Initial Plan

Machine learning of where people look in images

Author: Shing Chung Lee
Supervisor: Dr. Hantao Liu
CM3203: One Semester Individual Project
Total Credits: 40

Description

Do you know where you are looking at while watching online videos? Do you know where the computers think you are looking at? Modelling visual saliency - predicting where human eyes pay attention to in visual content - has been a very active research area over the past few years in both academia and industry. This project aims to analyse visual attention data in emerging applications in image and vision computing, such as computer rendered images, high-dynamic-range (HDR) imaging, medical imaging, and mobile.

Human vision is a complex system that allocates attention to their visual input. Since the brain is so highly attuned to processing visual information, most human-computer interfaces rely heavily on the capabilities of the human visual system. [1] As such understanding and attempting to model this system can greatly assist humans in various computer vision tasks. With increasing computational power, machine learning has become a popular and useful tool in analysing and modelling visual saliency. [2]

This project aims to creating one (or more) machine learning model for modelling visual saliency, then using assessing the model(s) effectiveness with different datasets like computer rendered images, medical imaging, mobile and web interfaces.

Aims and Objectives

1. **To understand neural networks and visual saliency.**

Researching materials and references that could be used in the final report. Learn the tools and technologies that will be used in this project and declare which ones to use.

2. **To be able to implement machine learning model(s) for modelling visual saliency.**

Understand and be able to implement current available models, and be able to implement own model (either from scratch or making use of other models, depending on the difficulty). If time and resources allow, create multiple models for comparison.

3. **Analyse the models created.**

    1. Compare the effectiveness of different models in modelling visual saliency

    2. Investigate the effectiveness of different models in modelling visual saliency of different datasets, such as: computer rendered images, medical imaging, mobile and web interfaces.

Work Plan

Overview

Week	Date	Task
1	2021-02-01 - 2021-02-07	Research, Gather resources and tools
2	2021-02-08 - 2021-02-14	-
3	2021-02-15 - 2021-02-21	Learn to use tools
4	2021-02-22 - 2021-02-28	Implement current available models
5	2021-03-01 - 2021-03-07	Implement VGG-16 model
6	2021-03-08 - 2021-03-14	-
7	2021-03-15 - 2021-03-21	**Implement other models (ResNet50)**
8	2021-03-22 - 2021-03-28	-
Easter	2021-03-29 - 2021-04-18	Easter Break, Finish unfinished work, **Create more models**, **Create web application modelling visual saliency for websites**
9	2021-04-19 - 2021-04-25	Evaluation of solution, comparison of models
10	2021-04-26 - 2021-05-02	Evaluation of solution, comparison of models, Write up for final report
11	2021-05-03 - 2021-05-09	Write up for final report
12	2021-05-10 - 2021-05-13	-
End of Project	2021-05-14	Deadline for final report

* Tasks in **bold** will be done with enough time and resources.

Important dates

2021-02-08 Submit initial plan

2021-05-14 Submit final report

Milestones

Research, Gather resources and tools (2 weeks)

Research about the project to familiarise with the topic and plan for future tasks through the initial plan.

Gather resources and tools, like training datasets, libraries and language environments for future implementation.

Learn to use tools (1 week)

To be able to use the tools researched and be able to implement machine learning models through learning from online tutorials and books.

Implement current available models (1 week)

To be able to implement some of the pre-trained models available on the internet, this will provide a standard for how effective the models are, and also give insight on how to implement own model.

Implement own model (2-4 weeks)

Using tools and knowledge gained from earlier milestones, create own machine learning models. Depending on the time and resources needed, more or less models may be made.

Current plan is to create a VGG-16 and a ResNet50 model, using the Python programming language.

(Optional) Create web application

Create a practical application for the models, a web application that takes a url as input and creates a saliency map with the models as output.

Evaluation of solution, Comparison of models (2-3 weeks)

Test the different models created during the project, comparing the effectiveness of each model against each other and the effectiveness of each model with different datasets.

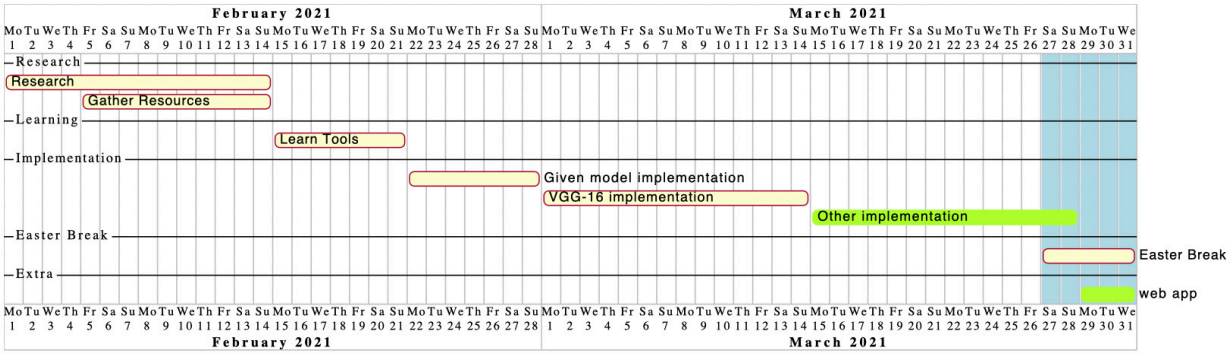
The results will be formatted to be easier to read and visualise. Some data analysis will be conducted to compare the models. And the results will be used in the final report

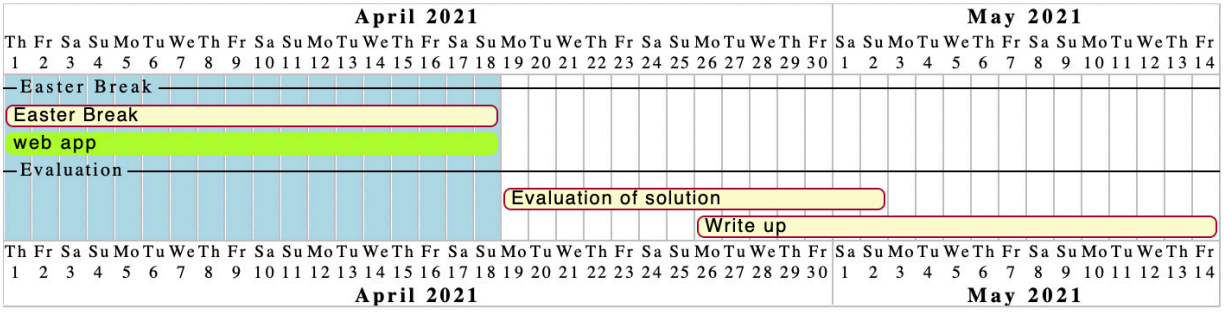
Writing up final report (2-3 weeks, throughout entire project)

Report writing, throughout the project important achievements and milestones will be written down. And further writing will be made at the end of the project, including

finalising in words the results obtained in the evaluation milestone. A project diary and meeting minutes with the supervisor will be used throughout the project to help facilitate this process.

Gantt Chart





* Tasks in GreenYellow will be done with enough time and resources.

References

[1] Matzen, L. E., Haass, M. J., Tran, J., & McNamara, L. A. (2016). Using eye tracking metrics and visual saliency maps to assess image utility. Electronic Imaging, 2016(16), 1-8.

[2] Zhao, Q., & Koch, C. (2013). Learning saliency-based visual attention: A review. Signal Processing, 93(6), 1401-1407.