Project Proposal: Solving class-agnostic counting problem using deep learning.

Submitted October 26, 2024, in partial fulfillment of
the conditions for the award of the degree **BSc Hons Computer Science with Artificial Intelligence.**

Yuzhe Wu
20411994

Supervised by Jianfeng Ren

School of Computer Science University of Nottingham Ningbo China

Chapter 1

Aims and Objectives

The primary objective of this project is to address the problem of object counting using deep learning techniques. Specifically, we focus on a subtasks within object counting: Class-Agnostic Counting and plar-Free Counting. The proposed method aims to learn and estimate the number of instances of a specific object class in a given image scene using zero-shot or few-shot exemplars. Additionally, it is expected that the method will be presented at an academic conference. To achieve the main objective, it can be divided into several sub-goals:

1. Investigate and analyze the sub-tasks and related tasks, and clearly define each task.
2. Explore and evaluate existing deep learning methods to identify an efficient approach for accurate object counting.
3. Identify the key challenges inherent in the object counting task.
4. Investigate recent advancements and new frameworks in artificial intelligence for potential improvements.
5. Design and develop a novel method by addressing the identified challenges to enhance object counting performance.
6. Compare and evaluate the performance of existing methods against the proposed approach.

Chapter 2

Background and Motivation

Object counting aims to determine the number of instances of a specific object class in an image [1]. The acquired knowledge can be applied across various counting scenarios, *e.g.*, vehicles [2], crowds [3], and cells [4]. Object counting can be broadly categorized into three types, *i.e.*, **Class-Specific Counting (CSC)**, **Class-Agnostic Counting (CAC)**, and **Exemplar-Free Counting (EFC)**.

Class-Specific Counting (CSC)

CSC involves counting specific categories, such as fruits [5] and animals [6]. Solutions based on object detectors have been extensively explored for these categories. However, a major drawback of current CSC methods is that they require large annotated training datasets for each object class, which is often impractical.

Class-Agnostic Counting (CAC)

CAC responses to overcome the limitations of CSC, particularly the need for extensive annotated training datasets, making it applicable to new, previously unseen classes with minimal annotations. It counts objects based on provided visual exemplars [1, 7, 8] or text prompts [9, 10], specializing in the target object category at test time using a few user-provided exemplars.

Exemplar-Free Counting (EFC)

EFC addresses the limitations of both CSC and CAC, as exemplars may introduce sample bias and cannot cover the entire distribution. EFC counts objects without relying on exemplars, presenting a significant challenge in identifying countable objects and determining their repetitions [11, 12, 8]. It shows promise for applications in automated systems such as wildlife monitoring [13], healthcare [14], and anomaly detection [15].

Despite recent advancements, existing CAC methods [8, 16, 17] largely follow the framework proposed by [1], *i.e.*, Feature Extraction, Similarity Matching, and Density Prediction. These methods often require explicit exemplars to count similar objects, which limits their performance in dense scenarios or when dealing with highly similar objects within a single image. EFC methods, such as RepRPN-Counter, avoid the need for exemplars by generating regions through region proposals [11]; however, they still struggle to effectively differentiate countable objects from other objects.

Given the current challenges and limitations in CSC, CAC, and EFC, this work aims to explore potential improvements, specifically focusing on the CAC and EFC tasks.

Chapter 3

Project Plan

The project plan is divided into several phases to systematically design and validate a deep learning-based object counting method, focusing on Class-Agnostic Counting (CAC) and Exemplar-Free Counting (EFC) tasks:

1. **Preliminary Work:** Complete the ethics form, project plan, and project proposal.
2. **Literature Review:** Conduct an in-depth analysis of the object counting problem, clearly defining the sub-tasks involved. Identify the key challenges within each sub-task and review state-of-the-art methods. Perform an extensive literature review to explore existing models and recent advancements in object counting. Evaluate these methods to identify limitations, particularly concerning the performance of CAC and EFC tasks. Investigate novel mechanisms in other related tasks for potential application in model design.
3. **Design:** Develop a novel method to address the challenges in CAC and EFC tasks. The proposed design will focus on overcoming the limitations of existing methods. This phase will also involve creating new modules or enhancements to improve proposed paradigm, *i.e.*, *feature extraction*, *similarity matching*, and *counting accuracy*.
4. **Interim Report:** Document the findings from the preliminary work, literature review, and initial design concepts. The interim report will provide a detailed overview of the project's progress and outline the remaining tasks.
5. **Experiment:** Implement the proposed method and conduct experiments on benchmark datasets, such as FSC147 and CARPK. The experiments will focus on evaluating the performance of the proposed method compared to existing models.
6. **Conference Paper:** Prepare and submit a conference paper summarizing the progress made. The goal is to present the proposed method at a leading computer vision or artificial intelligence conference.
7. **Final Report:** Complete the final report, summarizing the project, including the research conducted, methodology developed, experiments performed, and results obtained. The final report will also provide conclusions and potential directions for future work.

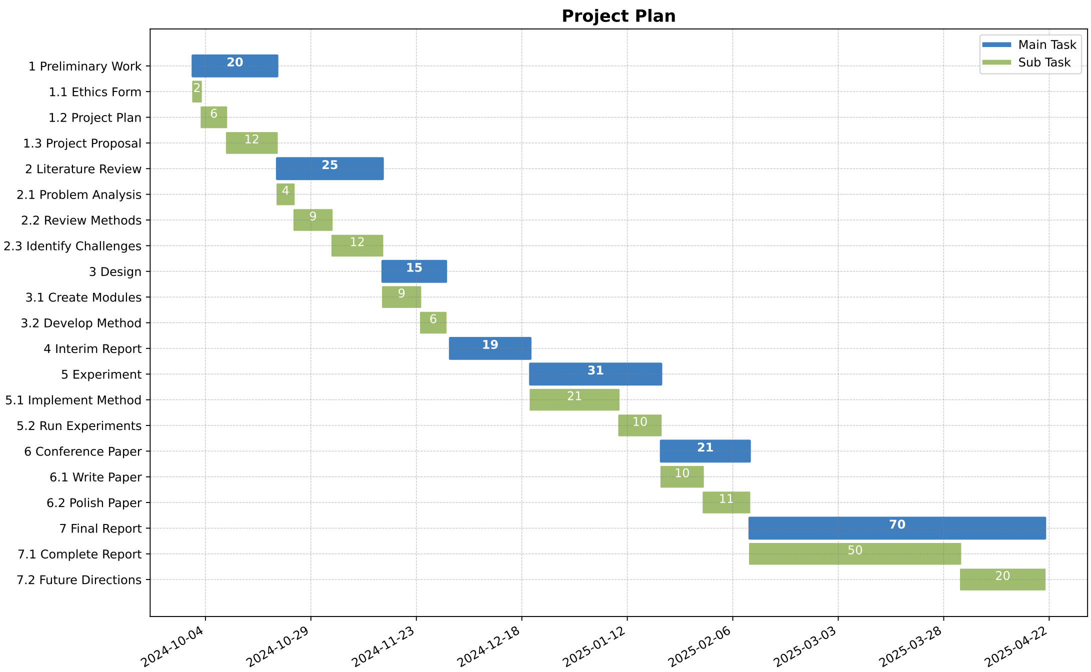

Figure 3.1: Project Plan

References

[1] Viresh Ranjan, Udbhav Sharma, Thu Nguyen, and Minh Hoai. Learning to count everything. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 3394–3403, 2021.

[2] Ersin Kilic and Serkan Ozturk. An accurate car counting in aerial images based on convolutional neural networks. *J. Ambient Intell. Humanized Comput.*, pages 1–10, 2023.

[3] Mingliang Dai, Zhizhong Huang, Jiaqi Gao, Hongming Shan, and Junping Zhang. Cross-head supervision for crowd counting with noisy annotations. In *IEEE Int. Conf. Acoust. Speech Signal Process.*, pages 1–5, 2023.

[4] Weidi Xie, J Alison Noble, and Andrew Zisserman. Microscopy cell counting and detection with fully convolutional regression networks. *Comput. Methods Biomech. Biomed. Eng. Imag. Vis.*, 6:283–292, 2018.

[5] Enrico Bellocchio, Thomas A Ciarfuglia, Gabriele Costante, and Paolo Valigi. Weakly supervised fruit counting for yield estimation using spatial consistency. *IEEE Robot. Autom. Lett.*, 4:2348–2355, 2019.

[6] Jayme Garcia Arnal Barbedo, Luciano Vieira Koenigkan, Patricia Menezes Santos, and Andrea Roberto Bueno Ribeiro. Counting cattle in UAV images—dealing with clustered animals and animal/background contrast changes. *Sensors*, 20:2126, 2020.

[7] Min Shi, Hao Lu, Chen Feng, Chengxin Liu, and Zhiguo Cao. Represent, compare, and learn: A similarity-aware framework for class-agnostic counting. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 9529–9538, 2022.

[8] Liu Chang, Zhong Yujie, Zisserman Andrew, and Xie Weidi. CounTR: Transformer-based Generalised Visual Counting. In *Brit. Mach. Vis. Conf.*, 2022.

[9] Jingyi Xu, Hieu Le, Vu Nguyen, Viresh Ranjan, and Dimitris Samaras. Zero-shot object counting. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 15548–15557, 2023.

[10] Seunggu Kang, WonJun Moon, Euiyeon Kim, and Jae-Pil Heo. Vlcounter: Text-aware visual representation for zero-shot object counting. In *Proc. AAAI Conf. Artif. Intell.*, volume 38, pages 2714–2722, 2024.

[11] Viresh Ranjan and Minh Hoai Nguyen. Exemplar free class agnostic counting. In *Proc. Asian Conf. Comput. Vis.*, pages 3121–3137, 2022.

[12] Michael Hobley and Victor Prisacariu. Learning to Count Anything: Reference-less Class-agnostic Counting with Weak Supervision. In *IEEE Conf. Comput. Vis. Pattern Recog.*, 2023.

[13] Hüseyin Gökhan Akçay, Bekir Kabasakal, Duygugül Aksu, Nusret Demir, Melih Öz, and Ali Erdoğan. Automated bird counting with deep learning for regional bird distribution mapping. *Animals*, 10:1207, 2020.

[14] Vitjan Zavrtanik, Martin Vodopivec, and Matej Kristan. A segmentation-based approach for polyp counting in the wild. *Eng. Appl. Artif. Intell.*, 88:103399, 2020.

[15] Zhikang Liu, Yiming Zhou, Yuansheng Xu, and Zilei Wang. Simplenet: A simple network for image anomaly detection and localization. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 20402–20411, 2023.

[16] Nikola Djukic, Alan Lukezic, Vitjan Zavrtanik, and Matej Kristan. A low-shot object counting network with iterative prototype adaptation. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 18872–18881, 2023.

[17] Jer Pelhan, Vitjan Zavrtanik, Matej Kristan, et al. DAVE-A Detect-and-Verify Paradigm for Low-Shot Counting. In *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 23293–23302, 2024.