Smartphone Software Retina

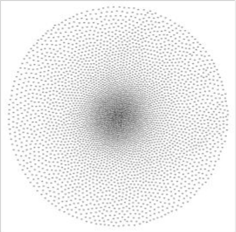
Ryan Wong

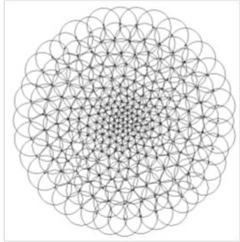
Introduction & Motivation

- Many deep learning architectures use the raw image data for training of the deep neural networks.
- A biologically inspired method reduces the memory requirements while increasing their invariance to scale and rotation changes.
- Robotic vision systems require an efficient approach detect and analyse points of interest.

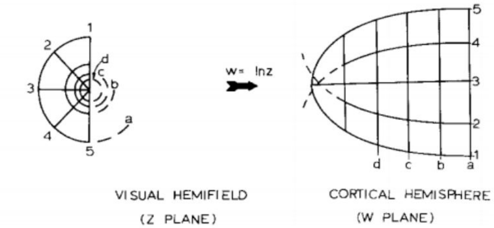
Background

Back Projected Images - foveated images generated by mapping the *receptive fields* onto the original image plane





Cortical Transformed Images - Mapping of the receptive field centres onto a new ‘cortical’ space by performing a *forward warp*



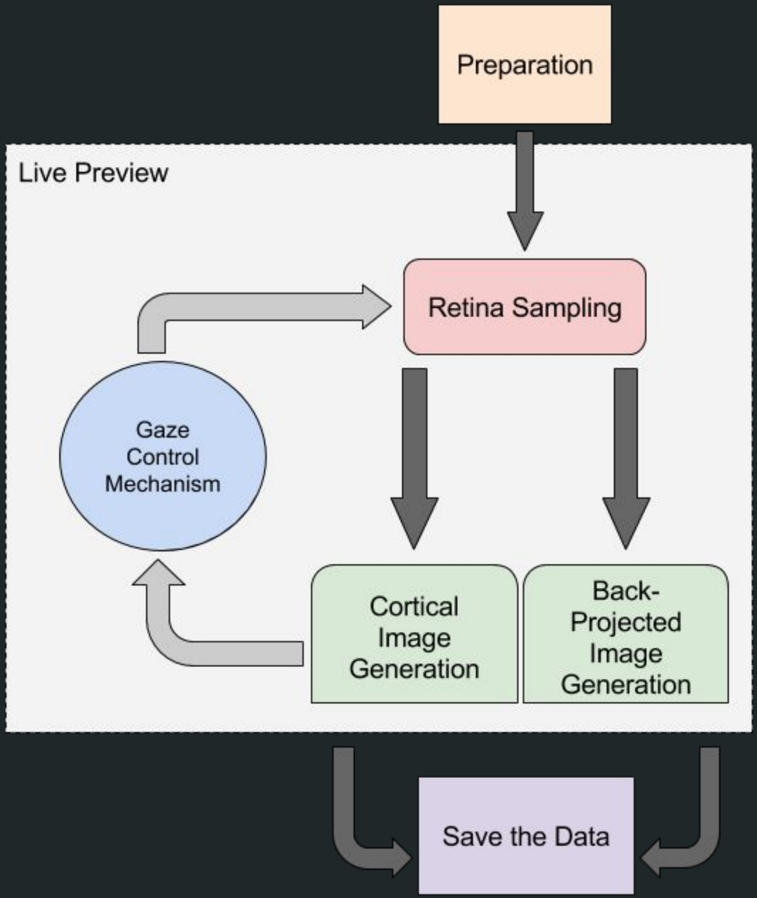
Objectives

1. A live preview of the retinal image transforms of each frame captured by the smartphone video camera.

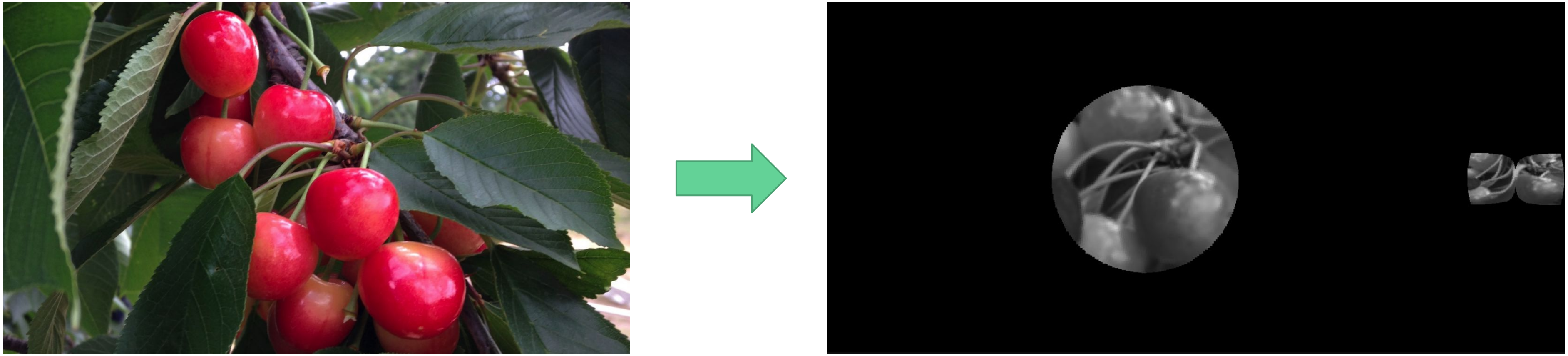
2. Gaze control mechanism to find points of interest

3. Effective recording mechanism

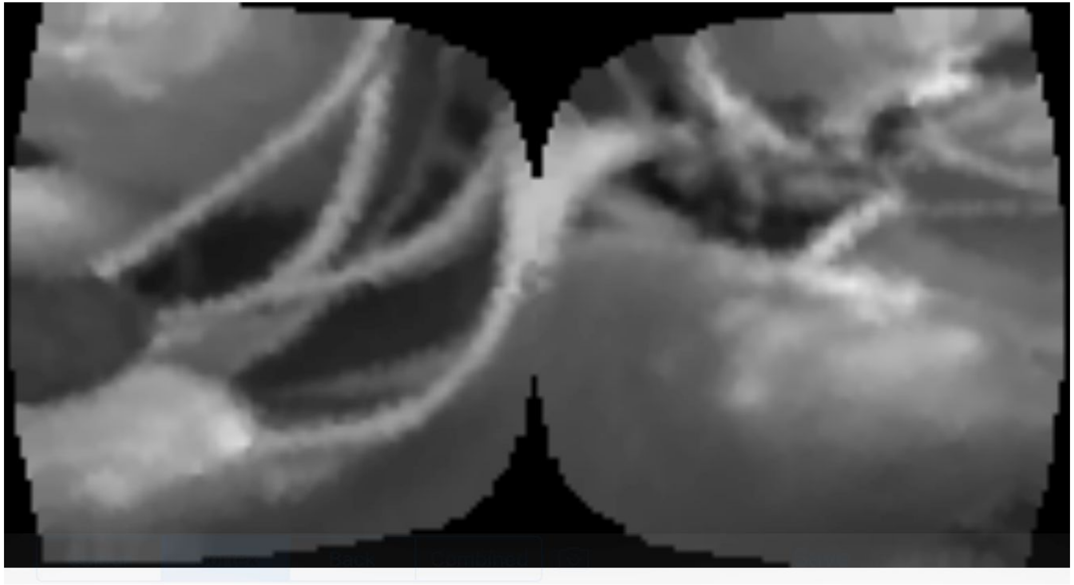
Architecture

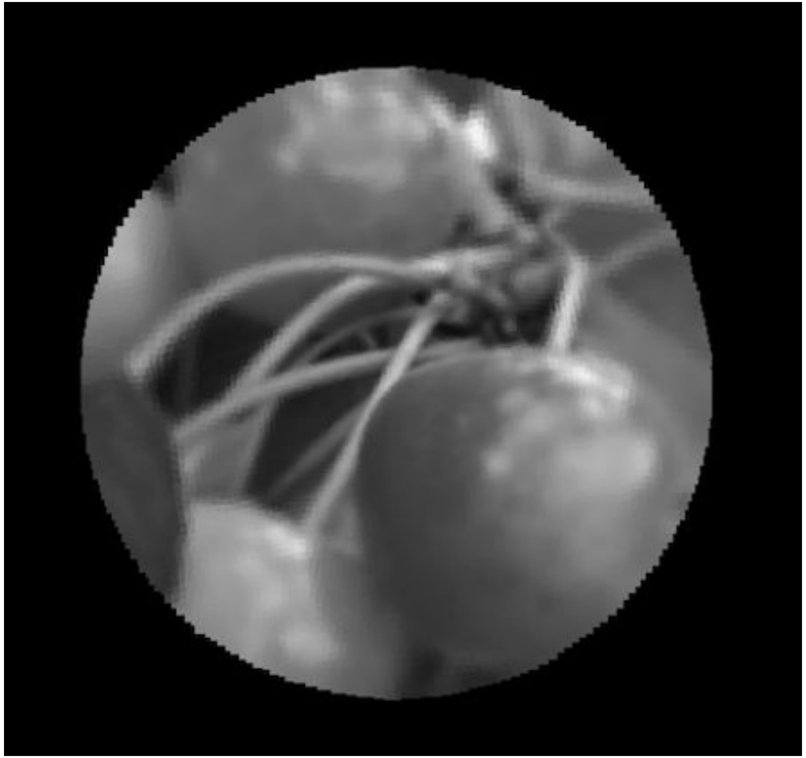


Retinal Sampling

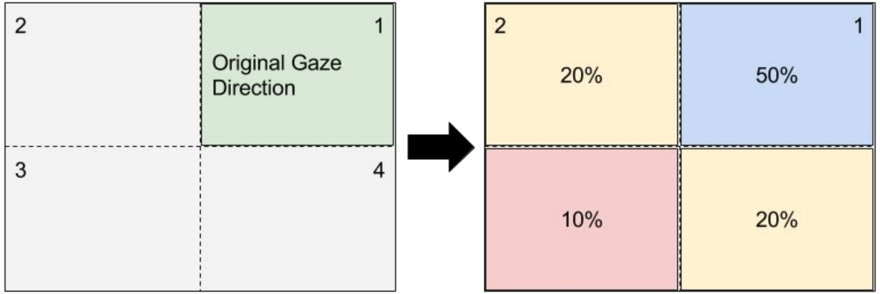


Cortical and back-projected images



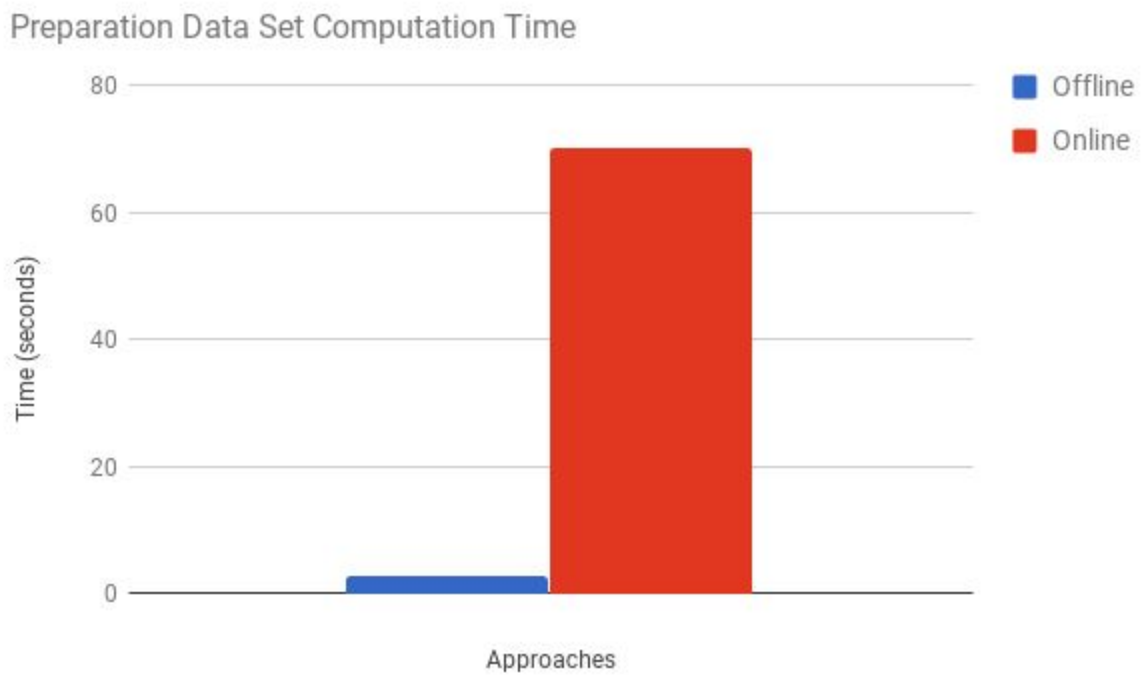


Gaze Control Mechanism

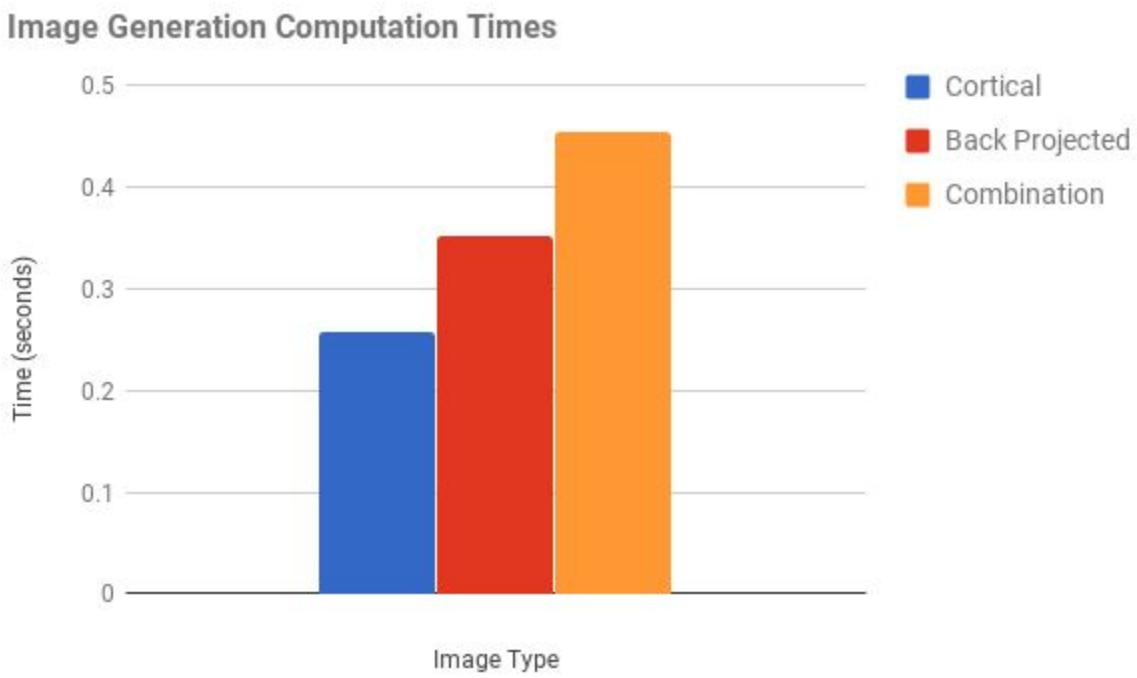


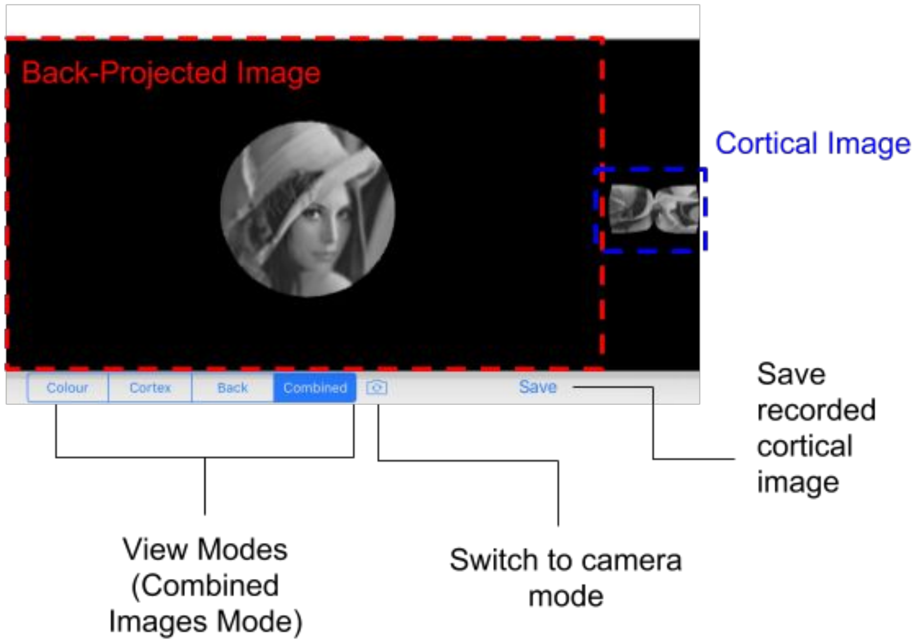
Results

Preparation



Retina Transformed Image Generation





iOS Application

Conclusion of Master’s Project

Image Processing

Data Acquisition

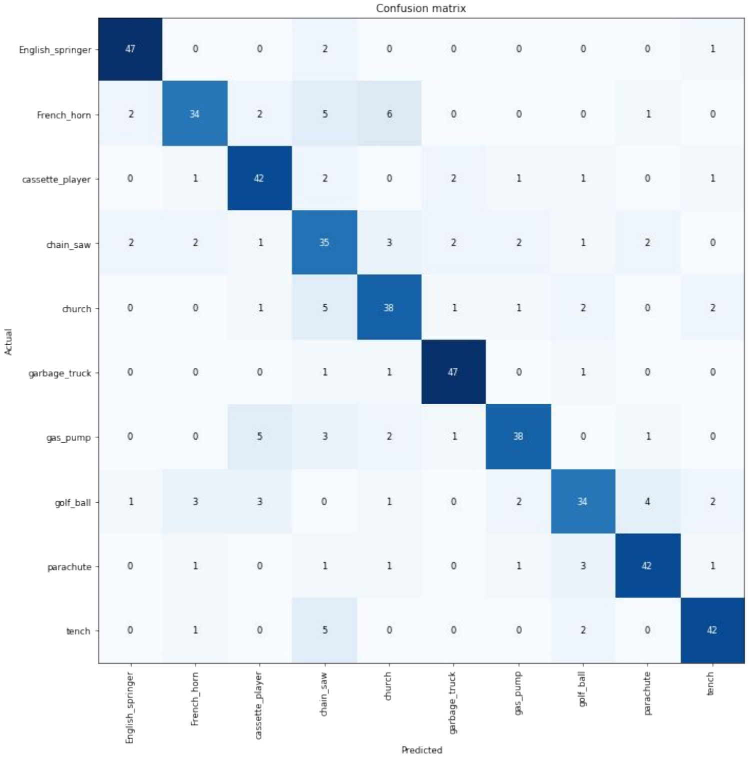
Future Research

Research Question: Can we train existing image classification deep learning architectures with the cortical images as input and achieve similar results to the original image dataset?

- Imagenette (10 classes)
- Squeezenet (Mobile neural network architecture)
- Same hyperparameters
- No data augmentation

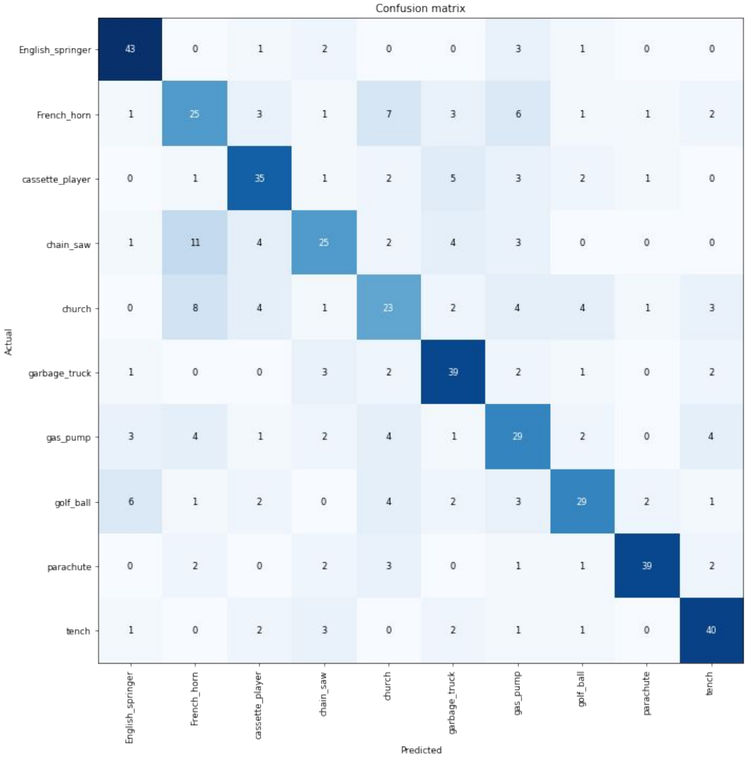
Dataset

Original Dataset



79.8%

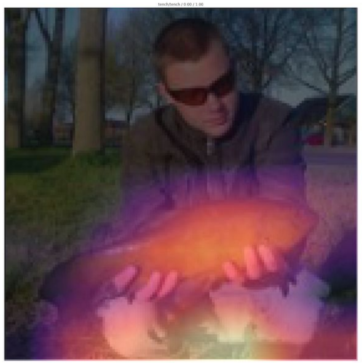
Cortical Transformed Dataset



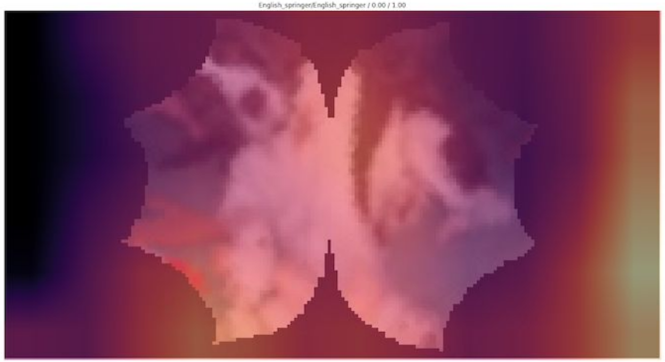
69.2%

Analysis

Original Dataset



Cortical Transformed Dataset



Discussions & Conclusions

- Results around 10% worse for the cortical transformed images.

- Current neural network architectures do not seem suitable

- Requirement of identifying a focal point before applying transform

Any Questions?

References

1. Schwartz, Eric L. "Spatial mapping in the primate sensory projection: analytic structure and relevance to perception." Biological cybernetics 25.4 (1977): 181-194.
2. Balasuriya, L. S., and J. P. Siebert. "An artificial retina with a self-organised retinal receptive field tessellation." Proceedings of the AISB 2003 Symposium: Biologically Inspired Machine Vision, Theory and Applications, Aberystwyth, UK. 2003.
3. Selvaraju, Ramprasaath R., et al. "Grad-cam: Visual explanations from deep networks via gradient-based localization." Proceedings of the IEEE International Conference on Computer Vision. 2017.