1 PROBLEM

Traditional semantic segmentation methods can recognize at test time only the semantic classes that are present in the training set (closed set). This is a significant limitation for segmentation algorithms mounted on intelligent autonomous systems. Regardless of how many classes the system has seen at training time, it is inevitable that unexpected, unknown objects (anomalies or out-of-distribution object) will appear at test time. Detecting and localizing such objects is crucial for safety critical applications such as perception for automated driving, especially if they appear on the road ahead.

Anomaly segmentation (AS) aims to segment objects unseen to the model. (a) Road Anomaly, (b) Road Obstacle and (c) anomalies (light-blue) are all regions unmatched with any class prototype.
2 Existing Methods

1. Bayesian Deep Learning (BDL) based methods [Lakshminarayanan et al. (2016), Gal & Ghahramani (2016), Atanov et al. (2019)] are used to estimate uncertainty of predictions as anomalous image regions are expected to correlate with high uncertainty. BDL model parameters are treated as distributions.

2. Anomaly segmentation via generative models [Xia et al. (2020), Lis et al. (2019)] that resynthesize the original input image. The intuition is that the reconstructed images will better preserve the appearance of regions containing known objects than that of unknown regions. Pixel-wise anomaly detection is then performed by identifying the discrepancies between the original and reconstructed image.

3. Other Methods + Code: [https://segmentmeifyoucan.com/leaderboard](https://segmentmeifyoucan.com/leaderboard)

4. Benchmark for Anomaly Segmentation: [https://segmentmeifyoucan.com/](https://segmentmeifyoucan.com/)

3 Proposed Idea

Recent weakly supervised approaches that only require class tags have been proposed by Sawatzky et al. (2019). Sawatzky et. al.’s model use image caption to generate class tags. To leverage textual context, a multi-modal network that learns a joint embedding of the visual representation of the image and the textual representation of the caption is needed.

In this internship, we are interested mainly in segmenting the anomaly objects which is never seen. Hence, we propose to rely on weak form of supervision using class tag or scene caption description as a prior information beside the closed set semantic classes.

4 Goals

1. Study state of the art method related to semantic Segmentation.
2. Reproduce some relevant work results using their online code.
3. Generate image caption for a scene using existing NLP model.
4. Use textual representation as a prior to aid the segmentation of anomalies. Such representation can be embedded on the top of any existing Semantic Representation method, or re-build a new method.

5 Skills

1. Image Processing.
2. Machine Learning and in particular Deep Learning (DL) (if not, the first month will be dedicated to study the basics of DL).
4. Knowledge in DL libraries is appreciated such as: TensorFlow or PyTorch.
5. Command Line skills is a plus.
REFERENCES


