CROWD COUNTING
Multi-Scale Convolutions Using Co-prime Dilation Rates

Goutam Das, Department of Computer Science & Engineering, IIT Kanpur
Mentor : Bebina Hidangmayum Devi

01 Introduction

- Counting people in dense crowds is important for safety and planning, but traditional methods struggle with dense crowd and scale changes.
- This model uses a Counting Network which uses convolutions to estimate crowd density from images, aiming for accurate predictions across different crowd densities.

02 Objective

- Develop a model that accurately predicts crowd count using multi-scale features.
- Focus on improving counting in complex, dense scenes with diverse crowd sizes.

03 Approach

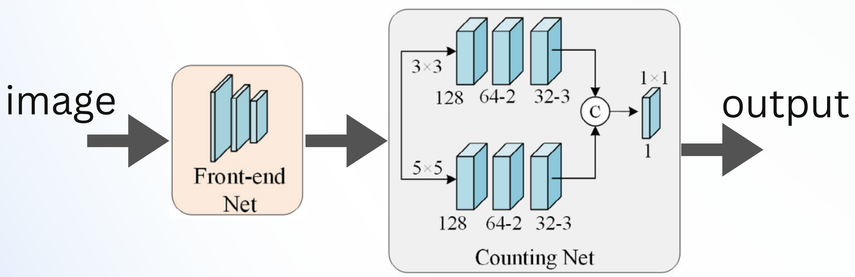
Counting Network: A counting network is implemented, focusing on improvement over traditional counters.

- 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.

04 Model Description

The Counting Network combines:

- 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.



05 Model Training and Evaluation

- Dataset: Uses the ShanghaiTech dataset (Parts A and B) to train and test on different crowd densities.
- Loss Function: Mean Squared Error (MSE) measures differences between predicted and actual GT-dot maps.
- This measures pixel wise difference between the counter output and the ground truth map.
- PyTorch is used for implementation of this model.

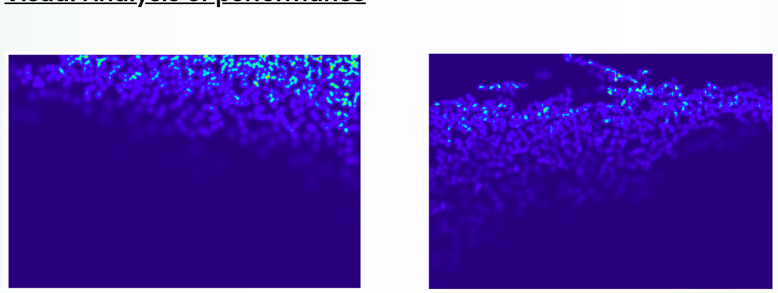
06 Results & Analysis

Comparative performance with Adaptive Density Map Generator

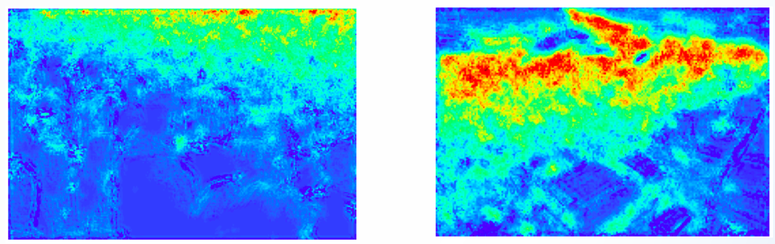
Dataset	CSRNet with Generator	Current Model
SanghaiTech A	MAE : 238	MAE : 187
	MSE : 397	MSE : 253
SanghaiTechB	MAE: 85	MAE : 78
	MSE: 121	MSE : 96

Note : The current model trained only for 30 epochs

Visual Analysis of performance



Ground Truth Density Map



Predicted Map for Counter (*the scale is different*)

Refrences:

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