**Single-Image Crowd Counting via Multi-Column Convolutional Neural Network***Yingying Zhang, Desen Zhou, Siqin Chen, Shenghua Gao, Yi Ma*

Methods:  
The authors propose a Multi-column Convolutional Neural Network (MCNN) that estimates crowd density using density maps. Each column in the network uses convolutional filters of different sizes, allowing the model to adapt to heads of varying scales. The architecture consists of three parallel CNN branches with identical structures but distinct receptive field sizes.  
They also compare their model to other methods, such as:

* LBP+RR (Local Binary Pattern + Ridge Regression): A traditional method where features are extracted from the image and then used to predict the crowd count via ridge regression.
* MCNN-CCR: An extended version of MCNN that includes an additional regression module for improved crowd counting.
* Zhang et al. [33]: Another CNN-based method designed to handle variations in scenes for crowd estimation.

Dataset:  
The model is trained and evaluated on the ShanghaiTech dataset, which contains 1,198 annotated images with a total of 330,165 labeled head positions. The dataset is split into two parts: Part A (more crowded, 300 training and 182 testing images) and Part B (less crowded, 400 training and 316 testing images). The model can be adapted to another dataset, so they also used the UCF CC 50 dataset, the UCSD dataset and the WorldExpo’10 dataset for comparing MCNN but with other methods.

Performance is measured using Mean Absolute Error (MAE) and Mean Squared Error (MSE). The results show that MCNN outperforms other approaches significantly, achieving the best MAE and MSE across both parts of the ShanghaiTech dataset.

**Chapter by chapter short description**

**Abstract**

They say about proposing the MCNN model and why this model and introduce us to experiments with the model. MCNN outperforms other models and trained on one dataset can be used on other datasets.

**Introduction**

They say about existing methods and their drawbacks. They propose new solution and make assumptions about new method. They also introduce us to dataset.

**Multi-column CNN for Crowd Counting**

* **Density map based crowd counting**

They choose density map instead of head count.

* **Density map via geometry-adaptive kernels**

They choose kernel and explain the operation of it.

* **Multi-column CNN for density map estimation**

They say about the multi-column idea and a perspective problem on images.

* **Optimization of MCNN**

They say about pre-training the model.

* **Transfer learning setting**

They use transfer learning and tunel ast few layers and the first layers are pre-trained from domain.

**Experiments**

* **Evaluation metric**

Evaluation using MAE and MSE

* **ShanghaiTech dataset**

Dateset description, comparison of methods

* **The UCF CC 50 dataset**

Dateset description, comparison of methods

* **The UCSD dataset**

Dateset description, comparison of methods

* **The WorldExpo’10 dataset**

Dateset description, comparison of methods

* **Evaluation on transfer learning**

Something about pre-trained and non-pre-trained model

**Conclusion**

Summary of the method.