CS 4391 Introduction to Computer Vision
Homework 2

Instructor: Yu Xiang

February 17, 2026

Download the homework2_programming.zip file from eLearning, Assignments, Homework 2. Finish the following programming problems and submit your scripts to eLearning. You can zip all the data and files for submission. TA will run your scripts to verify them.

It is recommended to use Anaconda `https://www.anaconda.com/` to build a python environment. For example, to create an environment named “cs4391”:

- conda create -n cs4391 python=3.10
- conda activate cs4391

Install the Python packages needed by

- `pip install -r requirement.txt`

Install PyTorch (CUDA 13.0 example):

- `pip install torch torchvision --index-url  https://download.pytorch.org/whl/cu130`

Make sure the CUDA version matches your system. Check available versions at: `https://pytorch.org/get-started/locally/`

If you don’t have a NVIDIA GPU with CUDA, you can install a cpu version.

Here are some useful resources:

- Python basics `https://pythonbasics.org/`
- Numpy `https://numpy.org/doc/stable/user/basics.html`
- OpenCV `https://docs.opencv.org/4.x/d6/d00/tutorial_py_root.html`
- PyTorch `https://docs.pytorch.org/tutorials/beginner/basics/intro.html`

Problem 1

(4 points) Preparing the dataset.

1.1 Download the dataset

We will use the objects in the FewSOL dataset `https://irvlutd.github.io/FewSOL/` for classification. Check the following link:

- `https://utdallas.app.box.com/folder/166466844982?v=FewSOL-Dataset`

Download the real_objects.zip file to the folder “homework2_spring26”. Unzip the file.

1.2 Data understanding

Finish the TODOs in the hw_part1_dataset_understanding.py script, and then run the following command to test it:

- `python hw_part1_dataset_understanding.py --root real_objects`

You will see some print and a figure similar to the following one.

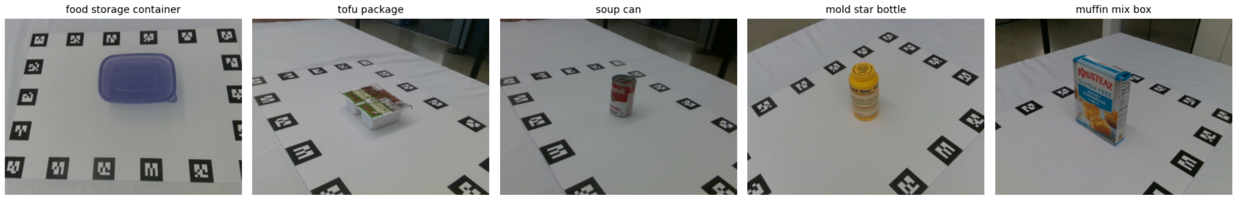

Figure 1: Sampled classes.

Submission: Upload your hw_part1_dataset_understanding.py file and an image of the dataset visualization to eLearning.

Do not submit the dataset to eLearning. That is too large.

Problem 2

(6 points) Network training and testing for image classification.

2.1 Generate a split of the dataset

First, we will use the following script to generate a split of the dataset. Run the python script:

- `python make_split.py --root real_objects --shots 5 --val 2 --test 2 --seed 0 --out split.json`

The dataset will be split into a training set, a validation set, and a test set. The split information will be save into a "split.json" file.

2.2 Training and testing

Finish the TODOs in the hw_part2_train_scratch_vs_pretrained.py script, where you need to implement a simple CNN for image classification.

After your implementation, do the following three runs and check the results.

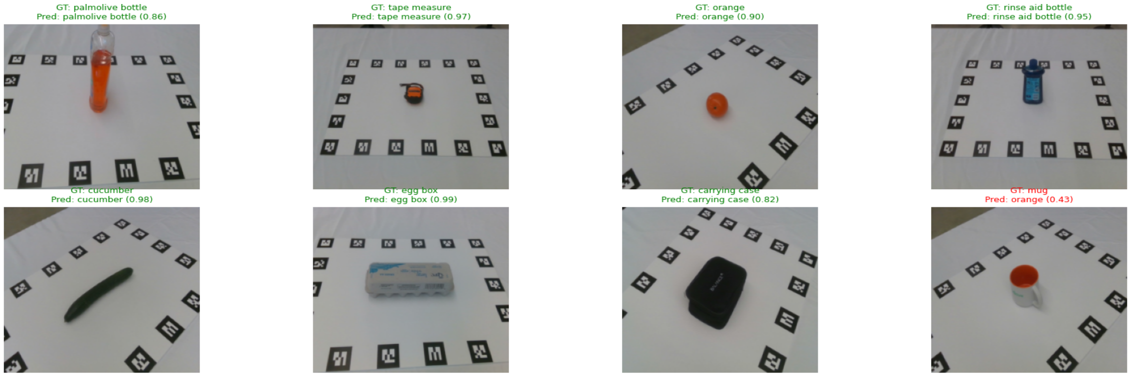

Figure 2: Classification results.

- Train a simple CNN from scratch:

```
python hw_part2_train_scratch_vs_pretrained.py --split split.json --model scratch
--epochs 10 --tag scratch
```

- Use a pretrained ResNet18 and then fine-tune the whole network:

```
python hw_part2_train_scratch_vs_pretrained.py --split split.json --model pretrained
--pretrained --epochs 10 --save_best --tag pretrained_ft
```

- Use a pretrained ResNet18, freeze the backbone and then fine-tune the FC layer only:

```
python hw_part2_train_scratch_vs_pretrained.py --split split.json --model pretrained
--pretrained --freeze_backbone --epochs 10 --save_best --tag pretrained_freeze
```

After training, the model checkpoints and training curves will be save to a folder “runs/part2”. You can check these outputs.

2.3 Test the trained model

Run the following script to test your trained model based on ResNet18 and see the visualization:

- `python visualize_test_predictions.py --split split.json`

  `--ckpt runs/part2/pretrained_ft/best.pt --num_samples 8`

You can try different checkpoints. You will see a figure similar to Fig. 2.

Submission: Upload the following items to eLearning.

- hw_part2_train_scratch_vs_pretrained.py
- the 3 training loss plots according to the 3 training commands
- a test result image. Use your best trained model.

Do not submit the dataset to eLearning. That is too large.