Assignment 8

1 Training and Evaluating U-Net for Image Segmentation

1.1 Dataset and Preprocessing

- Used the Oxford-IIIT Pet dataset with images and segmentation masks.
- Images and masks resized to 128×128 pixels.
- Masks converted to binary masks (pet = 1, background = 0).
- Images normalized to the range $[0, 1]$.

1.2 Model and Training

- Implemented a U-Net architecture with encoder-decoder blocks and skip connections.
- Compiled the model using Adam optimizer and binary cross-entropy loss.
- Trained for 10 epochs with a batch size of 32.

1.3 Results

- **Training Accuracy:** 89.81%
- **Training Loss:** 0.2486
- **Validation Accuracy:** 90.01%
- **Validation Loss:** 0.2415
- **Test Accuracy:** 90.05%
- **Test Loss:** 0.2408

The model converged well with high accuracy on both training and validation datasets, demonstrating effective learning of the segmentation task.

1.4 Visualization

Sample visualizations were generated showing the input images, ground truth masks, and predicted masks side-by-side. These confirm that the U-Net model effectively segments the pet regions from the background.

Figure 1: Sample visualization showing (left) input image, (middle) ground truth mask, and (right) predicted mask.

2 Training and Evaluation of U-Net for Crowd Counting

Training Details

The U-Net model was trained for 10 epochs. The training and validation loss improved steadily:

- Initial training loss: 0.0011
- Final training loss: 0.0004473
- Final validation loss: 0.0007950

Evaluation Metrics

The trained U-Net model was evaluated on the test set, yielding the following metrics:

- Mean Absolute Error (MAE): 408.21
- Root Mean Squared Error (RMSE): 687.21

Visualization

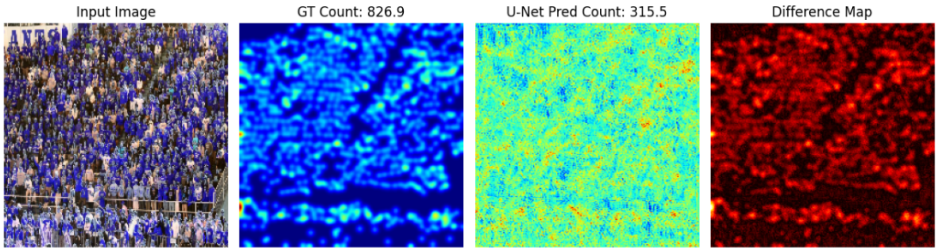

Figure 2: Sample visualisation showing (left) input image, (middle) ground truth count, U-net prediction count and (right) difference map.

Task 3: Training and Evaluation of MCNN for Crowd Counting

Training Details

The MCNN model was trained for 10 epochs. The training and validation loss were stable:

- Initial training loss: 0.0037
- Final training loss: 0.0007649
- Final validation loss: 0.0008582

Evaluation Metrics

The MCNN model evaluation on the test set produced:

- Mean Absolute Error (MAE): 395.87
- Root Mean Squared Error (RMSE): 592.29

Visualization

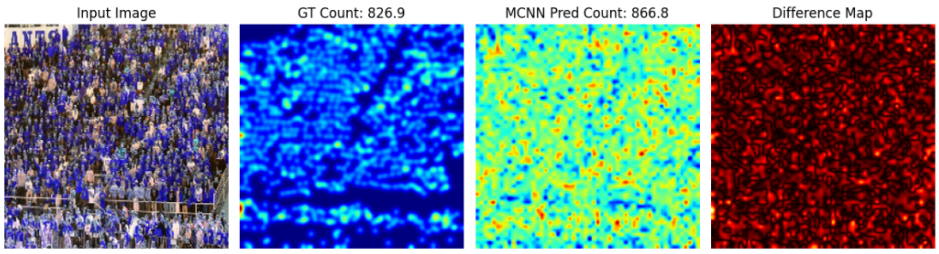

Figure 3: Sample visualisation showing (left) input image, (middle) ground truth count, MCNN prediction count and (right) difference map.

Comparing MCNN-based and U-Net based crowd counters.

Model	MAE	RMSE
U-Net	408.21	687.21
MCNN	**395.87**	**592.29**

Table 1: Comparison of MAE and RMSE metrics for U-Net and MCNN models.

MCNN outperforms U-Net on both MAE and RMSE, indicating higher accuracy and robustness. Visualizations support these findings, with MCNN producing smoother and more consistent density maps.

You can access the file at Google Colab.