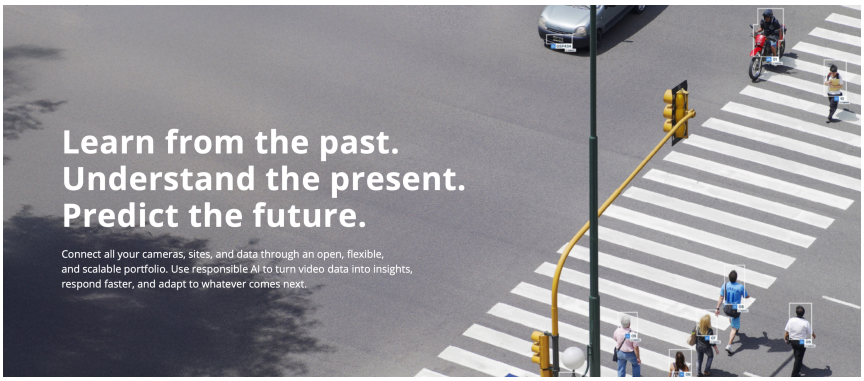
Are you interested in collaborating with a market leading company on developing tomorrows' video analytics algorithms?

Milestone Systems A/S, part of the Canon Group, is a world-leading provider of video management systems, video analytics, and VSaaS (Video Surveillance as a Service) with more than 500 000 installations around the world. The company is headquartered in Denmark but has sales offices in all the continents. We are about 1500 employees.
Milestone Systems A/S main product, XProtect, records video frames of connected cameras on recorder(s) and provides the possibility to retrieve the recorded data efficiently. Furthermore, it provides an open platform for third party video analytics companies to develop their algorithms without struggling with different camera drivers. Third parties of Milestone systems span a wide range of video analytics companies for many different purposes, for example, looking at people for security and safety purposes, healthcare, etc.

Furthermore, Milestone's own video analytics suit, BriefCam, provides one of the most selling and advanced set of video analytics algorithms.

On the following pages, multiple different topics of our interest within video analytics algorithms is presented. **For each project we deliver data and a proof-of-concept code that students can start with.** If you are interested in any of these topics please reach out to Kamal Nasrollahi, Professor at AAU and Director of Research at Milestone Systems at kn@create.aau.dk.



Generative-AI-based advanced data augmentation methods for model training

Background

Robust performance in computer vision models often hinges on the diversity and quality of training data. However, collecting large-scale, varied datasets—especially for tasks involving people or vehicles—can be costly and time-consuming. Data augmentation offers a scalable solution by synthetically expanding existing datasets.

Traditional augmentation techniques (e.g., flipping, cropping, rotation) are limited in their ability to simulate real-world variability. This project explores complex augmentations that modify pose, background, and lighting conditions to generate realistic and diverse training samples. In addition, we consider appearance-based transformations, including gender and ethnicity modification, clothing changes, and even alterations in overall identity or facial appearance. Such transformations can provide richer diversity in training data, reducing bias and improving model generalization across populations and environments.

These augmentation strategies are relevant for many applications such as various classifiers,recognition tasks, and object-specific detectors (YOLO-based). Furthermore, the project invites exploration of additional augmentation ideas—whether environmental, behavioral, or appearance-based—that could further enhance the robustness of vision systems under real-world variability.

Objectives

- Develop a pipeline for generating augmented data by manipulating pose, background, and lighting in existing images.
- Evaluate the impact of these augmentations on model performance for tasks such as classification and detection.
- Explore domain-specific augmentation strategies for both **human-centric** and **vehicle-centric** datasets.

Deliverables

1. **Literature Review -** Overview of current data augmentation techniques, with emphasis on pose transformation, background replacement, and lighting simulation.
2. **Augmentation Framework -** A modular system that applies complex augmentations using tools such as 3D pose estimation, GANs, or neural rendering.
3. **Dataset Expansion -** Augmented versions of an existing dataset annotated and ready for training.
4. **Domain Adaptation Study -** Insights into how augmentations affect generalization across domains (e.g., indoor vs. outdoor scenes, day vs. night).

Super-Resolution for Improving Low Resolution People Re-Identification

Background

People re-identification is an important task in various Video Analytics applications. Its purpose it to identify the same people in various scenes as long as they have not changed their clothes regardless of changes in lighting, environments, distance to the camera and the angle towards it, resolution, etc.

This technology can be used to track suspects or missing people throughout a city or analyze customers' movement around a shopping mall to improve their experience or optimize sales.

The problem is that in many scenarios the people in the camera's field-of-view (FOV) are too small, i.e. in low resolution, such that the re-identification network cannot work properly, and its accuracy drops.

A possible solution can be to take several low-resolution samples from a person's track in a camera's FOV and use super-resolution algorithms to generate one high-resolution image of him to feed the re-identification network.

Objectives

Enhance the accuracy of a re-identification network on low-resolution people using super-resolution. The student will be given a pre-trained people re-identification network and a dataset containing people walking across multiple cameras in various sites.

Deliverables

- Literature review: Summary of current super-resolution methods and their contribution to various tasks, specifically to people re-identification
- Implementation of one or more super-resolution algorithms.

- Measuring the effect of the chosen super-resolution algorithms on the accuracy of the re-identification network on the given dataset