11761 – Image and Video Analysis

Project 1

In summer, beaches are monitored by lifeguards, who gather information about the occupancy, sea state, wind, etc. Image processing techniques can be used to automate this task. In this project, we will estimate the amount of people present in a beach (in computer vision, "crowd counting"), focusing on all aspects of an image processing project: data annotation, algorithm design and implementation, and validation.

1. **Data annotation.** Manually annotate images, which can be done with the online tool
   https://www.makesense.ai/
   a. Actions > Edit labels > Add label "Person"
   b. With "Point" selected, click on every person's head
      i. (If in doubt, click if you are 90% sure it is a person)
   c. Actions > Export annotations > Single CSV file.
2. **Algorithm design and implementation** to i) identify every person and ii) count the number of people in every photograph. This is (intentionally) left as a very open-ended question: you should decide which techniques to apply (and why!) Some ideas:
   a. Image averaging.
   b. Histogram equalization/CLAHE.
   c. Pixel-wise comparison with "empty" beach (image averaging).
   d. Spatial filters: derivatives, laplacian.
   e. Kernels specifically designed for crowd counting
   f. Morphological thinning.
   g. Edge detectors.
   h. ...
3. **Validation.**
   a. Image-level: consider the mean squared error (MSE) to compare the number of people in the manual annotations and the number of people detected for the algorithm.
   b. Person-level: consider the average accuracy to check whether people was correctly identified.

For more information, check the following articles:

[1] Gao, Guangshuai, et al. "Cnn-based density estimation and crowd counting: A survey." arXiv preprint arXiv:2003.12783 (2020).
[2] Idrees, Haroon, et al. "Multi-source multi-scale counting in extremely dense crowd images." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2013.
[3] Chan, Antoni B., Zhang-Sheng John Liang, and Nuno Vasconcelos. "Privacy preserving crowd monitoring: Counting people without people models or tracking." 2008 IEEE conference on computer vision and pattern recognition. IEEE, 2008.

You should submit:
- Functional source code, in a public github repository,
- A 4-page summary, which should focus on the technical part: the algorithms, their implementation, and their performance.
- Material for the oral presentation (slides), which should focus on the process: what decisions were taken and why, which strengths/shortcomings were found, and a critical discussions of the results.