# Crowd Counting System Using Unmanned Aerial Vehicle (UAV)

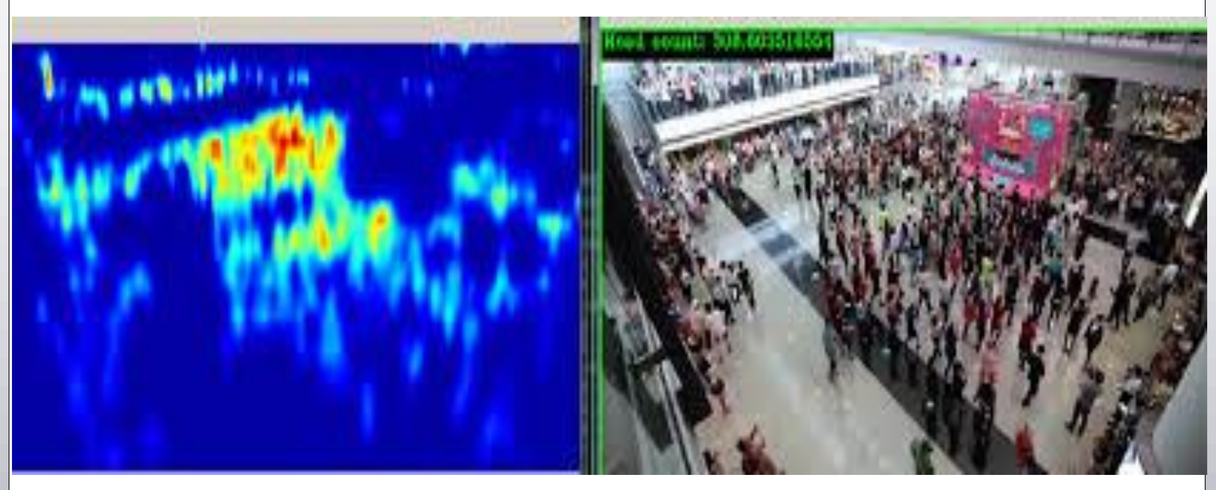
**-by Prince Makwana and Vibhor Choudhary**
Under the supervision of
**Dr. Gaurav Singal**

## Abstract

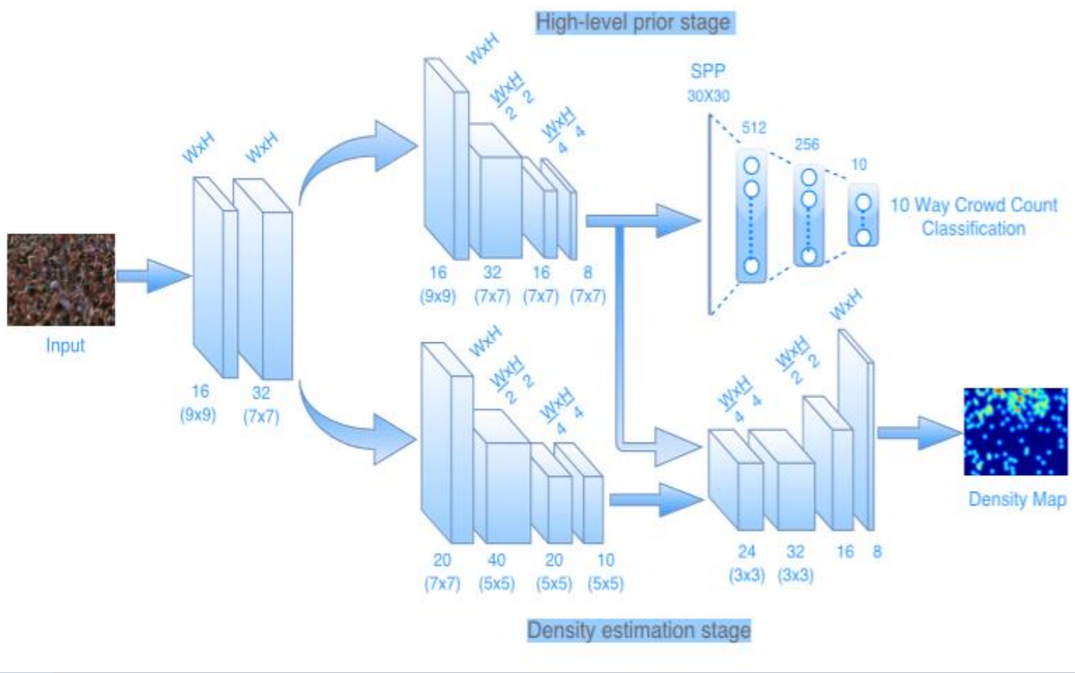
Unmanned Aerial Vehicle (UAV) has created much significance in recent times. It is the booming technology which are mostly used in military applications, their use is rapidly expanding to commercial, scientific, agricultural and other applications. In this research, we proposed a model to estimate the density of crowd from aerial view. Our model used Gaussian Kernels to create the heat map (density map), which is used to estimate crowd in an image. We had deployed our model on raspberry pi to estimate crowd in real time from aerial view which has significant accuracy.

## Introduction

Crowd Analysis has gained a lot of interest in recent years due to it's variety of applications such as video surveillance, public safety design and traffic monitoring. Estimating crowd is difficult task due large variations in scale and appearance of objects that occurs due to severe perspective distortion of the scene. Many different approaches are there such as density-estimation, regression-based, detection-based approach. In our research, we used density-estimation approach because this approach has got better accuracy in previous researches. For this, we used pre-trained model of Multi-Cascaded CNN of ShanghaiTech.



## Proposed Method




## Gaussian Kernel Density Estimator Function

$$\widehat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i) = \frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{x - x_i}{h}\Big),$$

where, K is kernel, a non-negative function and h>0 is smoothing function also called estimator.

The proposed model is developed in pytorch. We have given input images and its ground truth value numpy array as labels to network. Network predicts the density map of input image and maps it with its corresponding ground truth value linearly.

**Link for pretrained ShanghaiTech A model:**
https://www.dropbox.com/s/irho4laltre9ir5/cmtl_shtechA_204.h5?dl=0

## Experimental Results and Discussion

For the purpose of evaluation, the standard metrics used by many existing methods for crowd counting were used. These metrics are defined as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - y_i'|, \quad MSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|y_i - y_i'|^2},$$

### Training of Model

```
epoch: 1999, step 1000, Time: 0.0000s, gt_cnt:  5.8, et_cnt:  3.6
epoch: 1999, step 1500, Time: 0.0000s, gt_cnt: 201.7, et_cnt: 179.3
epoch: 1999, step 2000, Time: 0.0000s, gt_cnt:  7.6, et_cnt: 13.3
epoch: 2000, step    0, Time: 0.0000s, gt_cnt: 26.4, et_cnt: 33.5
epoch: 2000, step  500, Time: 0.0000s, gt_cnt: 45.8, et_cnt: 47.3
epoch: 2000, step 1000, Time: 0.0000s, gt_cnt:  5.8, et_cnt:  2.7
epoch: 2000, step 1500, Time: 0.0000s, gt_cnt: 201.7, et_cnt: 194.1
epoch: 2000, step 2000, Time: 0.0000s, gt_cnt:  7.6, et_cnt: 14.0
EPOCH: 2000, MAE: 9.5, MSE: 17.7
BEST MAE: 9.0, BEST MSE: 16.9, BEST MODEL: cmtl_shtechA_1372.h5
root@15ef8e3b24b5:/home/dgxuser122/data/Crowd Estimation using CCMTL#
[prince_te0:docker*
```

Best MAE: 9.0, Best MSE: 16.9

### ShanghaiTech Dataset

| | Part A | | Part B | |
|---|---|---|---|---|
| Method | MAE | MSE | MAE | MSE |
| MCNN [32] | 110.2 | 173.2 | 26.4 | 41.3 |
| Single stage CNN 1 | 130.4 | 190.9 | 29.3 | 40.5 |
| Proposed method | 9.0 | 16.9 | 18.9 | 28.3 |

## Conclusions

Crowd Counting is great area of research in different areas as estimating crowd is crucial for public safety, monitoring traffic, in political gatherings and rallies.

In this project, we used pretrained model and experimentally evaluated on two datasets. Our results show that, creating density maps of images is very efficient technique to estimate crowd.

## References

https://en.wikipedia.org/wiki/Kernel_density_estimation
https://www.crcv.ucf.edu/data/ucf-cc-50/
https://www.dropbox.com/s/fipgjqxl7uj8hd5/ShanghaiTech.zip?dl=0

Team No. - 44