Project Title

Semantic Segmentation for Memory-Constrained Edge Devices

Project abstract

The aim of this project is to design a lightweight semantic segmentation network architecture that: (1) is designed to use minimal resources, (2) while maintaining high accuracy, and (3) fits the Midokura production environment. The candidate will work under the mentorship of AI engineers.

Extended abstract

This project aims to develop a lightweight semantic segmentation model that can be deployed on memory-constrained edge devices. Semantic segmentation is a computer vision task that involves identifying and labeling each pixel in an image with a corresponding class label. Edge devices such as smartphones, drones, and surveillance cameras have limited computing resources and memory, making it challenging to run complex models for tasks like semantic segmentation.

The focus of this project is to create and optimize a semantic segmentation model for deployment on Sony edge devices. Midokura, a company specializing in edge computing, will provide the necessary guidance and expertise to create an efficient model that can run on Sony devices. The internship will offer an opportunity to work closely with the Midokura team to develop and deploy the model.

Midokura will provide all necessary resources, and the project will be supervised by a team of experienced AI engineers. The student will have access to an A100-GPU server to perform experiments needed for proper investigation. The project will be conducted using an agile methodology based on sprints, with regular updates to the team on project progress.

The project will consist of several stages. First, the student will conduct research by reviewing relevant literature and learning about state-of-the-art models in the field. Second, the student will create or adapt SOTA models to meet our hardware limitations. Third, the student will conduct various experiments to identify an optimal model for our specific case. Finally, the student will integrate the model into our production environment.

The goal is to create a model that can accurately perform semantic segmentation while remaining lightweight and efficient. The resulting model will be used for semantic segmentation in the Midokura production environment. Overall, the project will contribute to the advancement of edge computing and bring us one step closer to enabling intelligent devices in various industries.