CROWD_SZ: A Large-scale Multi-view Crowd Counting Semantic Dataset

Jiajia Wu[1], Yufeng Lin[2], Cheng Wu[1], Jin Zhang[1], Lijun Zhang[1]

[1] School of Rail Transportation, Soochow University

[2] Command Center of Suzhou, Industrial Park Public Security Bureau

Contact email: chengwu@suda.edu.cn

Soochow University

IARIA

Author

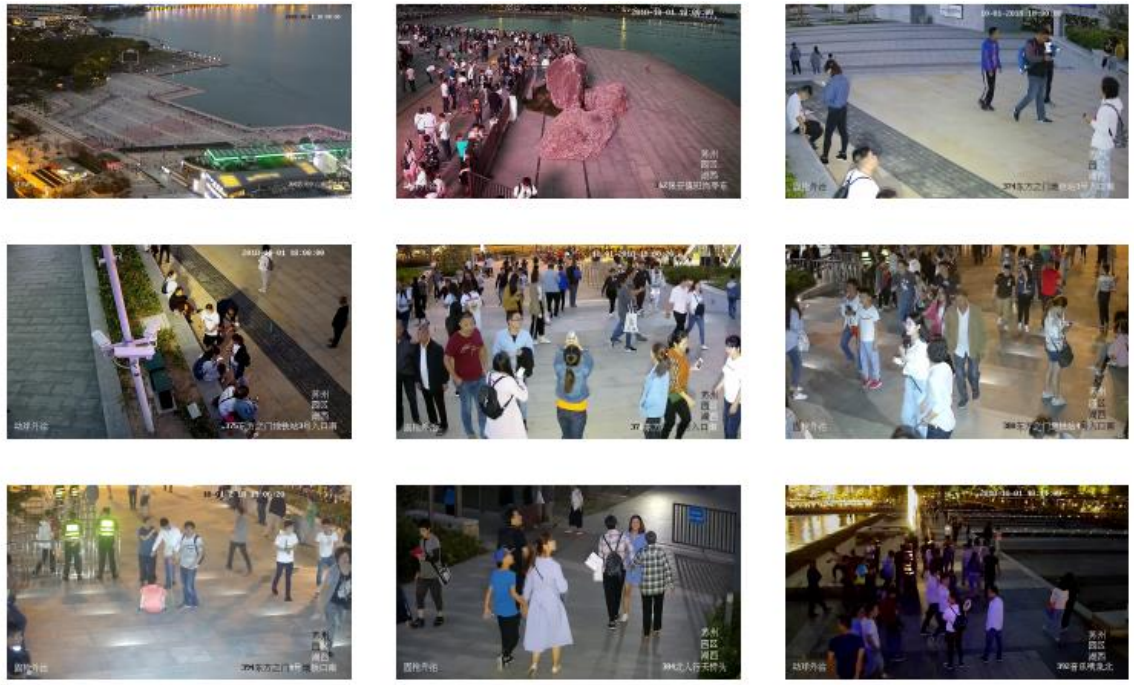
JIAJIA WU is studying in Soochow University, majoring in traffic information engineering and control.
Her research interests lie in the field of machine vision in artificial intelligence, mainly crowd counting algorithms, density estimation and multi-source information fusion.

Contributions of our paper

1. We propose a new large-scale multi-view dataset for crowd counting and density estimation. The dataset includes 5,610 images and 1,738 videos, which makes up for the lack of diversity in traditional datasets;
2. For the stock data of the dataset, a large number of complete head annotation files are carefully prepared;
3. For the key elements of dense crowd vision, different scene video frames with good statistical dispersion are provided;
4. Fully considering the basic principle of "the high-altitude perspective must include the low-altitude perspective coverage area", our dataset provides global and local images, and allows to calculate the number of people in wide-area places according to the time correlation and spatial complementarity between them.

Dataset

This dataset is collected from The Fountain Square of Jinji Lake in Suzhou, which covers an area of nearly 30000 square meters. Fountain screening activities are held regularly, during which a large number of tourists will be attracted to watch. There are many subway stations around and the terrain of entrance and exit is complicated, so the flow of people changes quickly and the crowd count is difficult.



Example from CROWD_SZ

Image collection location and pixel size

Compared with other datasets, the original sizes of the images in our dataset are 1920*1080, 2560*1440, etc., while the image sizes of other existing datasets mostly do not exceed 1k*1k.

The installation location and pixel size

Density distribution

Low density: between zero and fifty.

Medium density: between fifty and five hundred.

High density: between five hundred and one thousand.

Ultra high density: count value above one thousand.

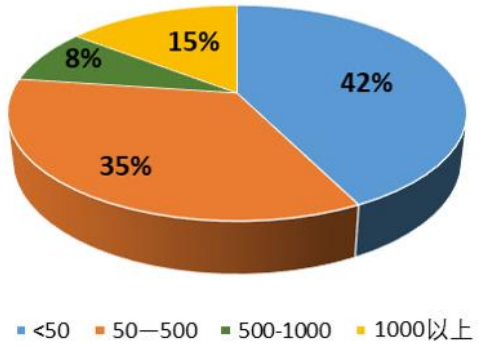
Low density

Medium density

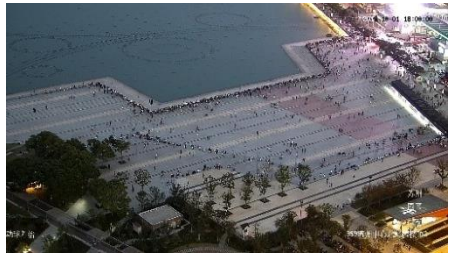
High density

Ultra high density

Density	Low	Medium	High	Ultra high
No. of images	2389	1941	460	820



Diversity



High altitude

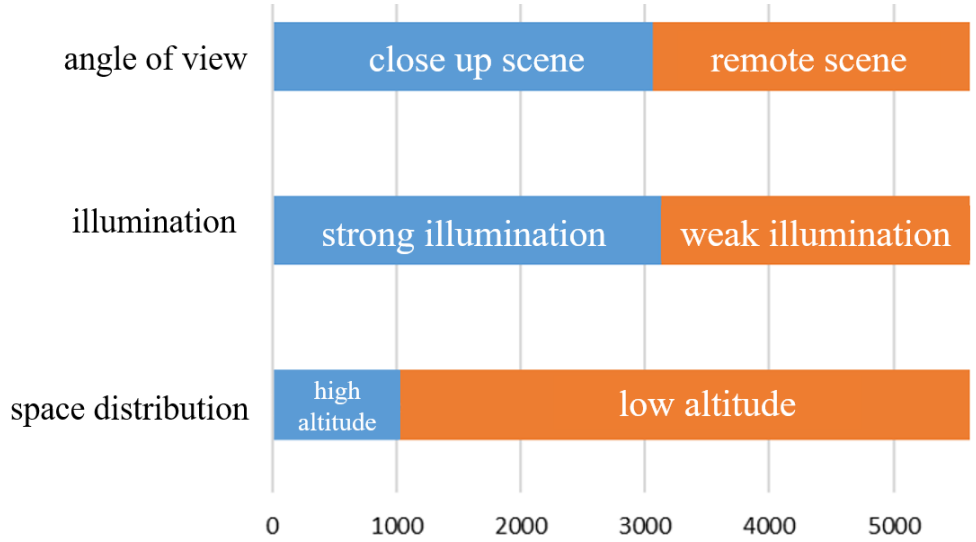
Low altitude

Strong illumination

Weak illumination

Close up scene

Remote scene

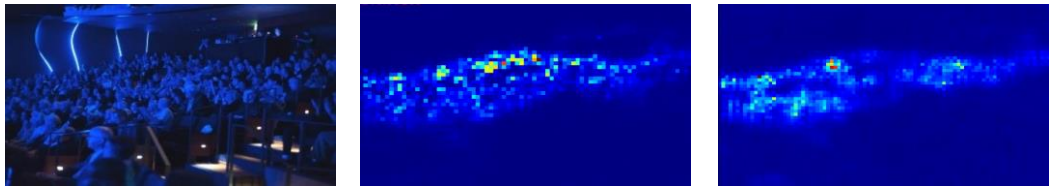


Scene diversity is an important attribute of the dataset. Our dataset contains images taken by cameras at different heights and angles.

IARIA

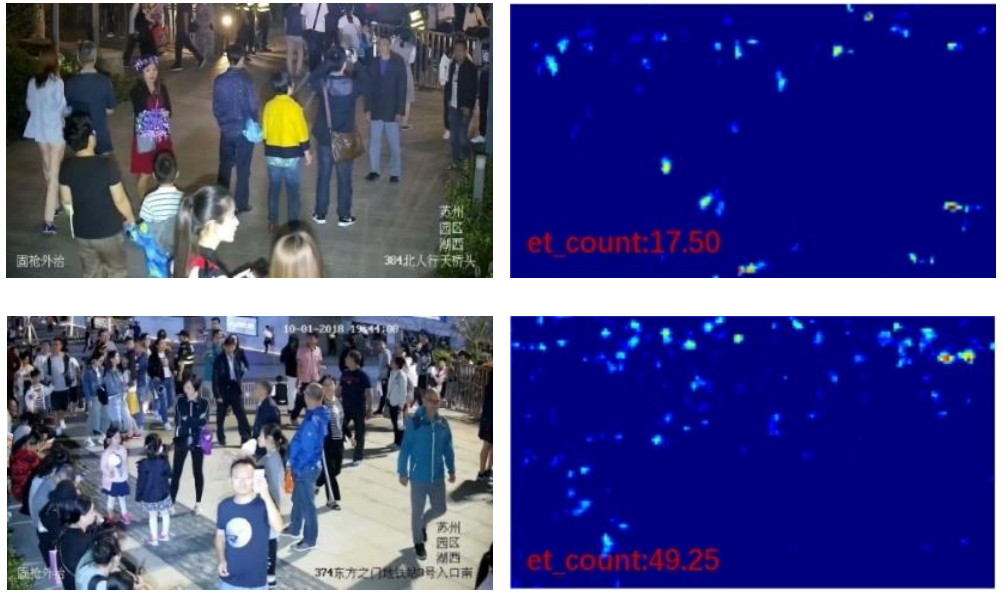
Evalution

We evaluate some typical crowd counting algorithms on the CROWD_SZ, select a low-altitude image as the input data to get the specific number of people and the output density map. To recognize the pedestrian characteristics better, we preprocessed the images, the image needs to be grayed first to obtain the density map.



a b c

Image preprocessing: a is the original image, b is the density map obtained from the input original image, and c is the density map obtained by graying the original image



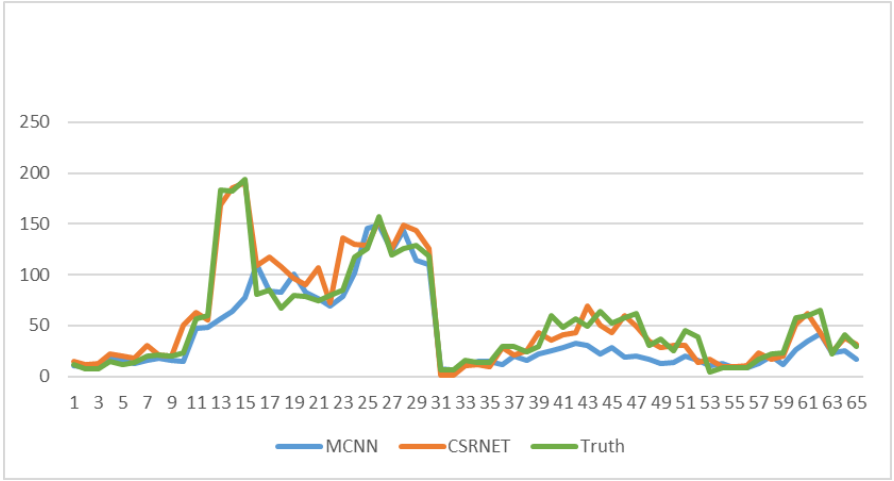
Input Output

Evalution

Choose MCNN and CSRNet as the benchmark algorithms, we list the estimated number of the same image obtained by two algorithms, both of which estimate the image with higher accuracy. In crowded scenarios, CSRNet has a higher accuracy. On the contrary, MCNN has a higher accuracy in sparse scenes.

Comparison of MAE values between MCNN and CSRNet in different scenarios

Scene \ MAE	MCNN	CSRNet
Sparse scene	0.02839	0.1429
Crowded scene	0.2979	0.1415



The comparison of prediction results of MCNN and CSRNet

IARIA

Compare with other datasets

Comparison between CROWD_SZ and other crowd datasets.

Dataset		Resolution	Images	Min	Max	Multi-view	Density change
Shanghaitech	Part A	different	482	33	3139	✓	✓
	Part B	768*1024	716	9	578		
UCF_CC_50		different	50	94	4543	✓	✗
UCSD		158*238	2000	11	46	✗	✗
Mall		480*640	2000	13	53	✗	✓
WorldExpo_10		576*720	3980	1	253	✓	✓
CROWD_SZ		different	5610	1	673	✓	✓

CROWD_SZ: A Large-scale Multi-view Crowd Counting Semantic Dataset

Jiajia Wu[1], Yufeng Lin[2], Cheng Wu[1], Jin Zhang[1], Lijun Zhang[1]

[1] School of Rail Transportation, Soochow University

[2] Command Center of Suzhou, Industrial Park Public Security Bureau

Contact email: chengwu@suda.edu.cn

Soochow University

IARIA