Programming Assignment - Edge Detection

TA: Ishan M (ep20b017)
Modern Computer Vision

March 15, 2025

Not all edges are equally important to the human eye. For segmentation, we need to detect images through edges that humans perceive as important. The Berkeley Computer Vision group has an annotated dataset Berkeley Segmentation Data Set and Benchmarks 500 (BSDS500).. The goal of this assignment is to build a deep learning network that works for edge detection on this dataset.

Task 1: Canny Edge Detection
First, download the dataset and write a canny edge detection algorithm to extract edges. Use different values of the blurring parameter to see how close you can get to the ground-truth edges. Your answer should contain plots of the image, ground truth and canny output for various values of blurring parameter. Does the canny detector work well for edge detection? Why/Why not?

Task 2: Simple CNN Model
Now, we use deep learning to try and improve our results. Create a dataloader for the BSDS dataset and train a simple 3 layer CNN (ReLU activation in hidden layers, kernel size 3, padding=1, 8 filters in hidden layer 1, 16 in hidden layer 2, num_epochs =100) using the training dataset. Use this trained network to plot outputs for the test dataset and compare with ground truth. Apply an appropriate threshold on the output to binarize the image.
Read the paper on holistically nested edge detection (linked in task 4) and use the class balanced loss as described in equation 2 of section 2.2.

Task 3: VGG16 Model
The second network we will try is the VGG16 convolutional network. Ensure the images are the right format for the input of VGG16. Import VGG 16 without the last max-pooling layer and the fully connected layers. Instead, use a tranpose convolution(transposeConv2D in pytorch) decoder at the end of the network to restore output size to original image size. (How well does only bilinear interpolation upsampling work? Try this out and compare). Train this network like you did in Task 2 (add threshold as required to binarize the output edge map). Use this trained network to plot outputs for the test dataset and compare with ground truth. Use class balanced loss function as above.

For tasks 2 and 3 your report should contain the following:

1. Describe the loss function you used at the output layer for such a task and why is it better than a simple binary cross entropy loss? What activation function did you use for the output layer and why?

2. Plot train and validation loss vs epoch for your network.

3. Plot test image, ground truth and model output for a couple of test images.

4. Write your observations on the performance of the model. (Bonus: If you have time try changing the hyperparameters to see if model can be improved)

5. Compare the outputs of VGG16 and CNN. Which network performed better?

Task 4 Holistically Nested Edge Detection

We now implement a model similar to the state-of-the-art HED model for edge-detection. For this task, download the pre-trained VGG without the fully connected layer and the final max-pooling layer.

Before each pooling layer of the VGG model, extract a side-output. Upsample (using simple bilinear upsampling to original size) the side outputs and fuse them with appropriate weights. These weights are also parameters learned during training.

For the loss function refer to Section 2.2 of the HED paper and apply a loss function as described. Describe what you used in your report. Train the network for 50 epochs.

Visualize your output for a train image after appropriate thresholding to binarize the output.

For task 4 your report should contain the following:

1. Plot train and validation loss vs epoch for your network.

2. Plot test image, ground truth and model output for a couple of test images.

3. Plot side outputs and print learned weights. What is your observation?

4. Write your observations on the performance of the model.

5. Compare the outputs of this model with VGG16 and CNN. Which network performed the best? Recall that our goal is not to detect all edges, but the ones that are ”important” as described in the ground-truth image.

6. How does simple canny compare to HED?