Introduction to Computer Vision 2025
Innopolis University

Computer Vision Project Guidelines

Issue Date: 09 September 2025

Your class project is an opportunity for you to explore an interesting computer vision problem. Each one of you will work in a group of **three or four students on a single project**. The instructor and teaching assistants will consult with you on your ideas, but of course the final responsibility to define and execute an interesting piece of work is your responsibility. Your project will be worth 35% of your final class grade. Each group chooses the project from the provided list of topics or proposes their own idea. It is possible, more than one group can select the same dataset but need to apply different techniques. Please write the details in the following link (till the end of this week). If you will use existing code from recent work in the literature then you need to mention clearly what is add-on to the existing development.

Project Topics

Expectation: Each project should be developed in python and rigorous **analysis** is required.

Deliverables

1. **Preliminary Proposal:** 1 page **[Points: 3%]**
    a. Title
    b. Project idea: (Should be understandable to anyone in the course)
    c. Technique/Method
    d. Dataset explanation and link
    e. Timeline with each member's individual task.
    f. References - no less than 5

    Template: used which one suits you well (i.e., MS Word, Latex, Overleaf) from the following link: https://www.ieee.org/conferences/publishing/templates.html

    Submission: Upload *.pdf* proposal in moodle

    Deadline: 30 September 2025
2. **Final Submission**
    a. **Project Report [Points: 15%]**

    (Template: https://www.ieee.org/conferences/publishing/templates.html)

    i. Introduction
    ii. Related Work
    iii. Methodology
    iv. GitHub link: (Double check the link it should work properly)
    v. Experiments and Evaluation
    vi. Analysis and Observations
    vii. Conclusion

viii. Reference - no less than 15

b. Project Demo Presentation [Points: 15%] 10 minutes

i. Presentation of the project idea, motivation, contribution, methods, and results - 4-5 minutes.
ii. Run a live demo - 1-2 minutes.
iii. Questions and Discussions (each student should answer some questions about their project) - 3-4 minutes.

c. GitHub Maintenance [Points: 2%]

i. .md file explaining the project details such as title and one paragraph of about the project. (Strictly follow this)
ii. Good coding practice should be followed

Submission: Upload *.pdf* report (Minimum page limit: 4 Pages) including the link to repo, and presentation in .pdf or pptx on Moodle.

Deadline: 18 November 2025

NOTE: If your code is not working during the project demo then you whole submission will be evaluated in 50% marks of the "Final Submission".

- **Policy**
  - **No late submission is allowed.**
  - Your code should be properly commented
  - If plagiarized code == True then "0"

List of Topics

Project: Text2Scene: Generating Compositional Scenes from Textual Descriptions
http://openaccess.thecvf.com/content_CVPR_2019/papers/Tan_Text2Scene_Generating_Compositional_Scenes_From_Textual_Descriptions_CVPR_2019_paper.pdf

Project: 3D human pose and shape estimation from a single image
http://openaccess.thecvf.com/content_CVPR_2019/papers/Kolotouros_Convolutional_Mesh_Regression_for_Single-Image_Human_Shape_Reconstruction_CVPR_2019_paper.pdf

Introduction to Computer Vision 2025
Innopolis University

Project: Long-Term Feature Banks for Detailed Video Understanding
http://openaccess.thecvf.com/content_CVPR_2019/papers/Wu_Long-Term_Feature_Banks_for_Detailed_Video_Understanding_CVPR_2019_paper.pdf

Project: Object detection in aerial images
http://openaccess.thecvf.com/content_CVPR_2019/papers/Ding_Learning_RoI_Transformer_for_Oriented_Object_Detection_in_Aerial_Images_CVPR_2019_paper.pdf

Project: Human Activity Detection and Recognition for Video Surveillance

Project: Visual words for image retrieval/classification task

Project: Video classification using CNNs

Project: Multiclass Food Classification
https://www.kaggle.com/theimgclist/multiclass-food-classification-using-tensorflow/data

Project: YOLO based application for object detection

Project: Content-Based Image Retrieval
Description: Create a system that can retrieve images from a large database based on their visual content, rather than just keywords.
What Is Content-Based Image Retrieval? | Baeldung on Computer Science

Project: Document Image Dewarping
Description: This project focuses on correcting geometric distortions in images of documents, such as those captured by a camera.
Axis-Aligned Document Dewarping

Project: Object Tracking in Videos with Occlusions
Description: Build a system that can track objects in a video, even when they are temporarily hidden from view. This is a challenging problem with applications in surveillance and robotics.
Robust Occlusion Handling in Object Tracking

Introduction to Computer Vision 2025
Innopolis University

Project: Multi-Object Tracking

Description: Develop a system that can track multiple objects in a video, which has applications in surveillance, sports analytics, and autonomous driving.

(PDF) Deep Learning for Multiple Object Tracking: A Survey

Project: Object Detection in Aerial Images

Learning RoI Transformer for Oriented Object Detection in Aerial Images

********** Good Luck **********