

# Autophotographer: aesthetic judgements from image data

---

Report Name	Project Outline
Author (User Id)	Oscar Pocock (osp1)
Supervisor (User Id)	Hannah Dee (hmd1)
Module	CS39440
Degree Scheme	G401 (Computer Science (with integrated year in industry))
Date	February 11, 2022
Revision	1.0
Status	Release

---

## 1 Project description

The main aim of the "Autophotographer" project is to develop a program which will select aesthetic pictures from a given video file. The user will give the program a video file, the program will then process the file, and select the most aesthetic frames using data analysis and machine learning techniques.

The program will be written in Python and will utilise machine learning libraries (OpenCV [1], PyTorch [2], TensorFlow [3]) to help select the most aesthetic frames in the last stage of the process. The program will initially be a CLI, if there's time I would like to develop a cross-platform GUI in something like Flutter [4] to improve accessibility.

A Convolution Neural Network (CNN) will be trained with an existing dataset (AVA [5] or AADB [6]) and used to predict and select the most aesthetic frames based on the strength of certain aesthetic features detected in the image. The majority of the program will only use the CNN once the original set of input frames has been reduced using conventional non-machine learning techniques.

If there's more time, I would also like to incorporate a machine learning editing system, which will attempt to enhance certain detected features automatically by editing the picture. For example, cropping to increase a match to the 'rule of thirds' or blurring the background via edge detection to increase a shallow depth of field. This would be an optional setting and would be non-destructive editing.

The final goal of the project is to provide a tool that minimises the manual process of selecting aesthetically significant pictures from a video. This can help improve automatic thumbnail selection for online videos, or create a photo album from a video for a different way to experience the same memory. Personally, I'm very interested in using this technology as a way to aid people with certain motor and sensory conditions/disorders, who might find it difficult/impossible to take aesthetic pictures naturally.

## 2 Proposed tasks

- **Context specific research into machine learning** - I will need to research further into CNNs and how to build and finetune them, this should be done in tandem to learning a machine learning library. Work on this should start during the 3rd week.
- **Setup tools for development** - I will need to setup tools to help me practise development methodologies. This will help me keep track and plan ahead. This should be finished by the start of the 3rd week.
- **Learn machine learning libraries** - For the machine learning aspect of the project I will need to learn how to use machine learning libraries. I will use these to process a reduced set of the frames and determine how aesthetic each of them are. This should be done in tandem to researching more about machine learning and CNNs. This should start in the 3rd week.
- **Create a CLI tool in Python** - As I'm working with computer vision, I'll be writing the program in Python which will be the most important task. Work has already started on this and this should be expected to be finished in the last week of the project.

- **Enhance aesthetic properties of selected photos** - If I have more time, I will also look at using the knowledge obtained through machine learning to improve the selected frames by enhancing their most significant aesthetic properties. This should only start once the original work has been completed and should be finished by the end of the project.

### 3 Project deliverables

- **CLI program written in Python** - A CLI Python program, with built in help that takes a video file as an input and outputs image files to a specified output folder.
- **User and Developer Manual** - A manual for users to help them understand how to use the program and tune the parameters to their preference. Another manual for developers will be written to help people extend or make modifications to the program.
- **Final report** - A final report that discussed the initial problem, details of the technical work achieved and it's development process, and a critical evaluation and insight into the project as a whole.

## Annotated Bibliography

- [1] Intel Corporation, Willow Garage, and Itseez, "OpenCV." [Online]. Available: <https://opencv.org>

An open source library used for real-time computer vision.

- [2] Adam Paszke, Sam Gross, Soumith Chintala, and Gregory Chanan, "PyTorch." [Online]. Available: <https://www.pytorch.org>

An open-source machine learning framework developed by Facebook (based on the Torch library)

- [3] Google Brain Team, "TensorFlow." [Online]. Available: <https://www.tensorflow.org/>

Open-source library for machine learning and AI developed by Google.

- [4] "Flutter - Build apps for any screen." [Online]. Available: <https://flutter.dev/>

- [5] N. Murray, L. Marchesotti, and F. Perronnin, "AVA: A large-scale database for aesthetic visual analysis," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, June 2012, pp. 2408–2415, iSSN: 1063-6919.

A paper that discusses the creation of the AVA dataset. It's composed of 250 thousand aesthetic images collected from a photography competition website along with grades for aesthetic features.

- [6] S. Kong, "Photo Aesthetics Ranking Network with Attributes and Content Adaptation," Jan. 2022, original-date: 2016-06-05T06:08:10Z. [Online]. Available: <https://github.com/aimerykong/deepImageAestheticsAnalysis>

A tool that uses the AADB dataset to predict the aesthetics of a given picture using a CNN.

- [7] H.-J. Lee, K.-S. Hong, H. Kang, and S. Lee, "Photo Aesthetics Analysis via DCNN Feature Encoding," *IEEE Transactions on Multimedia*, vol. 19, no. 8, pp. 1921–1932, Aug. 2017, conference Name: IEEE Transactions on Multimedia.

This paper discusses the use of a Convolution Neural Network to predict how aesthetic a given picture is.

- [8] Fing, "AVA Dataset," Feb. 2022, original-date: 2016-11-13T02:20:32Z. [Online]. Available: [https://github.com/imfing/ava\\_downloader](https://github.com/imfing/ava_downloader)

A project with tools required to build the AVA dataset.

- [9] "Image Quality Assessment," Feb. 2022, original-date: 2018-06-12T14:46:09Z. [Online]. Available: <https://github.com/idealo/image-quality-assessment>

An open source tool developed by Idealo which uses a CNN to predict the aesthetic value of a given image.