COMP9517: Computer Vision

2025 Term 2

Group Project Specification

Maximum Marks Achievable: 40

The group project is **worth 40% of the total course mark**.

Project work is in Weeks 6-10 with deliverables due in Week 10.

Deadline for submission is Friday 8 August 2025 18:00:00 AET.

Instructions for online submission will be posted closer to the deadline.

Refer to the separate marking criteria for detailed information on marking.

Introduction

The goal of the group project is to work together with peers in a team of 5 students to solve a computer vision problem and present the solution in both oral and written form.

Group members can meet with their assigned tutors once per week in Weeks 6-10 during the usual consultation session hour to discuss progress and get feedback.

Description

Forests are an essential part of the natural environment, serving as habitats for many wildlife species and regulating local and global ecosystems. Unhealthy forests can cause various natural disasters and pose a major risk also for human life. For example, too much dead wood contributes to the rapid spread of wildfires. Thus, it is important to monitor the health of forests. In recent years, technologies such as remote airborne sensing or satellite imaging have enabled large-scale forest monitoring without human involvement. This has caused a growing need for computer vision methods to analyse the data produced by these technologies.

<u>Task Statement</u>

The goal of this group project is to develop and compare different computer vision methods for segmenting standing dead trees in aerial images of forests.

<u>Public Dataset</u>

The public dataset to be used in this group project is available from Kaggle (see references with links at the end of this document). It contains 444 aerial images of forest areas with corresponding ground-truth segmentation masks manually created by a group of forest health experts. The images contain four-band data, including near-infrared (NIR) and RGB channels, any or all of which can be used for the task. Further details about the dataset can be found on the Kaggle page and in the associated paper (see also the references at the end).

Method Development

Many traditional, machine learning, and deep learning-based computer vision methods could be used for this task. You are challenged to use concepts taught in the course and other techniques from literature to develop your own methods and test their performance.

At least two different methods must be developed and tested. For example, you could compare one feature extraction and machine learning method and one deep learning method, or two deep learning methods using different neural network architectures.

We do not expect you to develop everything from scratch, but we do encourage you to try different methods, or modifications of existing ones, or use more state-of-the-art methods that have not been presented in the lectures. As a general guideline, the more you develop things yourself rather than copy straight from elsewhere, the better. In any case, always cite your sources, including all papers, tools and repositories you have used.

Training and Testing

If your methods require training (that is, if you use supervised rather than unsupervised segmentation approaches), you must ensure that the training images are not also used for testing. Even if your methods do not require training, they may have hyperparameters that you need to fine-tune to get optimal performance. In that case, too, you must use a separate training and test set. Using (partly) the same data for both training/fine-tuning and testing leads to biased results that are not representative of real-world performance.

A standard approach is to do an 80-20 split of the dataset. This means randomly selecting 80% of the images for training and keeping the remaining 20% for testing.

To assess the performance of your methods, compare the segmented regions quantitatively with the corresponding manually annotated masks by calculating the mean intersection over union (IoU), also known as the Jaccard similarity coefficient (JSC). Furthermore, compare the computation times needed for training and testing of the different methods.

Results and Discussion

Show the quantitative results in your video presentation and written report (see deliverables below). Also show representative examples of successful segmentations as well as examples where your methods failed (no method generally yields perfect results). Give some explanation why you believe your methods failed in these cases. Also, if one of your methods clearly works better than the other(s), discuss possible reasons why. Finally, discuss some potential directions for future research to further improve segmentation performance.

Practicalities

The dataset is less than 300 MB in total, so method training and testing should be feasible on a modern desktop or laptop computer. If more computing resources are needed, you could consider using the free version of Google Colab. Otherwise, you are free to use only a subset of the data, for example 75% or even just 50%. Of course, you can expect the performance of your methods to go down accordingly, but as long as you clearly report your approach, this will not negatively impact your project mark.

Deliverables

The deliverables of the group project are 1) a video presentation, 2) a written report, and 3) the code. The deliverables are to be submitted by only one member of the group, on behalf of the whole group (we do not accept submissions from multiple group members).

Video

Each group must prepare a video presentation of at most 10 minutes (any content beyond that will be ignored). The presentation must start with an introduction of the problem and then explain the used methods, show the obtained results, and discuss these results as well as ideas for future improvements. For this part of the presentation, use PowerPoint slides to support the narrative. Following this part, the presentation must include a demonstration of the methods/software in action. Of course, some methods may take a long time to compute, so you may record a live demo and then edit it to stay within time.

All group members must present (points will be deducted if this is not the case). It is up to you to decide who presents what (introduction, methods, results, discussion, demonstration). In order for us to verify that all group members are indeed presenting, each student presenting their part must be visible in a corner of the presentation (live recording, not a static head shot), and when they start presenting, they must mention their name.

Overlaying a webcam recording can be easily done using either the video recording functionality of PowerPoint itself (see for example this YouTube tutorial) or using other recording software such as OBS Studio, DaVinci Resolve, Adobe Premiere, and many others. It is up to you (depending on your preference and experience) which software to use, as long as the final video satisfies the requirements mentioned above and below.

The entire presentation must be in the form of a video (720p or 1080p MP4 format). Note that video files can easily become quite large (depending on the level of compression used). To avoid storage problems for this course, the video upload limit is 100 MB per group, which should be more than enough for this type of presentation. If your video file is larger, use tools like HandBrake to re-encode with higher compression.

The video presentations will be marked offline (there will be no live presentations). If the markers have any questions, they may contact the group members by email for clarification.

Report

Each group must also submit a written report (in 2-column IEEE format, maximum 10 pages of main text, with any additional pages of references).

The report must be submitted as a PDF file and include:

1. Introduction: Present a brief introduction including the relevance and your understanding of the task and the dataset.
2. Literature Review: Review relevant techniques in literature, along with any necessary background to understand the methods you selected.
3. Methods: Motivate and explain the selection of the methods you implemented, using relevant references and theories where necessary.
4. Experimental Results: Explain the experimental setup you used to test the performance of the developed methods and the results you obtained.
5. Discussion: Provide a discussion of the experimental results and findings, comparing the different methods you developed.
6. Conclusion: Summarise what worked and what did not work, any limitations of the developed methods, and recommend future work.
7. References: List the literature references and other resources used in your work. The references section does not count toward the 10-page limit.

The upload limit for the report is 10 MB per group. If your initial report is larger than that, you may need to resize images or increase the compression.

Code

The source code of the developed software must be submitted as a ZIP file. The file should mainly contain your own code. Libraries or code obtained from other sources should be clearly described in a README file. Include proper documentation about how to run the code, use inline comments, and make the code well structured.

The upload limit for the source code is 25 MB per group. To stay within this limit, do not include trained models, input images, or result images.

Student Contributions

As a group, you are free in how you divide the work among the group members, but all group members must contribute roughly equally to the method development, coding, making the video, and writing the report. For example, it is unacceptable if some group members only prepare the video and report without contributing to the methods and code.

An online survey will be held at the end of the term allowing students to anonymously evaluate the relative contributions of their group members to the project. The results will be reported only to the LIC and the Course Administrators, who at their discretion may moderate the final project mark for individual students if there is sufficient evidence that they contributed substantially less than the other group members.

References

Dataset: Mete Ahishali. Aerial Imagery for Dead Tree Segmentation. Kaggle, April 2025.

Paper: Mete Ahishali, Anis Ur Rahman, Einari Heinaro, Samuli Junttila. ADA-Net: Attention-Guided Domain Adaptation Network with Contrastive Learning for Standing Dead Tree Segmentation Using Aerial Imagery. arXiv:2504.04271, April 2025.

Released: 6 July 2025