such as product placement in movies and video games [2]. This is bringing the prediction made in 1991 come true.

Today, most companies that produce eye-tracking hardware offer eye-tracking based off of image-based techniques [3]. Less intrusive models (e.g. a device that sits on the computer) have become most popular among consumers. These devices commonly use a combination of near-infrared light and a camera to capture data. The IR light reflects off of the inside of the user's eye, where the reflected light is captured in images of the eyes. These images are then analyzed in order to produce a gaze direction. This gaze direction data is then used to provide raw input on the computer, based on the devices location on the computer.

Eye-tracking as a new technology provides human-computer interactions a new source of input in the form of the human visual experience. Because of it's potential applications in neuroscience [4], psychology [5], and computer science [6], eye-tracking has been both more prevalent and more developed in recent years. Tobii Tech, a branch of the larger tech company Tobii, has been primarily focused on evolving this technology and is currently the world leader in eye-tracking technology. It's success has only been limited by its costs; recently, however, consumer demand has brought this barrier down to the point where eye-tracking software is almost commercially available from any location. Some of the biggest companies offering eye-tracking hardware at the time this paper was written are Gazepoint (based in Vancouver, B.C., Canada), Natural Point (based in Corvallis, OR, USA), and Tobii (based in Stockholm, Sweden).

With eye-tracking hardware and software becoming more widely available, it is important that the technology is very accurate to the user's actual eye position. This paper is focused on evaluating the accuracy of the Tobii 5 eye-tracker (See Fig. 1), the most recent publicly available device at time of writing capable of tracking both head and eye movement. The device is plugged into the computer, mounted on either the top or the bottom of the monitor, and then pointed towards the user. The device then gives a visual que to where the user is looking on screen. With an operable distance of 45-95cm, a sampling rate of 133Hz and infrared illuminators, Tobii claims that the device is accurate and performs exceptionally well under any lighting conditions at determining an individual's point of gaze (POG). This paper will specifically provide data on the accuracy of the Tobii 5 in the different quadrants of the screen (upper left, upper right, lower left, lower right) and test whether the eye-tracker is more or less accurate towards the edges of the screen versus the center of the screen.

The testing was performed using both the Tobii 5 eye-tracker hardware and the Unity game engine. There is a free game included in Unity that tracks your POG via the Tobii 5 eye-tracker and returns the exact coordinates (in pixels) as text fields in the game. Using a combination of this game and an image

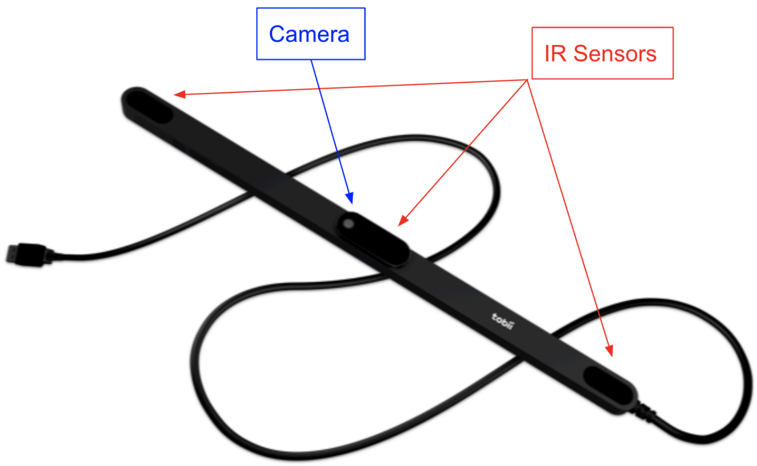

Fig. 1: A labeled diagram of the Tobii 5 Eye-Tracker hardware used in the experiment.

containing 21 fixed points of known location, we were able to manually collect data for testing.

Accuracy is defined in this study as the average difference between the known location and the eye-tracker's measured location. Precision is defined as the standard deviation of the data in each trial; the eye-tracker's capability in consistently producing the same output from the same location (See Fig. 2).

When examining the data from this study, we employed both basic statistics and a k-means clustering algorithm to gain insights into the performance of the Tobii 5 eye-tracker. Overall, we found that the Tobii 5 eye-tracker strays on average 35 pixels (visual angle of .74 degrees) from the target and has a standard deviation of 18 pixels (.39 degrees) according to the data.

The data was acquired by having 3 people each perform 4 trials in which they looked at 21 fixed points on the screen for a total of 252 points, distributed at even intervals across the screen. Next, we compared the actual location of the point in pixels to the location returned by the Tobii 5 eye-tracker and were able to do our data analysis, using both excel and python (pandas, scikit-learn, matplotlib libraries were used).

In addition to simply examining the accuracy of the Tobii-5 eye-tracker, solutions are also explored for improving the overall accuracy of the device through software and neural networks. A Multi-Layer Perceptron Regressor was implemented to determine the error for the POG anywhere on the screen. This neural network helps to provide more accuracy to the Tobii-5 eye-tracker hardware through the implementation of software.

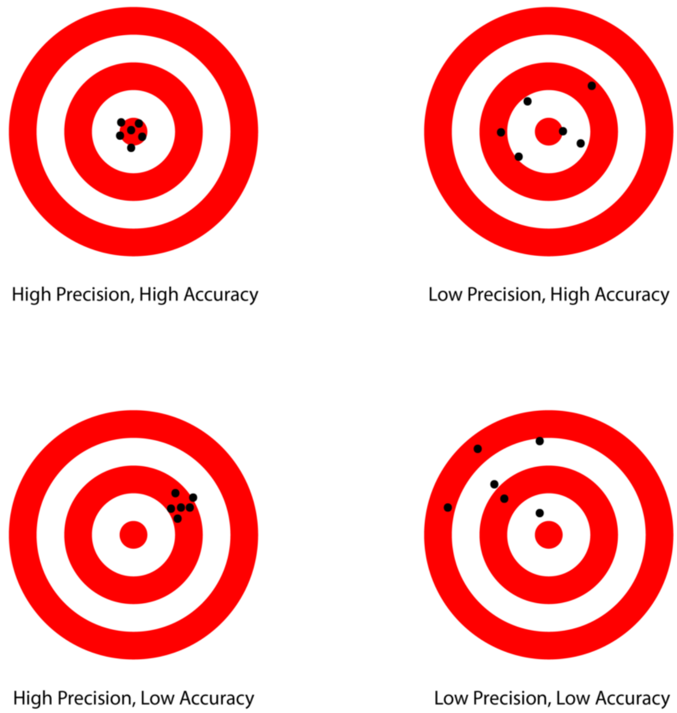

Fig. 2: Distinction between high accuracy and high precision from [7].

2 Related Work

Eye tracking technology has become a popular area of research within the field of Human Computer Interaction and new applications are emerging everyday. From research [3], to psychology [5], to making the lives of disabled people less difficult [8], a large number of people are gradually adopting eye tracking technology and making it a more prevalent tool in everyday life.

In terms of research, one subset of study focuses heavily on examining the accuracy, precision, and overall performance of the eye tracking technology [9]. These types of studies collect large amounts of data and attempt to quantify the performance of the eye tracking devices. They include studies regarding efficacy when user is wearing glasses [9], efficiency of eye tracking in comparison to other multi-modal human-computer interfaces [10], and even the accuracy between specific brands of eye-tracker [11].

Another subset examines the actual metrics used to determine what "accuracy" [12], "precision" [13], and "performance" [14] actually means in regards to eye-tracking sensors in general. The Tobii website has published their specific methodology in regards to testing their own device, with accuracy measured through Euclidean distance between points, and precision measured through the root mean square (RMS) between successive sample points. [15]. Another study conducted by Microsoft uses the standard deviation to illustrate the spread of points around the "average" [16].

Beyond the examination of accuracy and precision of eye-tracking technology, work has been done to seek improvement of certain flaws that exist within eye-trackers [17]. This study made advancements in three areas of improvement: smoothing gaze data in real time, accounting for sensor lag in real time, and providing focus points for the user's eyes. Two of the three improvement attempts showed moderate success (data smoothing and sensor lag adjustments). Another study suggests a new mechanism to convert eye-gaze-tracking data to mouse-pointer data [18]. This study in particular finds that although using eye-gazing data as a primary input is fast, it is still much less accurate and stable than most other input devices [19] due to saccades, the rapid movement of eyes between two different fixation points. Saccades are the concurrent movement of both eyes and result in very unpredictable pathing from one point to another. However, the work done in the study confirms that the use of eye-gaze data as a computer input is feasible with the implementation of algorithms that reduce the effects of the jittery saccade behavior of the eye.

With the right implementation, there are some potentially useful applications of eye-tracking technology that have been discussed. Using eye gaze data as a means for target selection on a computer screen is an application that has been explored in a recent study [20].

As eye-tracking performance improves, its application to other technologies could help make new devices more accessible. One study has shown specifically that human testers using an eye-tracking device can perform just as well as human testers using mouse and keyboard [8]. The biggest impact on eye-tracking user performance is primarily precision and accuracy, rather than speed or ease of use. Users were able to quickly adjust to the eye-tracking technology.

With major advancements being made in eye-tracking technology and its applications, there are still some limitations to the type of data collected. Most eye trackers are not necessarily portable and only are for use on computer or laptop screens. One study examines how gaze data could be collected without any external hardware [21]. With a significant data set of over 2.5 million frames, this study uses the data to attempt to predict the location of the users' gaze. The overall goal of the study is to make eye-tracking technology available to all mobile devices without any sort of external hardware. This, however, is some ways away.

3 Experiment

3.1 Technology Specifications

Tobii Experience was downloaded onto a Dell PC running Windows 10 with a 3.0 GHz Intel core 2 Quad processor with 32 GB of RAM on a 24-inch thin film transistor liquid crystal display, 1920 x 1080 pixels. The Tobii 5 hardware was connected through a USB port and mounted to the bottom of the monitor facing the user. With no other operations running in the background, the test was performed using an application that plots exact x and y coordinates of the eye-tracker's calculated viewpoint. This application can be found on a project in the Tobii Unity SDK for Windows.

3.2 Procedure

For this study three participants were screened. The participants were aged in their twenties and none require glasses or wore any type of accessory that would potentially interfere with the Tobii Eye Tracker. All three consenting participants were informed ahead of time about the study's methods and nature. To test the accuracy and precision of the eye-tracking technology, known locations were marked on the screen for testers to look at down to the exact pixel. Upon installation of the device drivers, the user is provided with various demo projects in Unity. One of these applications is to plot the exact x and y location, in pixels, of where the eye tracking device is reading the users point of gaze to be. Given the environment was in 1920p by 1080p, a transparent .png image was created with the matching dimensions of 1920p by 1080p with 21 predetermined points plotted out (See Fig. 3) sporadically across the screen. This image was produced fully in Adobe Photoshop. Each of the points consisted of a light blue circle with a black dot on the exact pixel of each point used for testing.

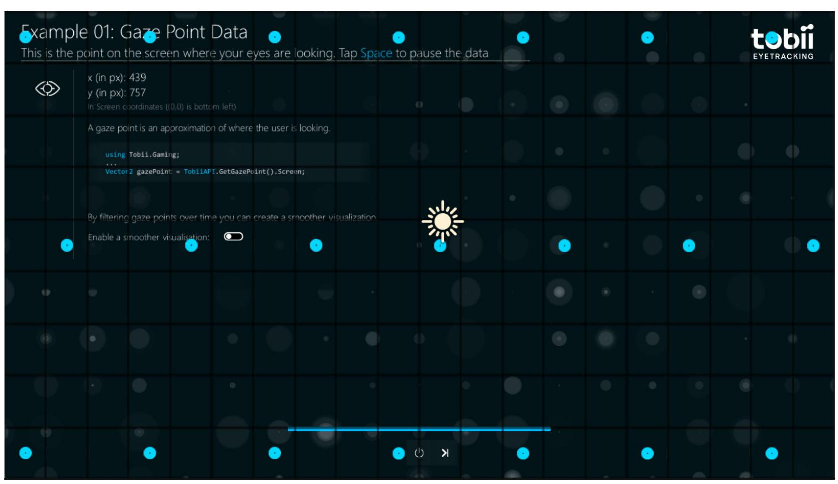

Fig. 3: The Unity environment used for obtaining data points. The user, with the eye-tracking device enabled, observes each point on the screen in reading order; left to right, top to bottom.

With the user seated the recommended 45-95cm away, the participants began by calibrating the device using the native Tobii Experience calibration tool. From there, they went through each point row by row and another individual recorded the determined POG by the eye tracker. This given point was then compared with the known location, and both locations were entered manually into a Microsoft Excel document.

3.3 Data Processing

By calculating the Euclidean distance between the determined POG and the given location, the approximate accuracy of the sensor can be calculated in terms of the visual angle. The test was repeated multiple times with multiple people to compare accuracy between the middle of the screen, the four corners, the top/bottom edges, and the left/right edges.

The test used was a grid of points made in Adobe Photoshop and the accuracy was calculated using the Euclidean distance to find the visual angle between the sensor's determined point and the actual point.

The Euclidean distance can be calculated using the x and y coordinates between the determined point and the actual point using the following equation:

$$Euclidean\ Distance = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{1}$$

The visual angle can be calculated using the Euclidean difference between the two points and the distance of the observer to the screen. The relationship is as follows:

$$Visual\ Angle = 2 * \arctan(\frac{\frac{Difference}{2}}{Distance}) \tag{2}$$

To evaluate accuracy, the average Euclidean distance across all test points was found using Microsoft Excel, along with the average across all corner points, the average across all points on the top/bottom edges, and the average across all points on the left/right edges.

To evaluate precision, the standard deviation of the Euclidean distance was taken across all points for each trial in the same categories as mentioned above.

The data was also processed using a neural network in python in an attempt to correct the Tobii data to display the actual coordinates where the user was looking. This neural network takes in Tobii gaze data as its input and outputs the predicted actual gaze location. The neural network is implemented in the scikit-learn package in python as an MLPRegressor.

The data was first scaled using scikit-learn's minmax scale method and then ten percent of the data was put into a test set at random while the remaining ninety percent was put into the training set. The labels for the training in test set were the known locations on the screen that the users looked at during testing and that input data was the output from the Tobii 5 eye-tracker.

One limitation to using a neural network to process this data is the amount of data collected. The amount of data collected only amounted to 252 points due to time constraints on the project. Ideally, we would want to have much more data (preferably in the thousands) with many more known points in order to get a more accurate neural network to better improve our results.

4 Results

The first set of results to look at are the statistics from the raw data from the experiment. A histogram of the calculated euclidean distances can be seen in Figure 4. From this histogram, we can see that the distribution of euclidean distances is skewed to the right so the measure of center and spread to use are median and interquartile range. The median of this data set is 30.70 pixels (.65 degrees) and the interquartile range is 33.37 pixels (.71 degrees).

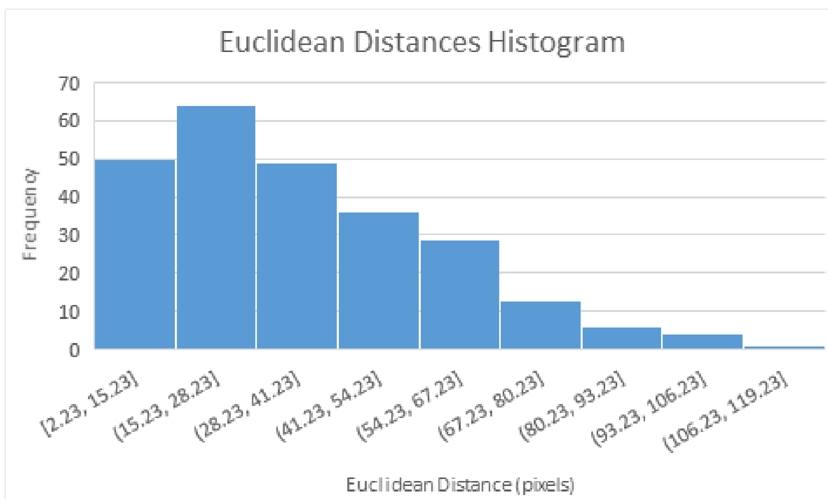

Fig. 4: Histogram of the euclidean distances between the data from the Tobii 5 eye-tracker and the actual location of the user's POG.

Using the scikit-learn KMeans clustering algorithm with 4 clusters, the clusters that appeared were the 4 different quadrants of the screen as shown in Figure 5. The choice of 4 clusters was made after using the elbow method for KMeans clustering which showed no obvious number of clusters to use; four clusters seemed to work the most consistently though. After looking at the euclidean distances from each of the actual fixed points to each test point in each of the clusters, however, there was no significant difference in the distribution between any of the quadrants. There was a slightly higher variance in the top left and bottom right quadrants than in the top right and bottom left quadrants, but this may only be due to the amount of data points in each of the clusters.

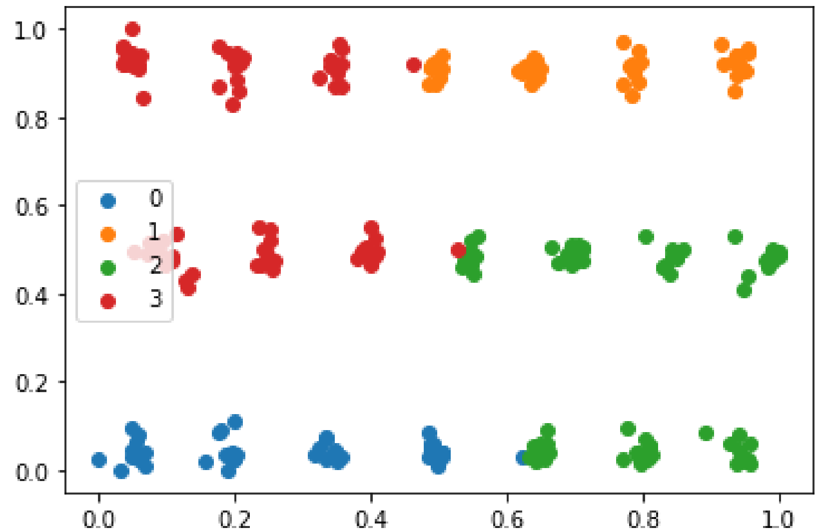

Fig. 5: A graph of the clusters that came from using the KMeans clustering algorithm.

A Multi-Layer Perceptron Regressor was also used to analyze the data. We decided not to use a CNN or RNN because our data was numeric and did not include a time component so a simple ANN would be good enough for our data set. This neural network attempts to take the Tobii-5 eye-tracker data and make it more accurate to where the user is actually looking. The neural network has a hidden layer size of 20 and a logistic activation function. The network also used a set random state for splitting the training and test data so it is trained on the same data every time. The loss curve from training the network can be seen in Figure 6. After training the network, it appears to be working as intended. For example, looking at the point (1808, 48), which was given from the Tobii-5 eye-tracker when the user was looking at the point (1776, 60), the network updated this point to be (1771, 57). The euclidean distance in this case decreased from 34.18 to 5.77 (1.15 degrees to 0.19). This decrease is similar across all data points in the test data. It is clear that the network's output is much more accurate to the actual point of gaze of the user than what the Tobii-5 eye-tracker had originally predicted.

There are some limitations to this method of improving the accuracy of the Tobii-5 eye-tracker however. One such limitation is the amount of data collected prior to training and testing. Due to time constraints on the project, only 252 data points were used for training and testing. With so few data points, it will be hard for the neural network to get highly accurate results. The data collection also only involved 21 known points. Better results could be obtained by collecting data from more known points. This would allow for us to see more clearly the

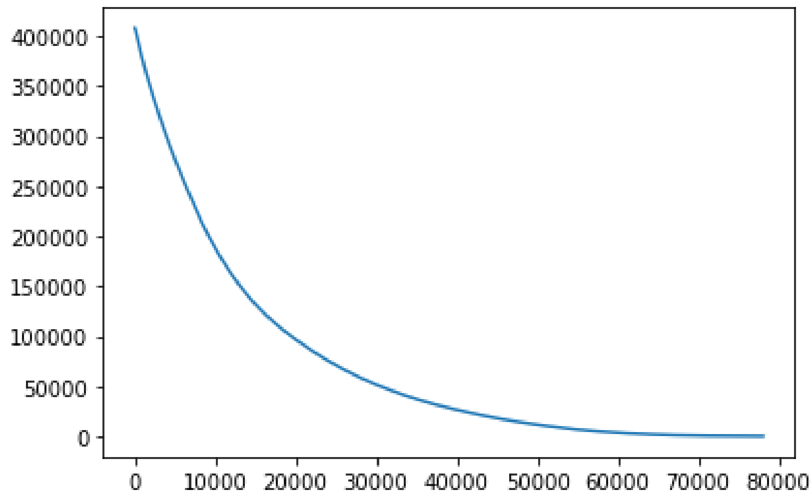

Fig. 6: The Loss Curve from the neural network. Along the horizontal axis is the training iterations that have occurred. Along the vertical axis is the loss of the neural network.

accuracy of the Tobii-5 eye-tracker in different regions of the screen. Another potential limitation of the neural network is that it only seemed to work very well on points that were already inaccurate. On points that were already fairly accurate, the neural network did not improve much or got slightly worse in terms of accuracy.

5 Conclusion and Future Work

From these results, we can conclude that the Tobii-5 eye-tracker is typically accurate to around 30 pixels (1.01 degrees) of the actual point of gaze of the user. This accuracy tends to be consistent throughout the entire screen. In other words, eye-tracker does not get more or less accurate when the user looks at different regions of the screen.

As previously stated, one of the biggest limitations with the Tobii-5 to this project thus far is the number of data points collected. Only 252 data points were used for training and testing the neural network as well as collecting the accuracy data. Ideally the the results would be made based off of a few thousand points; however, time constraints restricted how many data points could be collected. Many more known points on the screen would also be used to better test the accuracy in different regions in the screen.

We also plan to have more individuals participate in future studies. For this study, we only had three participants. Having closer to 20 or 30 participants would allow for even more data points and would give a better overall picture of the accuracy and precision of the Tobii-5 eye-tracker. The best case scenario would allow us to randomly select 30 participants from a large pool of individuals.

Our method of data collection was to manually record each coordinate as indicated by the Tobii-5 eye-tracker. An automatic data collection program would have been noticeably more efficient and would have allowed us to collect many more data points in a shorter time-frame.

With a larger pool of data, our results would have been considerably better and our software using the Multi-Layer Perceptron Regressor would be able to reliably correct the location of someone's gaze for any point on the screen. Currently the neural network only reliably corrects points that are very inaccurate and does not tend to improve points that are already relatively close to the actual point of gaze of the user.

It would be beneficial to study if this was consistent across a variety of screen sizes, how much the screen size affects the accuracy of the Tobii Eye Tracker 5, and then apply a dynamic progressive model to correct the areas that are typically further off from the ones that tend to be more accurate. This can be especially useful when the user is using their eyes as a mouse for example, as the buttons the user could click would be used as correction points, rather than just random points on the screen.

After finding a method to collect larger pools of data and to efficiently correct that data, we would have the opportunity to go further with our hypotheses regarding the accuracy and precision of the Tobii-5 eye-tracker. With more participants and data, we would be able to look into how optical aids such as eyeglasses and soft contact lenses effect the accuracy of eye-tracking. Though this has been done before [9], we would first be able to confirm or deny the findings of this paper by having more participants in our study. We then could also take this further by checking whether our correction algorithm would be able to also correct for inaccuracies produced by prescription eyeglasses or contacts.

References

1. R. J. Jacob, "The use of eye movements in human-computer interaction techniques: what you look at is what you get," *ACM Transactions on Information Systems (TOIS)*, vol. 9, no. 2, pp. 152–169, 1991.
2. D. Leggett, "A brief history of eye-tracking: Ux booth," Jan 2010.
3. C. Weigle and D. Banks, "Analysis of eye-tracking experiments performed on a tobii t60," 01 2008.
4. J. F. Hopstaken, D. Van Der Linden, A. B. Bakker, and M. A. Kompier, "A multi-faceted investigation of the link between mental fatigue and task disengagement," *Psychophysiology*, vol. 52, no. 3, pp. 305–315, 2015.
5. e. a. Morgante, James D., "A critical test of temporal spatial accuracy of the tobii t60xl eye tracker," *Infancy*, vol. 17, no. 1, pp. 9–32, 2012.
6. M. Elhelw, M. Nicolaou, A. Chung, G.-Z. Yang, and M. S. Atkins, "A gaze-based study for investigating the perception of visual realism in simulated scenes," *ACM Transactions on Applied Perception (TAP)*, vol. 5, no. 1, pp. 1–20, 2008.
7. GeoCue, "Accuracy, precision, resolution," 2015. [Online; accessed 28 January 2021].

8. M. Tall, A. Alapetite, J. San Agustin, H. H. Skovsgaard, J. P. Hansen, D. W. Hansen, and E. Møllenbach, “Gaze-controlled driving,” in *CHI ’09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA ’09, (New York, NY, USA), p. 4387–4392, Association for Computing Machinery, 2009.
9. J. Dahlberg, “Eye tracking with eye glasses,” pp. 1–66, Jun 2010.
10. B. Trouvain and C. M. Schlick, “Comparative study of multimodal displays for multirobot supervisory control,” *Engineering Psychology and Cognitive Ergonomics*, pp. 184–193, 2007.
11. G. Funke, E. Greenlee, M. Carter, A. Dukes, R. Brown, and L. Menke, “Which eye tracker is right for your research? performance evaluation of several cost variant eye trackers,” in *Proceedings of the Human Factors and Ergonomics Society annual meeting*, vol. 60, pp. 1240–1244, SAGE Publications Sage CA: Los Angeles, CA, 2016.
12. K. Holmqvist, M. Nystrm, H. Anderson, and J. Weijer, “Eye-tracking data and dependent variables,” *Niepublikowana praca, Lund University*, 2011.
13. U. Demšar and A. Çöltekin, “Quantifying gaze and mouse interactions on spatial visual interfaces with a new movement analytics methodology,” *PloS one*, vol. 12, no. 8, p. e0181818, 2017.
14. M. A. Goodrich and D. R. Olsen Jr, “Metrics for evaluating human-robot interactions,” 2003.
15. J. Johnsson and R. Matos, “Accuracy and precision test method for remote eye trackers,” *Sweden, Tobii Technology*, 2011.
16. A. M. Feit, S. Williams, A. Toledo, A. Paradiso, H. Kulkarni, S. Kane, and M. R. Morris, “Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design,” in *Proceedings of the 2017 Chi conference on human factors in computing systems*, pp. 1118–1130, 2017.
17. M. Kumar, J. Klingner, R. Puranik, T. Winograd, and A. Paepcke, “Improving the accuracy of gaze input for interaction,” in *Proceedings of the 2008 symposium on Eye tracking research applications*, pp. 65–68, 2008.
18. A. Sesin, M. Adjouadi, M. Ayala, A. Barreto, and N. Rishe, “Effective data conversion algorithm for real-time vision based human computer interface,” in *Proceedings of the 6th WSEAS Int. Conf. on Software Engineering, Parallel and Distributed Systems*, pp. 16–19, 2007.
19. e. a. Xue, Jiguo, “A crucial temporal accuracy test of combining eeg and tobii eye tracker,” *Medicine*, vol. 96, March 2017.
20. R. Vertegaal, “A fitts law comparison of eye tracking and manual input in the selection of visual targets,” in *Proceedings of the 10th international conference on Multimodal interfaces*, pp. 241–248, 2008.
21. K. Krafka, A. Khosla, P. Kellnhofer, H. Kannan, S. Bhandarkar, W. Matusik, and A. Torralba, “Eye tracking for everyone,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2176–2184, 2016.