

IPR_PROJECT_2_DETAILS

November 6, 2024

1 Implementation Details

1.1 Model Idea

- Counting Network: A counting network is implemented, focusing on improvement over traditional counters.
 - A Front-End Net as a backbone
 - A Multi-scale convolutions Module : Coprime dilation rates (e.g., [1,2,3]) are used to capture details at various scales, ensuring all pixels are considered without gaps.
- Architectures :
 - Front-End Net: First 10 layers of VGG-16 are used. Extracts primary features from images.
 - Dual Convolution Columns: Two parallel convolution layers (3x3 and 5x5 kernels) with coprime dilation rates to capture multi-scale crowd features without redundancy.

1.2 Implementation Ideas

- **THE MODEL IS IMPLEMENTED FROM SCRATCH, NO REFERENCE IMPLEMENTATION AVAILABLE**
- ADAM optimizer is used for the model, with learning rate 1e-5. A weight decay rate of 1e-4 is also applied . (These values are just from the tested results on my own, more hyper parameter tuning might produce better results.)
- Batch size 1 is used for training and the model is trained for 30 epochs (GPU availability, RAM and Time constraints). Best is to run for 300 epochs.

2 Dataset

- **Dataset** : ShanghaiTech With People Density Map
- It contains 2 datasets : ShanghaiTechA and ShanghaiTechB
- In each dataset , there are 3 folder:
 - images: the jpg image file
 - ground-truth: matlab file contain annotated head (coordinate x, y)
 - ground-truth-h5: people density map

3 Results:

(These are best possible results i was able to get, A better implementation might give better results)

3.1 ShanghaiTechA

- MAE : 186.98
- RMSE : 253.12

3.2 ShanghaiTechB

- MAE : 78.21
- RMSE : 95.89

4 Changes Made from Previous Model

- Replaced the combination of CSRNet and an Adaptive Generator with a convolution based counter
- The coprime dilation layered network captures the detailed features of the image for better counter results

5 References:

- Crowd counting method via a dynamic-refined density map network : Yanbo Liu, Guo Cao, Zixian Ge, Yingxiang Hu
- Adaptive Density Map Generator
- Single-Image Crowd Counting via Multi-Column Convolutional Neural Network: Yingying Zhang; Desen Zhou; Siqin Chen; Shenghua Gao; Yi Ma
- Pytorch official