homework4

May 2025

Hi, everyone. This is the last homework of our computer vision class. This assignment includes only one part of programming. Here are the **requirements**:

- You are required to use Python for all the programming tasks.
- We recommend you typeset your report using LATEX and submit the PDF. You are asked to submit a single zip file on WebLearning. The zip file should be consisted of the report and your code.
- Referring to the public code on the internet is allowed, but copying is absolutely forbidden.
- Please finish the programming and report independently.
- We provide the code framework for you (see in src.zip attached). Zip all your files and name it as: name_id_hw_4.zip,
  eg: zhangsan_2024123456_hw_4.zip, and the structure of directory should be like this:

```
zhangsan_2024123456_hw_4
├── report.pdf:  your report file
└── src:  put your code
    ├── README.md:  more details for code implementations
    └── etc...
```

1 Challenge of Cityscapes: A Semantic Segmentation Task

Since you have learned semantic segmentation task in computer vision, you are required implement your own semantic segmentation models for this assignment.

1.1 Dataset

Cityscapes is a new large-scale dataset that contains a diverse set of stereo video sequences recorded in street scenes from 50 different cities, with high quality pixel-level annotations of 5,000 frames in addition to a larger set of 20,000

weakly annotated frames. For the preprocessing of the dataset, you can refer to this codebase.

1.2 Requirements

Here are the requirements:

- Settings: **Cross-validation** is required for this project.
- Measurement: Mean Intersection over Union(**MIoU**) score is the only measurement of the performance.
- Models: Feel free to choose semantic segmentation models (vanilla or SOTA is ok). But we **DO NOT** recommend you to choose some large networks, such as mask r-cnn.
- Experiments: If you use some techniques to improve the performance, please do some **ablation study** for this changes compared with your baseline model.
- Visualizations: **Visualization** of your results is required, and please put down your experimental settings (hyperparameters) and training/testing curves.
- Others: Design an **original** architecture is highly encouraged. (**Bonus**: if it can work well)

1.3 Scores

Your score of this assignment is divided into three parts:

- Performance (60%)
- Code (20%)
- Document (20%)

For the performance part, we will rank all the students and the score of this part will come from your ranks following this rules:

- We will choose a vanilla baseline model as the **benchmark**.
- Suppose the performance of the benchmark is y and your rank is r. You will get 60 scores if your performance $\geq y$. Otherwise, you will lose 6 scores for 10% performance drop (e.g., $[0.9y, y) \rightarrow 54$ scores). The top 5 students will receive extra scores: Bonus / r.

For the code part, any **plagiarism** will make you fail. And we will checkout the codebase and **test** the code to verify the results in your report **randomly**. You will also fail if the **gap** between your results and the our testing results is too large (over 20%).

Your documents should include these parts:

- Introduction
- Architecture/Algorithm
- Results
- Discussion(Analysis)
- Conclusion