Expanding the reach of eye tracking in psychology using webcam data collection

Anatolii Evdokimov[1], Arryn Robbins[1]

[1]University of Richmond, VA

Background

- Eye tracking has become an **important tool for research** in many subdisciplines of psychology, including cognitive, developmental, and clinical psychology.
- Eye tracking often requires **specialized equipment** (cameras, computer screens, chin rests) and software to be set up stationary in the lab which creates bottlenecks in data collection.
- Webcam eye tracking is becoming more prevalent as a tool for **remote data collection** which helps **increase the reach** of the eye tracking method and **diversify the participant pool**.

Current Study

- The goal of this project is to **develop a research grade model** that can predict gaze location from webcam images.
- We also propose **a new eye movement dataset** that features gaze location information and face images collected online using participants' webcams on participant laptops and desktops.
- We trained a simple convolutional neural network serving as a **benchmark model** for comparing appearance-based eye movement datasets.

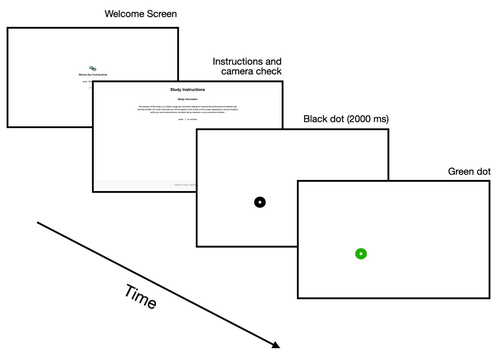
Dataset

Data Collection Task

- Dataset **collected online** in the participant browser using their webcam
- Participants are completing **a prosaccade task** (following a dot on the screen)
- Participants first see a **black dot** appear on the screen for 2000 ms and are instructed to look at it for as long as its color is black.
- Once the dot **changes color** to green, participants are instructed to press a space bar to see the next dot.

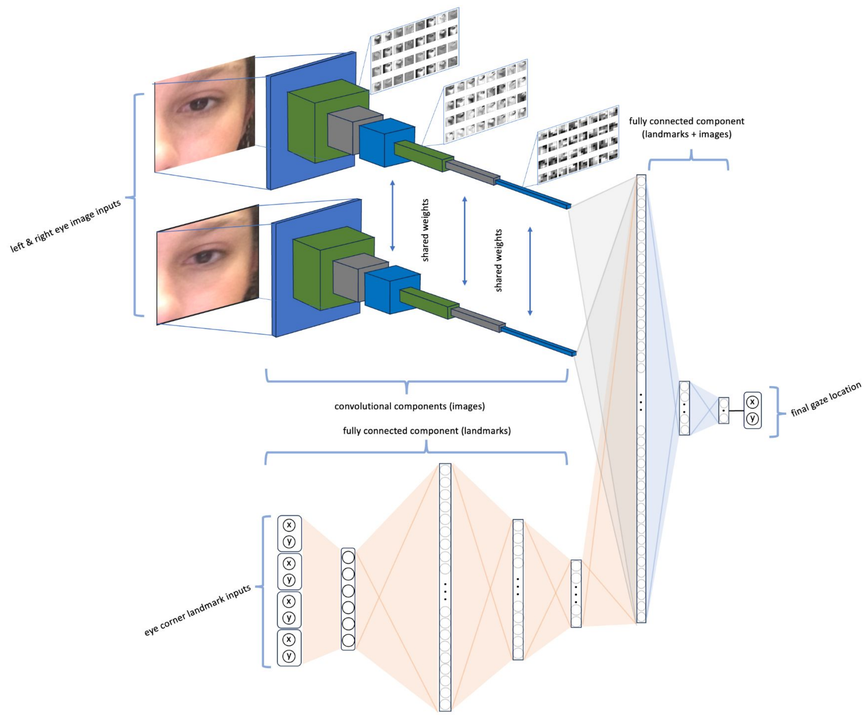
Dataset

- The **video frames** are synchronized with the dot locations and are turned into images
- The resulting dataset contains video frames, dot locations, and meta data about participants machine (screen information, browser, etc.)



Prosaccade Online Task Progression

Benchmark Model



Model Design

- We created a **benchmark CNN model** inspired by Valliappan et al. [1]
- The model takes in the **images of the left and right eye** as well as 4 **eye corner landmark** locations extracted using OpenCV's [2] Haar Cascades (8 landmarks total).
- The output of the model is an X and Y location of the fixation on the screen

Training

- We used **GazeCapture dataset** [3] to train the benchmark model
- Data were filtered based on the **number of recognized faces** on the image as well as the **device orientation**.

Model Results & Further Analyses

- Using the subset of GazeCapture with devices in portrait orientation, our model achieved **2.2cm accuracy** (~2 degree of visual angle) which is similar to results reported in Valliappan et al. (2022) on the model without personalization.
- We plan to use the online prosaccade data we collect **to fine-tune the model we trained** on GazeCapture to adapt our model for gaze prediction on laptop and desktop screens.
- We also plan to use it **to compare our dataset's performance** to GazeCapture since it is the same model architecture.
- In the future, we plan to create a new eye gaze prediction model based on the **transformer architecture** to further increase the accuracy of prediction.

References

[1] Heck, M., Becker, C., & Deutscher, V. (2023). Webcam Eye Tracking for Desktop and Mobile Devices: A Systematic Review. Proceedings of the 56th Hawaii International Conference on System Sciences, 6820–6829.
[2] Valliappan, N., Dai, N., Steinberg, E., He, J., Rogers, K., Ramachandran, V., Xu, P., Shojaeizadeh, M., Guo, L., Kohlhoff, K., & Navalpakkam, V. (2020). Accelerating eye movement research via accurate and affordable smartphone eye tracking. Nature Communications, 11(1), 4553. https://doi.org/10.1038/s41467-020-18360-5
[3] Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools.