

Project Outline and Plan

|  |
| --- |
| School of Computing  Faculty of Engineering AND PHYSICAL SCIENCES |

PObject Detection Website Based on Deep Learning

Yunjia Feng

**<BSc Computer Science>**

The candidate confirms that the work submitted is their own and the appropriate credit has been given where reference has been made to the work of others.

I understand that failure to attribute material which is obtained from another source may be considered as plagiarism.

(Name of student) <Yunjia Feng>

© <2021> The University of Leeds and <Yunjia Feng>

# 1. Introduction

<A brief overview of your project suitable for a non-specialised audience. You may mention background and motivation, but should avoid technical terms as far as possible. 1-2 paragraphs should be sufficient.>

Object detection has always been of great significance in the field of computer vision whose main objective is to enable the computer to accurately classify the objects in a given picture or video and find the position of each object. In recent years, many computer vision researchers at home and abroad have developed a large number of excellent object detection neural network algorithms, including Faster R-CNN, SSD, YOLO. The goal of this project is to select appropriate object detection algorithms and data sets and train the deep learning model, then develop a website for users which allows users to complete object detection tasks easily.

From the perspective of research, the significance of object detection is that object detection is one of the fundamental tasks in the field of computer vision, which is also the basis of many other high-level tasks: image classification, face recognition, target tracking, pedestrian re-recognition. Meanwhile, there is a large number of well-known research teams both at home and abroad, had been focused on the field of object detection, for instance, MIT Computer Science and Artificial Intelligence Laboratory, Stanford Computer Vision Lab, National Laboratory of Pattern Recognition of Chinese Academy of Sciences, LAMDA Institute of Nanjing University.

While from the perspective of the application, object detection has shown a wide range of practical usages: face detection technology, pedestrian detection technology applied in video surveillance, entrance and exit statistics, traffic sign detection technology, vehicle detection technology applied in aided driving, automatic driving. At the same time, major technology companies, for example, Microsoft, Google, Ali, and Baidu, have also spent a lot of manpower and material resources to explore the object detection field, which indicates the significance and prospect of object detection.

## 2 Aims, Objectives and Deliverables

## 2.1 Aims

<A brief summary of the primary aims of this project. Typically 1-3 sentences.>

## 2.2 Objectives

<A list of what you want to achieve *by the end of the project* - note this means ’learn how to…', ‘research into…` are *not* objectives, as they are intermediate milestones rather than final goals. All objectives should be measurable, *i.e.* it should be possible to provide evidence to confirm whether or not they have been achieved. 3-5 objectives is typical.>

## 2.3 Deliverables

<A list of what you will hand in at the end of the project. This will include the final report (possibly spread across multiple deliverables, if that makes sense for your project), code (possibly more than one version), and so on. Ideally the deliverables should be cross-referenced to the objectives. 2-3 deliverables is typical, but there can be more depending on the nature of the project.>

The proposal deliverable of the project should be a software form, which consisted of two main parts, a deep learning module, and a web module.

The deep learning module should include Pytorch based Yolo-v5 algorithm implementation, weights that trained from datasets VOC2007, code that used to transform dataset format, encapsulated API that process image and videos for object detection tasks.

The web module will contain both front-end and back-end parts. The front-end page will be implemented based on the Vue framework, providing interaction for users to upload images or videos and download object detection results. While the back-end will provide processing routers to call deep learning API, and also record corresponding logs during each process (time, process speed, size).

Additionally, a thesis will be constructed as a deliverable of this project, which including the background research, literature in the area and related work, objectives, experiment methods, evaluation, conclusion, reference, in order to demonstrate the whole process and experient details during the project.

## 3. Project Plan

<Provide a plan for the full project, from when you started until the submission of the final report. This should discuss the key stages of your project.>

9.20 - 10.8 Complete the project outline and plan, discuss the topic and laboratory usage in the project with the supervisor.

10.9 - 10.17 Object Detection background research:

(1) Basic background research about the general deep learning field on its significance, current application, bottleneck, and future development.

(2) Background research on object detection and related tasks: classification, location, instance segmentation, and read relevant papers.

(3) Research on object detection algorithms, for instance, Yolov5, Faster R-CNN (refer to R-CNN, Fast R-CNN as the basis), and SSD, then summarize the advantages and disadvantages of each algorithm.

(4) Investigate web front-end back-end technology, Python Flask backend framework, and Vue front-end framework.

(5) Research on object detection dataset, compare VOC, COCO, ImageNet, and select one data set for training.

(6) Investigate basic indicators and benchmarks related to object detection, read relevant papers, and acquainted with the process for subsequent model evaluation.

10.18 - 10.25 Prepare for the Thesis Proposal Defense:

(1) Write the Thesis Proposal Report.

(2) Prepare the Thesis Proposal Defense slide.

(3) Prepare the Thesis Proposal Defense presentation.

10.25 - 11.30 Determine the project structure and technology stack, built the development environment, and complete the basic demo:

(1) The basic construction of front and back end modules and deep learning framework (technology stack: Vue + Flask + Pytorch + Yolov5)

(2) Build Anaconda deep learning development environment and install related dependent libraries.

(3) Implement data set format conversion code.

(4) Complete the demo based on the pre-training model, test the object detection accuracy of pictures, videos, and cameras.

12.1 - 12.31 Start training on multiple pre-training weights:

(1) Record loss, AP, recall, and other important indicators of each training.

(2) Visualize the training results and record the continuous trend of loss, AP and Recall.

(3) Record the processing speed of each image predicted by the model.

(4) Evaluate the training results, use test sets to evaluate the training results, and formulate improvement plans according to the situation.

(5) Initial the setup of the front-end page(using Node.js and Vue framework).

(6) Write Project Outline and Plan.

1.1 - 1.31 Continue the remaining development tasks:

(1) Develop and beautify the web page, add status hints and other information to improve user experience.

(2) Complete the interface for accessing web pages at the back end, including uploading and downloading pictures.

(3) Complete the deep learning Yolo-v5 image object detection module and encapsulate the image object detection processing interface.

2.1 - 2.28 Continue remaining development tasks and thesis writing:

(1) Complete the front-end upload and download displaying layer for object detection based on the video.

(2) Implement a back-end interface to complete video detection.

(3) Complete the deep learning Yolo-v5 object detection module based on video and encapsulate the object detection based on the video processing interface.

(4) Design charts based on previous work (training and test samples, comparison with benchmark)

(5) Begin to write the abstract, introduction, and literature review of the thesis.

3.1 - 3.15 Continue to write the thesis and prepare for the mid-term examination:

(1) Continue to write the research method, experimental process, and conclusion part of the thesis.

(2) Prepare the slide for mid-term examination and defense.

(3) Prepare demo of mid-term examination project.

3.16 - 3.31 Roughly complete the thesis writing

(1) Complete writing the final parts: evaluation and citation of the thesis.

(2) Polish and modify the language and structure of the thesis.

4.1 - 4.20 Prepare for Thesis Defense

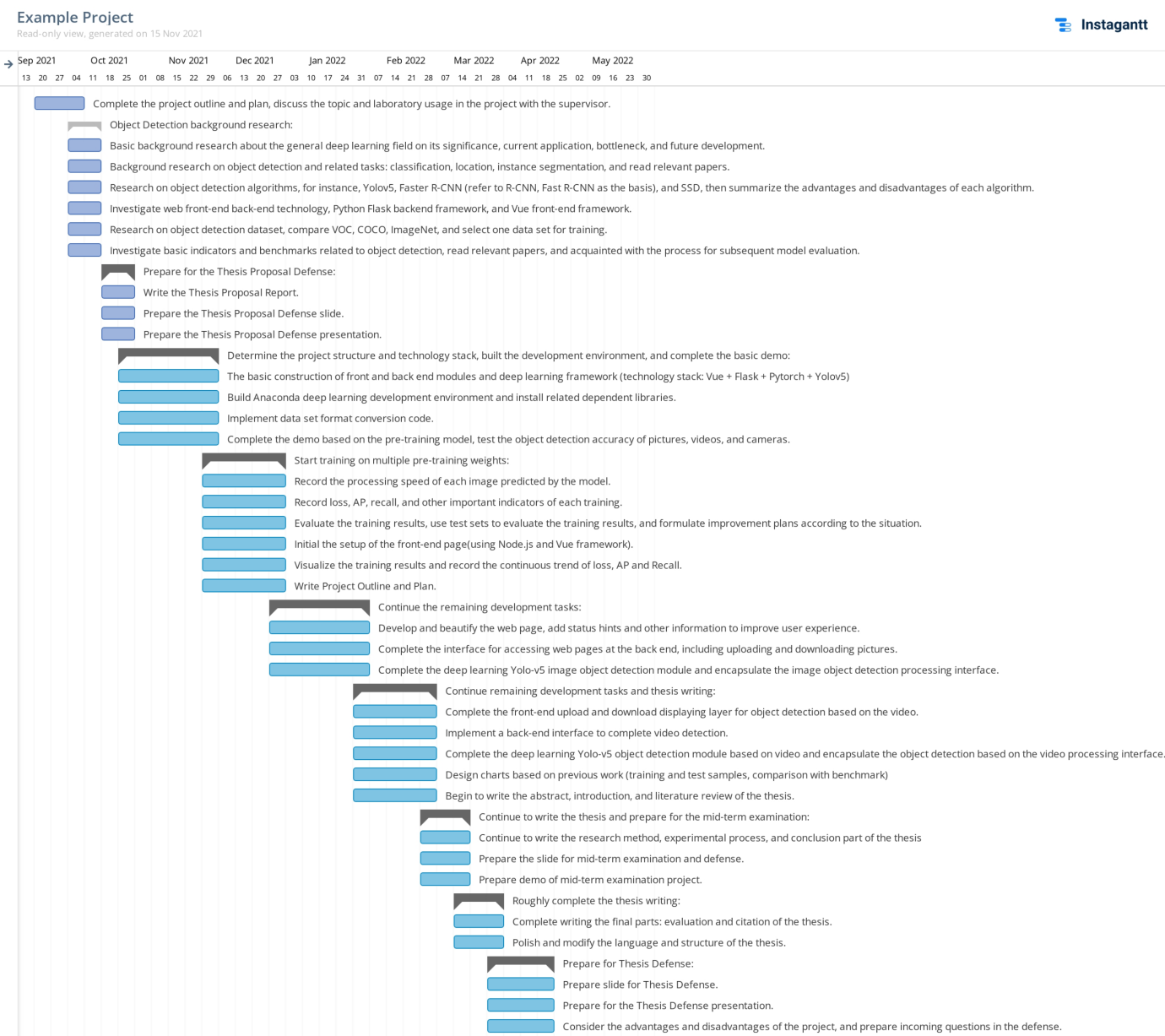
(1) Prepare slide for Thesis Defense.

(2) Prepare for the Thesis Defense presentation.

(3) Consider the advantages and disadvantages of the project, and prepare incoming questions in the defense.

## 3.1 Timeline

<A graphical description of your plan, often as a Gantt chart.>



## 4. Risk Mitigation

<Identify risks to your project, and what you would do if they arose.>

Training needs computer: 16G, displaycard => server in the school, google colab?

## 5. Ethics

<If your project has ethical issues (*e.g.*gathering of user consent forms), then you should state here how you intend to address them. If there are no ethical issues then explicitly state:"There are no ethical issues for this project.">

There are no ethical issues for this project.

# List of References

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in neural information processing systems, 2012, pp. 1097–1105.

[2] Vishwakarma S, Agrawal A. A survey on activity recognition and behavior understanding in video surveillance [J]. The Visual Computer, 2012: 1-27.

[3] Jiang, B., Luo, R., Mao, J., Xiao, T., & Jiang, Y. (2018). Acquisition of localization confidence for accurate object detection. In ECCV (pp. 784–799).

[4] Hei Law and Jia Deng. Cornernet: Detecting objects as paired keypoints. In Proc. Eur. Conf. Comp. Vis., pages 734–750, 2018.

[5] Liu, Wei, et al. "SSD: Single Shot MultiBox Detector." European Conference on Computer Vision Springer International Publishing, 2016:21-37.S

[6] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg. DSSD: Deconvolutional single shot detector. arXiv:1701.06659, 2016. 1, 2, 8

[7] Z. Li and F. Zhou. Fssd: Feature fusion single shot multibox detector. arXiv preprint arXiv:1712.00960, 2017.

[8] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection

and semantic segmentation. In CVPR, 2014.

[9] S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In NIPS, 2015.

[10] Cai, Z., Vasconcelos, N.: Cascade r-cnn: Delving into high quality object detection. arXiv preprint arXiv:1712.00726 (2017)

[11] Redmon, Joseph, et al. "You Only Look Once: Unified, Real-Time Object Detection." (2015):779-788.