Eye Tracking Experiments for Cybersecurity

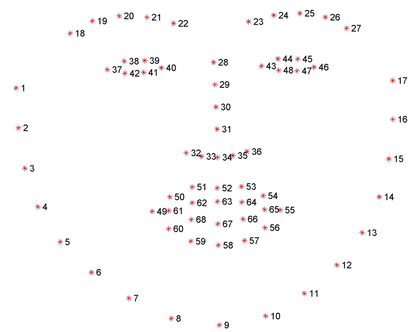
Xinyao Ma [1] Sunny Gandhi [1]

[1]Indiana University

Introduction

Our goal was to use eye tracking and head tracking software to determine if it was possible to draw conclusions that would allow web developers to integrate into their website designs to protecting against different potential cybersecurity breaches, such as phishing attacks.

Figure 1. Facial Keypoints



Our experiments involved centering on specific "keypoints" in this facial model that helped us determine where the user was looking.

Literature Review

The first step of the project — which took a considerable amount of time — was conducting a large scale literature review. We wanted to be able to understand three key things:

- **Experimental Design:** What kind of experimental designs have other researches used, and how would it serve our purposes?
- **Software Options:** What kind of softwares, technologies, and/or libraries exist that would allow us to conduct a successful experiment?
- **General Research Outcomes:** How much research and conclusions have been developed on this topic? (Spoiler! Not much at all!)

After much thought and review, we settled on developing a semi-custom program that would give us as much flexibility as possible, using the OpenCV-Python library to assist us.

An Aside on OpenCV Architecture

An important note is to understand what OpenCV is and why it was selected as fitting our criterion best for conducting this experiment. Put simply, OpenCV is a massive Python vision processing library that is powered by artificial intelligence and machine learning algorithms. We chose it for a few reasons:

- **Ease of Use:** It is remarkably easy to get started with OpenCV as there is no shortage of tutorials and other resources to assist someone with getting off the ground
- **Modern Compatibility:** OpenCV is easy to use with modern applications and operating systems that would make it simple to develop an experiment on the time crunch we had
- **Free Usage:** As we needed something that wouldn't cause any financial problems, OpenCV offered powerful processing ability for zero cost

Designing the Experiment

We decided upon utilizing subjects that were around 18 - 22 years old, as that was the easiest sample size to get and it would keep our findings consistent with a very similar age group. We also settled on using a Macbook Pro laptop, with Google Chrome as the browser.

We then created some basic fake web pages that looked extremely similar to real ones with minor differences that the user could navigate on. These web pages were hosted on a Raspberry Pi Apache2 web server.

The experimental process itself was extremely simple. We asked users to sit down and try to discern whether or not a web page was real or fake. We tracked their eye and head movement as they did this to determine what kinds of patterns users exhibit that could be used to draw conclusions on how to create webpages that protect against fake versions of themselves from, for example, scammers.

Building the Tracking Model

At first, the goal was to be able to just use eye tracking in order to determine where the user was looking at on a web page. But early on, it became difficult even theoretically to justify how that would work without some additional information. Multiple design ideas were contemplated, such as supplementing eye movement data with cursor (mouse) movement. Eventually, we decided to use head movement in tandem with the eye movement to determine exactly where the user was looking.

This was accomplished by using a Deep Neural Network (or DNN) model that had been pretrained to work on OpenCV. It was summarized through a Caffe file — a deep learning framework "skeleton" file that instructs OpenCV on how to create the DNN model. An example of the general structure of a DNN is shown below in Figure 2.

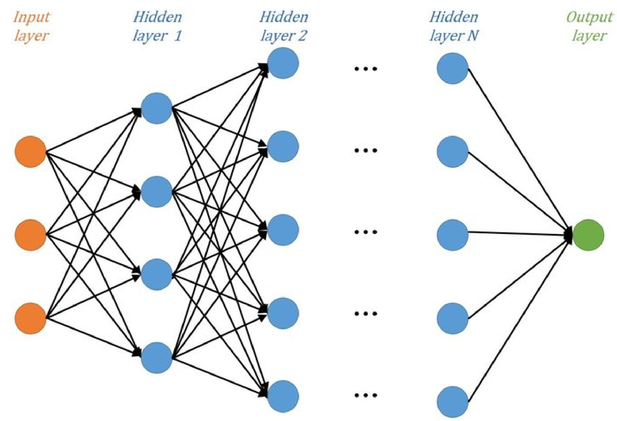


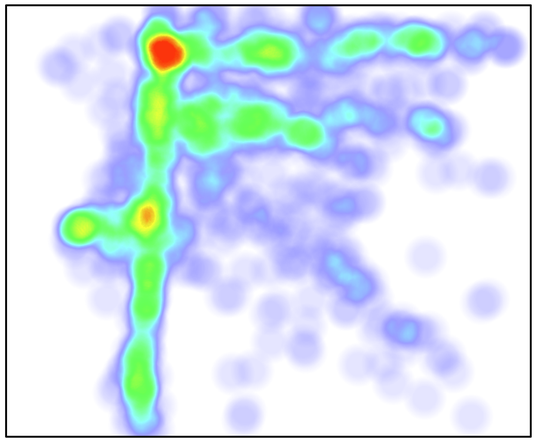
Figure 2. OpenCV Deep Neural Network (DNN) Model

This model was successful able to calculate and track — in real time — the movement of a users head and eyes, outputting a stream of "coordinates."

By cross-referencing these coordinates with a standardized model of the laptop screen, we were able to track exactly where experiment subjects were looking at through statistical data visualization done in through the Seaborn library.

Preliminary Results

These heat maps have been generated from an eye/head tracking evaluation of a very limited number of test subjects. It is important to note that this is not a representative sample, but rather extremely preliminary results that can be used as guidance for further research.



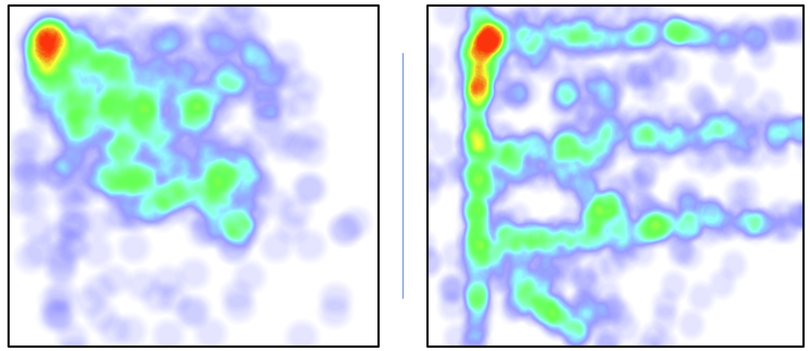


Figure 3. Generated Heatmaps

Interpreting these plots, users scan web pages in "F" or "E" shaped combinations. What does this mean? Users are extremely likely to scan web pages by reading the top line first — the address bar. Then they go down the page a little, read it, then down the page a little, read, and repeat. Eventually attention starts to drift but this pattern remains constant.

We plan on conducting on more testing with a greater sample size and diversified candidates to broaden our understanding of this extremely interesting and insightful experiment.

Thank you to the UROC program for organizing this fantastic experience!

sugandhi@iu.edu