S5 – Video Anomaly Detection and Understanding

6-month internship @ CEA List

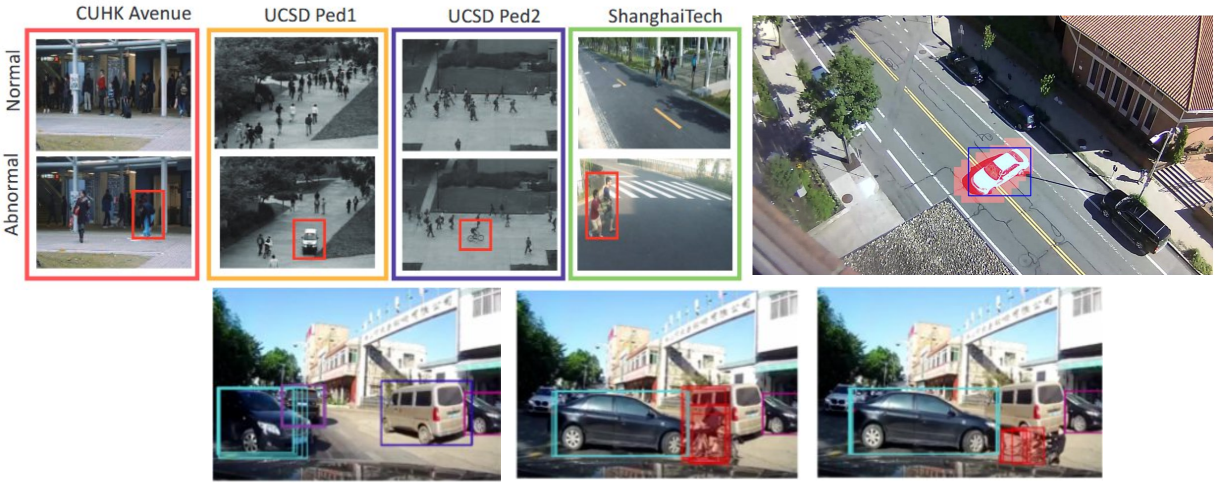
Internship context

Based in Saclay (Essonne), the LIST is one of the two institutes of CEA Tech, the technological research division of the CEA. Dedicated to intelligent digital systems, its mission is to carry out technological developments of excellence on behalf of industrial partners in order to create value.

Within the LIST, the Laboratory of Vision and Learning for Scene Analysis (LVA) conducts research in the field of computer vision and artificial intelligence for the perception of intelligent and autonomous systems. The laboratory's research themes include visual recognition, behavior and activity analysis, large-scale automatic annotation, and perception and decision models. These technologies are applied in major sectors such as security, mobility, advanced manufacturing, healthcare, and sports...

Missions

Video Anomaly Detection (VAD) has numerous applications in video monitoring, traffic control, industrial inspection or healthcare. This is an active research problem that seeks to automatically detect and localize abnormal events in videos, such as violence, dangerous or suspicious activities. However, the task is complicated due to the diverse and often unclear nature of abnormal events, their rare occurrence and the lack of training data and annotations.



Top-left: some of public VAD benchmarks for video surveillance; top-right: example of a static camera monitoring an urban scene from StreetScene dataset; bottom: normal and abnormal images from driving videos in DoTa dataset

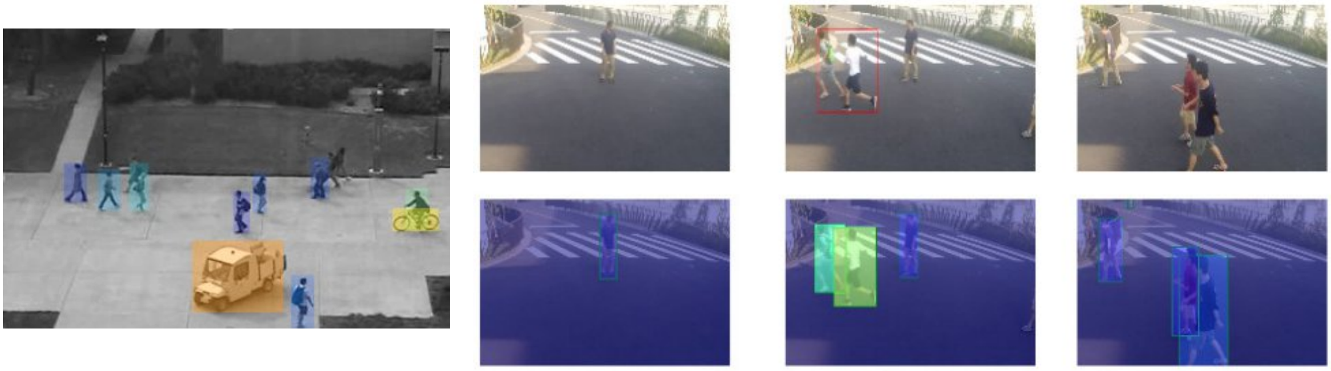
The most common approach to deal with VAD is One-Class learning [1] which consists in training anomaly detection models only on normal data, i.e. free from anomalies, to perform different auxiliary tasks such as reconstruction and prediction. The main assumption is that trained on normal data, the model fails to correctly reconstruct or predict video frames containing anomalies. However, such methods perform well only on rather simple datasets where anomalies can be defined by their visual appearance or motion, and fail on videos where abnormal events contain high-level semantic information (long-range trajectories, interactions between individuals or objects). Other approaches exist to deal with VAD such as Weakly-Supervised [2] or Few-Shot learning methods [3]. Such methods use some examples of anomalies in training which helps to capture more sophisticated anomalies which require understanding of high-level semantic information [4,5].

Recently, Vision Language models [6,7] have gained a lot of popularity because of their ability to deal both with images and text. Their applications include Visual Question Answering (VQA), image captioning, and Text-to-Image search. The advantage of using such models in VAD lies in the fact that they not only detect video anomalies, but also provide their description which helps to better understand and explain the anomalies occurred [8].

In this internship, we will aim to work with VAD methods that deal with anomalies requiring high-level semantic information by using some abnormal samples in training. Besides the anomaly detection task, we will address the problem of video anomaly understanding by exploiting VLM models.

References:

[1] Masked multi-prediction for multi-aspect anomaly detection, Y. Naji *et. al.*, Trans. Mach. Learn. Res. 2024
[2] Real-Time Weakly Supervised Video Anomaly Detection, H. Karim *et. al.*, WACV, 2024
[3] Few-shot scene-adaptive anomaly detection, Y. Lu *et. al.*, ECCV, 2020
[4] Not only Look, but also Listen: Learning Multimodal Violence Detection under Weak Supervision, P. Wu *et. al.*, ECCV, 2020
[5] Real-world Anomaly Detection in Surveillance Videos, W. Sultani, CVPR 2018
[6] Visual Instruction Tuning, H. Liu *et. al.*, NeurIPS 2023
[7]Learning Transferable Visual Models From Natural Language Supervision, A. Radford et. al., ICML, 2021.
[8] Text Prompt with Normality Guidance for Weakly Supervised Video Anomaly Detection, Z. Yang et. al., CVPR 2024



Left: examples of detected anomalies in UCSDped2 dataset; right: images and visualization of the corresponding VAD model predictions on Avenue dataset

Internship objectives

The proposed internship has the following objectives:

- Study state-of-the-art of Video Anomaly Detection and Explainability (weakly-supervised, few-shot learning and VLMs);
- Identify promising methods as baselines and perform experiments on public benchmarks.
- Propose improvements to the baselines and evaluate.
- Depending on the obtained results, the contributions of this internship may lead to an international conference or workshop publication.

Qualifications

- Students in their 5th year of studies (M2)
- Computer vision skills
- Machine learning skills (deep learning, LLM, VLM, generative AI…)
- Python proficiency in a deep learning framework (especially PyTorch or TensorFlow)

Job-related benefits

Join CEA List and LVA as an intern to:

- Work in one of the most innovative research organizations in the world (ranked in the global top 100, top 3 in France), addressing societal challenges to build the world of tomorrow
- Discover a rich ecosystem: privileged connections between the industrial and academic sectors
- Conduct research in an environment where autonomy and creativity are recognized, and where valorizing results is encouraged (publication of scientific articles, patents, and sharing of open-source code whenever possible).
- Join a young and dynamic team made up of research engineers, PhD students, post-doctoral researchers, and interns.
- Benefit from an internal computing infrastructure equipped with around 300 state-of-the-art GPUs.
- Receive a stipend between €1300 and €1400 per month.
- Have the opportunity to continue with a PhD or as a research engineer after the internship.
- Have the possibility of remote work.
- Receive a 75% (instead of 50%) reimbursement on public transportation costs, and benefit from the "mobili-jeune" aid to reduce rent costs…

To apply, contact the laboratory with a CV and cover letter: lva-stages@cea.fr