

## Single Image Crowd Person Counting

In our project we aim to work over the crowd person counting. This is an interesting problem it has applications in safe monitoring, disaster management, design of public places, intelligence gathering, virtual environment, forensic search, among others. The problem can be defined as follow: given an image find the number of persons that appears, as can be see in figure 1 the scenarios are really general and challenging.

The work of Wang et al. [1] uses the well known Alexnet architecture and changes the last layer for a unique neuron and applies transfer learning with data augmentation over this network, Zhang et al. [2] proposed ShanghaiTech dataset in order to boost the use of deep learning over this problem and Shang et al. [3] used GoogleNet and a Long Short Term Memory(LSTM) for this problem. Inspired by these works we intend to use an available network, initially we consider Caffenet, GoogleNet, YOLO or Resnet in order to apply transfer learning on this problem. We are also considering data augmentation with preprocessing as similar to [2].

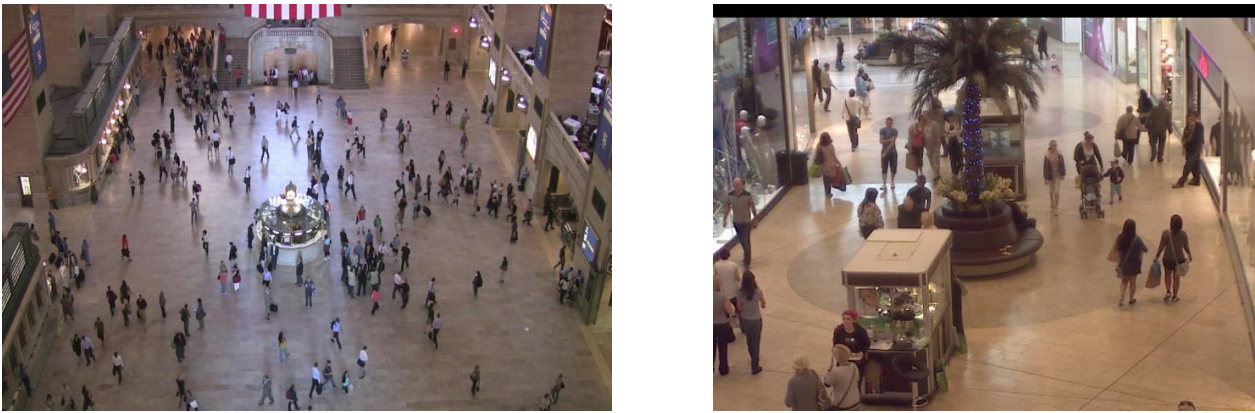


Figure 1: Examples of inputs for the crowd person counting.

In order to test our method we will use the datasets available for the problem, a technical resumo is shown in table 1

Dataset	Nro. of images	Resolution	Min.	Ave.	Max.	Total count
UCSD	2000	158x238	11	25	46	49,885
Mall	2000	320x240	13	-	53	62, 325
UCF_CC_50	50	Varied	94	1279	4543	63,974
WorldExpo'10	3980	576x720	1	50	253	199,923
ShanghaiTech Part A	482	Varied	33	501	3139	241,677
ShanghaiTech Part B	716	768 x 1024	9	123	578	88,488

Table 1: Available datasets for crowd person counting.

## References

- [1] Chuan Wang, Hua Zhang, Liang Yang, Si Liu, and Xiaochun Cao. Deep people counting in extremely dense crowds. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 1299–1302. ACM, 2015.
- [2] Yingying Zhang, Desen Zhou, Siqin Chen, Shenghua Gao, and Yi Ma. Single-image crowd counting via multi-column convolutional neural network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 589–597, 2016.
- [3] Chong Shang, Haizhou Ai, and Bo Bai. End-to-end crowd counting via joint learning local and global count. In *Image Processing (ICIP), 2016 IEEE International Conference on*, pages 1215–1219. IEEE, 2016.