OBJECTIVE

Without any labelled data, using only images of a traffic scene taken by a fixed camera:
- Detect all objects in the scene which are not part of the background
- Associate to each object its mask and a low dimensional feature vector

Why unsupervised?
- manage all kind of images (thermal, black and white, unusual camera locations..)
- Robustness

CHALLENGES

Scene complexity:
- Changing background (light, moving trees, etc.)
- High number of objects
- Complex object shapes

Specificity of the problem:
- Some objects are very small (pedestrians)
- Cars stopped at traffic lights should not be considered as background

STATE OF THE ART

Background subraction:
SOTA: SemanticBGS, IUTIS, SubSENSE, WeSamBE, PAWCS
- Either not differentiable or supervised
- Low mask quality
- Cannot handle small objects

Unsupervised object detection:
SOTA: SPAIR (2019), SPACE (2020)
- Fails when object sizes vary a lot
- Fails on real images with complex shapes

ACHIEVEMENTS

Background subraction:
- Real-time background reconstruction and subraction of 1280 x720 real-world images using a fully differentiable network

Input: traffic scene

Background reconstruction

Background subraction (normalized)

Differentiabilty → can be integrated inside any neural network performing vision tasks