

COMP5523 Computer Vision and Image Processing Project

VisionMate: Visual Model-Assisted Blind System

Motivation.

The group project is designed to consolidate students' understanding and practical expertise in computer vision. Students will gain extensive experience analyzing digital image data through state-of-the-art vision models and advanced pattern recognition approaches. The collaborative nature of the project cultivates leadership abilities, effective teamwork, and professional communication skills through the delivery of high-quality technical outputs. By bridging theoretical foundations with real-world applications, this project equips students with both the technical capabilities and professional competencies essential for success in computer vision and AI development. Thus, it aligns with the Intended Learning Outcomes (specifically, b, c, d, e) of COMP5523.

Task Specification.

In this project, let us develop a vision-based assistive system to help visually impaired individuals navigate and interact with their environment using advanced computer vision and deep learning models. This system will assist users by providing real-time audio descriptions of their surroundings, detecting obstacles, reading text, and identifying objects through a camera interface.

The system could consist of various tasks, such as real-time object detection, scene understanding, text recognition (OCR), depth estimation, and so on. For example, a user could ask "What is in front of me?" or "Read this text," and the system would analyze the camera feed to provide accurate audio descriptions. The focus will be on developing practical and reliable assistance for daily activities of visually impaired individuals.

To successfully complete the project, each group will go through the following key steps, using the knowledge and techniques learned in class and beyond:

- **Data preparation:** Collect and organize relevant datasets for testing (or training) visual models, including images of common objects, scenes, text samples, and navigation scenarios. Preprocess the data to ensure consistency and quality.
- **Algorithm design:** Explore state-of-the-art computer vision approaches including object detection, semantic segmentation, scene understanding, and so on. Implement efficient real-time processing pipelines and design effective audio feedback generation strategies. Integrate multiple vision models to provide comprehensive environmental understanding.
- **System implementation:** Develop the assistive system using frameworks such as PyTorch, OpenCV, and modern vision transformers. Create a user-friendly interface that processes camera input and provides clear and helpful audio feedback.
- **Performance evaluation:** Assess the system's accuracy, reliability, and usefulness for visually impaired users. Measure response times, detection accuracy, the clarity of audio descriptions, and so on.

Important Dates.

There are two important demonstrations that students need to prepare:

1. **In-Class Presentation with a Live Demonstration (18:30 to 21:20 November 25, 2025):** Each group will give a 15-minute presentation during the November 18, 2025 class. The presentation should showcase the task settings, challenges, methodologies, and their developed visual assistance system's functionality, features, and effectiveness. During the presentation, students will conduct a live demonstration, showcasing the system's capability to assist in real-world scenarios.
2. **Project Report Submission (by 23:59 on December 02, 2025, Tuesday):** Alongside the class presentation, each group should submit a comprehensive project report. The report should document the task setting, background, system development process, challenges, methodologies, outcomes, and evaluation results. Additionally, the group members' roles and contributions made by each group member (a maximum of 6 group members) should be delineated. The report should be **capped at 8 A4 portrait** pages with unlimited spaces for references in APA formatting. There is no need to include a Table of Contents and the cover page. The text should be *formatted in 12-size Times New Roman, 1.0 line space, and a single column with a 1-inch margin*, while there are no requirements for figures and tables (students are encouraged to put in figures and tables to make the writing more eye-catching and easy to understand).

Assessment Rubrics.

The project will take 15% of the final grade with the assessment rubrics as follows. The rubric assesses the project based on the given criteria, with a scale of 3% to 1%, where 3% represents the highest level of achievement and 1% represents the lowest. The descriptions provided in the rubric can be tailored to the project's specific requirements.

1. **Appropriateness (3%):**
 - Task settings, challenges, methodologies, and system functionality are highly appropriate and relevant. (3%)
 - Task settings, challenges, methodologies, and system functionality are mostly appropriate and relevant. (2.5%)
 - Task settings, challenges, methodologies, and system functionality demonstrate some level of appropriateness. (2%)
 - Task settings, challenges, methodologies, and system functionality are inadequate. (1.5%)
 - Task settings, challenges, methodologies, and system functionality are inappropriate. (1%)
2. **Soundness (3%):**
 - The project demonstrates a comprehensive and well-organized development process with clear and logical explanations. (3%)
 - The project demonstrates a mostly comprehensive and well-organized development process with mostly clear and logical explanations. (2.5%)
 - The project demonstrates some level of organization and logic in the development process, but with some unclear explanations. (2%)

- The project demonstrates limited organization and logic in the development process and provides inadequate explanations. (1.5%)
- The project is disorganized and lacks a clear or logical development process with unclear or illogical explanations. (1%)

3. Excitement (3%):

- The project presents innovative and engaging ideas that consistently capture the attention of the audience. (3%)
- The project presents engaging ideas that mostly capture the attention of the audience, but may lack consistency. (2.5%)
- The project presents some level of engagement, but with areas of improvement and inconsistent audience attention. (2%)
- The project lacks innovation and fails to capture the audience's excitement consistently. (1.5%)
- The project lacks excitement or fails to engage the audience. (1%)

4. Presentation (3%):

- The presentation is highly polished and professional, with excellent delivery and effective use of visual aids. (3%)
- The presentation is mostly polished and professional, with good delivery and adequate use of visual aids. (2.5%)
- The presentation has some areas of improvement in polish and professionalism, and delivery may lack consistency. (2%)
- The presentation lacks polish and professionalism, and delivery is weak. Visual aids are ineffective. (1.5%)
- The presentation is poorly executed, making it difficult to follow, with ineffective use of visual aids. (1%)

5. Writing (3%):

- The project report is well-written, with clear and concise explanations of ideas and proper use of grammar and formatting. (3%)
- The project report is mostly well-written, with mostly clear and concise explanations of ideas and mostly proper use of grammar and formatting. (2.5%)
- The project report is adequately written, with some clarity issues and inconsistent use of grammar and formatting. (2%)
- The project report is poorly written, with unclear or confusing explanations and significant errors in grammar and formatting. (1.5%)
- The project report is very poorly written, with numerous clarity and formatting errors, making it difficult to understand. (1%)