# Image and Video Analysis

## Project 1: Crowd Monitoring

12 Dec 2023

Abdulrahman Alsatouf & Amelia Yolanda

# Crowd Counting

- Calculating the number of objects in a single image, used for e.g. urban planning, public safety.
- Crowd counting technique categories:
  - detection-based,
  - regression-based,
  - density estimation,
  - CNN-based density estimation.

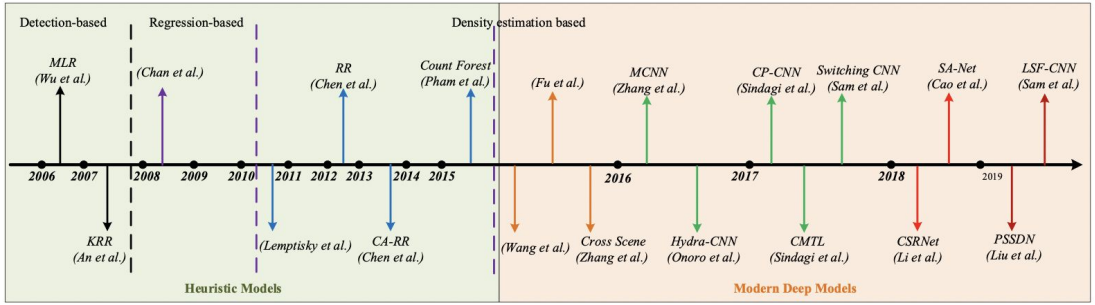



Fig. 1: A brief chronology of crowd counting. The first incorporation of deep learning techniques for crowd counting is from 2015. See Section 1 for more detailed description. Milestone models in this figure: MLR [52], KRR [53], Chan et al. [27], Lemptisky et al. [16], RR [40], CA-RR [54], Count Forest [48], Wang et al. [49], Fu et al. [50], Cross scene [51], MCNN [1], Hydra-CNN [2], CP-CNN [6], CMTL [55], switching CNN [5], CSRNet [12], SANet [11], PSSDN [56] and LSF-CNN [57]. The trend in the past few years has been designing crowd counting models based on multi-column (in green), single-column (in red) network architecture and object localization or tracking depending on counting techniques (in crimson), which are either contemporary and potential direction in future. While traditional heuristic methods are highlighted with the blue-shaded area and the modern CNN-based density estimation and crowd counting models are with the red-shaded backgrounds, respectively.

Gao, Guangshuai, et al. “Cnn-based density estimation and crowd counting: A survey.” arXiv preprint arXiv:2003.12783 (2020).

# Dataset

Image 1

Image 2

Image 5

Image 6

Image 3

Image 4

Image 7

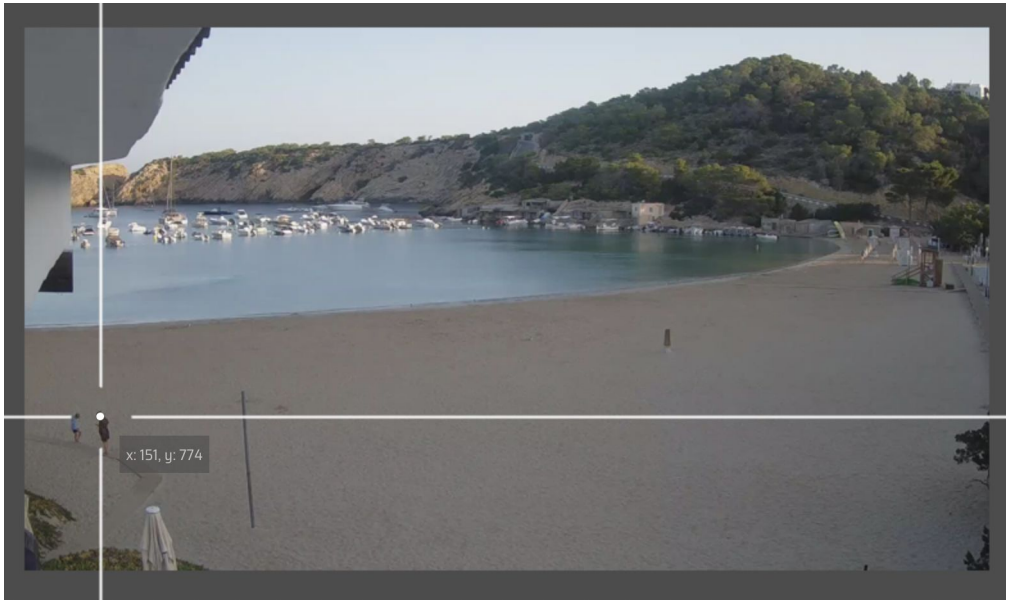
Image 8

Image 9

Image 10

# Annotation Process

https://www.makesense.ai/ > Object Detection




| object | x | y | image |
|---|---|---|---|
| Person | 148 | 766 | 1660284000.jpg |
| Person | 105 | 769 | 1660284000.jpg |
| Person | 621 | 719 | 1660287600.jpg |
| Person | 865 | 746 | 1660287600.jpg |
| Person | 929 | 608 | 1660287600.jpg |
| Person | 949 | 608 | 1660287600.jpg |
| Person | 820 | 513 | 1660287600.jpg |
| Person | 582 | 531 | 1660287600.jpg |
| Person | 1128 | 730 | 1660287600.jpg |

# Algorithm

- Separate images to exclude background (noise)

Cropped Image

Background Image

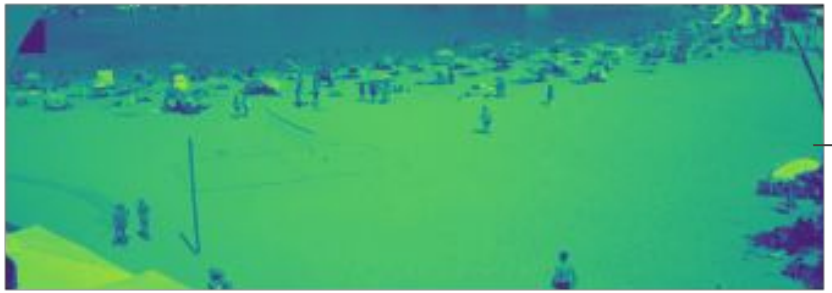
Cropped Image

Background Image

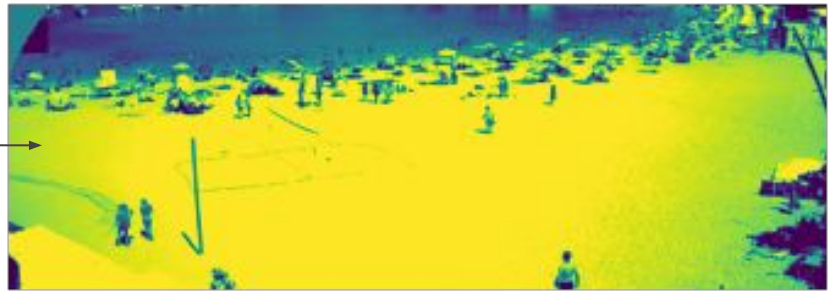
# Algorithm

Cropped Image

Grayscale Image

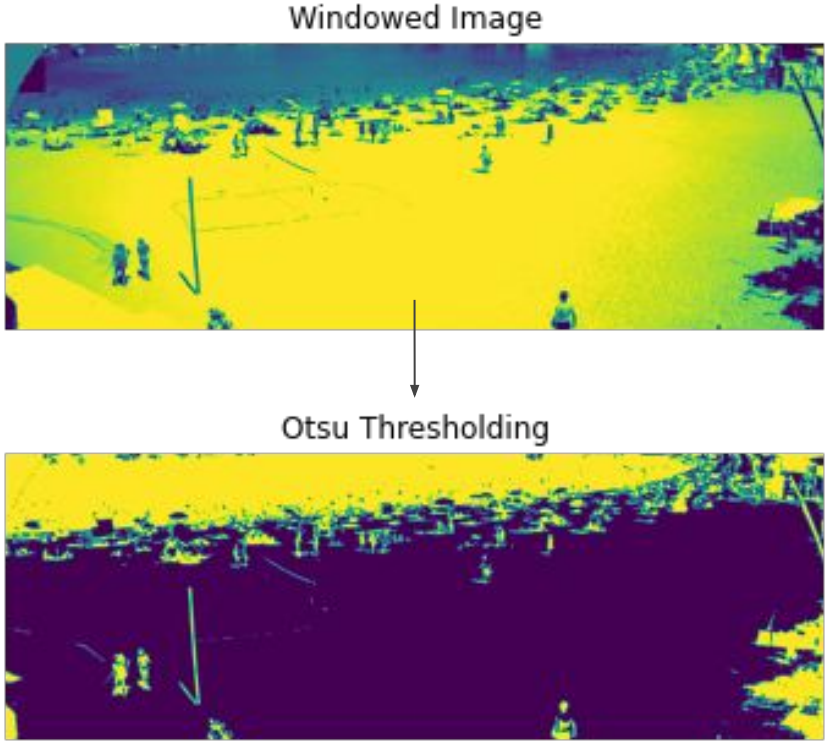


Windowed Image



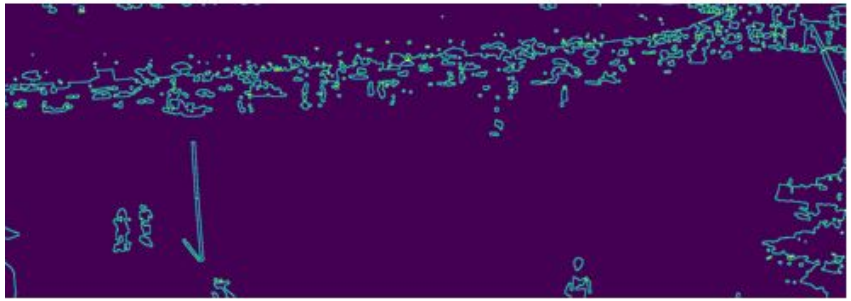
- Apply grayscale
- Apply windowing, for more contrast

# Algorithm




- Apply Otsu’s Thresholding & invert to negative
- Apply morphological opening & closing

# Algorithm



- Apply Sobel edge detection
- Filter contours based on area and aspect ratio
- Calculate the number of persons on the image

# Result

Output image with centroid of bounding box

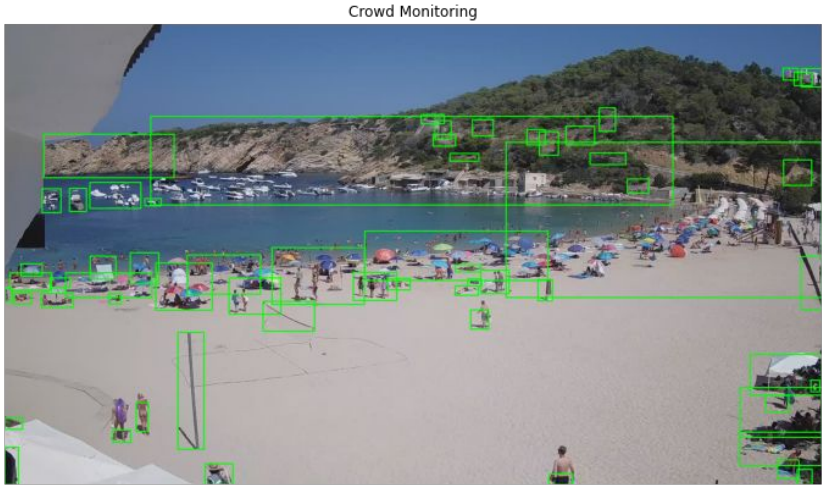
# Validation

- Image-level: Mean Squared Error of 3097.60
- Person-level: Average accuracy per image of 42.47%

| Image Number | Annotated Image | Detected from Algorithm | True Positives |
|---|---|---|---|
| 1 | 2. | 4 | 0 |
| 2 | 15. | 18 | 7 |
| 3 | 34. | 17 | 8 |
| 4 | 52. | 29 | 26 |
| 5 | 56. | 34 | 28 |
| 6 | 56. | 26 | 18 |
| 7 | 58. | 39 | 37 |
| 8 | 64. | 59 | 43 |
| 9 | 64. | 45 | 35 |
| 10 | 79. | 33 | 29 |

# What is interesting?

1. **Image Annotation:** not easy to decide whether it's a person or inanimate object.
2. **Algorithm:** trial and error process to achieve the current results.
3. **Result:** using centroid of bounding box allows us to obtain more accurate results.



# What needs improvement?

---

1. **Annotation process**: set a criteria to define proper label.
2. **Algorithm:** consider more complex algorithms (e.g. ML)
3. **Evaluation:** try different performance measures (e.g. precision and recall)

# Thank you