First Report: Problem Statement

Refer to PR1 as a guideline
See the website

LE MOAL STEVEN

–

What is your research topic?

Title + Set your aims and objectives

– Title: Detection and prevention of urban violence events in a distributed video surveillance system

Aims: Conception of a method capable of tracking meaningful objects and detecting events (especially violent behavior) happening in real-time in a multi-camera video system

Objectives:

- Extraction of objects and meaningful events (firstly violence than other human activity) from a video sequence using existing methods to create a semantic scene representation in a graph (with the PhD student)
- Experiment with urban violence prediction and tracking for distributed video surveillance system (multiple views) using the constructed graphs.
- Describe the result and propose a viable solution

Formulate research questions

–

- How can we learn and construct the relationship between high-level semantic representation of events and low-level (and mid-level) detections?
- How can we extract realistic descriptions and interactions from video footages?
- How do we narrow the semantic gap [1]?
- How can we link different scene semantic representation into a real-world description (meaningful interpretation)?

–

[1] Semantic gap: "*Understanding semantics is the most fundamental step in all kinds of computer vision problems as it paves the way for general artificial intelligence. The semantic gap is a general term widely used in content-based image retrieval (CBIR). It is defined in [134] as follows. "Humans tend to use high-level concepts in everyday life. However, what current computer vision techniques can automatically extract from image are mostly low-level features. In constrained applications, such as the human face and fingerprint, it is possible to link the low-level features to high-level concepts (faces or fingerprints). In a general setting, however, the low-level features do not have*

a direct link to the high-level concepts."

Article: Bridging Gap between Image Pixels and Semantics via Supervision: A Survey
Author: Jiali Duan and C.-C. Jay Kuo.

–

Your first literature review.

Present 3 key articles concerning your topic (reference + short summary – one paragraph – + how it guides you for your research)

Note: Here are 3 articles which give a good representation of my personal literature review and the direction of my work/research.

– Write here –

Article 1 –

Title: **State-of-the-art violence detection techniques in video surveillance security systems: a systematic review**
Authors: *Batyrkhan Omarov, Sergazi Narynov, Zhandos Zhumanov, Aidana Gumar and Mariyam Khassanova*
Link: link
Type: Survey
Journal: PeerJ Computer Science
Date of publication: April 2022

Reviews:
This paper provided an assessment of video violence detection problems described in the state-of-the-art researches (qualitative and quantitative) in terms of methodology (procedure), including datasets, and performance metrics. Additionally, challenges and potential future directions are also discussed.

(Why did I select this article? Getting an overview of my topic.)

Complementary article: **Human Action Recognition and Prediction: A Survey** (link)

– Write here –

Article 2 –

Title: **A Comprehensive Survey of Scene Graphs: Generation and Application**
Authors: *Xiaojun Chang, Pengzhen Ren, Pengfei Xu, Zhihui Li, Xiaojiang Chen, and Alex Hauptmann*
Link : link
Type: Survey

Journal: IEEE Transactions on pattern analysis and machine intelligence
Date of publication: March 2021
Reviews:
This paper provides an assessment of the existing scene graph generation and application approaches in the state of the arts. It discussed the scene graph generation approaches for visual relationship detection, as well as scene graph applications based on specific visual tasks. Furthermore, challenges in present scene graph generating methods and potential future directions are also discussed.

(Why did I select this article? Looking at the subject of action recognition using a promising approach.)

Complementary article: **A Survey on Graph Neural Networks and Graph Transformers in Computer Vision: A Task-Oriented Perspective** (link)

– Write here –

Article 3 –

Title: **Deep multi-view learning methods: A review**
Authors: *Xiaoqiang Yan, Shizhe Hu, Yiqiao Mao, Yangdong Ye and Hui Yu*
Link: link
Type: Survey
Journal and Rank: Neurocomputing
Date of publication: March 2021

Reviews:
This paper provides a review on the deep multi-view learning domain. It described the advancements from traditional methods to deep learning methods. Applications, datasets (mainly used) and performance are discussed. The challenges in deep multi-view learning domain are also addressed, to conclude with future potential directions.

(Why did I select this article? It will help me for the next step of the work.)

Theme that could be included: Transformer, Group activity detection, Constative Learning, Transfer learning, Self-Distillation, Fuzzy Logic (for realistic approach...), Domain Adaptation …...

Context

What is the socio-economic context of your work (3 paragraphs)

_

Violence is a part of the human experience, and its impact may be seen in several ways around the world. Individual (even self-inflicted) or social violence kills millions of people each year, with many more suffering non-fatal injuries. As a result, violence is one of the leading causes of mortality worldwide among persons between the age of 15 and 44 [1].
Violence (both verbal and physical) can occur at any moment and in any location. Violence has expanded swiftly at universities, whether among teaching staff, administration personnel, students, or others. Even still, violence against healthcare professionals is on the rise throughout the world: the World Medical Association defined violence against health employees as" an international emergency that undermines the very foundations of health systems and impacts critically on patient's health" [2].
Violent behavior may occur in any crowded setting, including shopping centers, malls, sports stadiums, prisons etc...
Because of the growth in security concerns throughout the world, the use of video cameras to monitor persons has become critical, and early discovery of these violent behaviors may considerably reduce risk.

[1]: Publication: World report on violence and health
Author: World Health Organization
Year: 2002

[2]: Publication: Violence Against Healthcare Workers: A Worldwide Phenomenon with Serious Consequences
Author: Sandro Vento, Francesca Cainelli, and Alfredo Vallone.

_

Given that violence might strike at any time, relying exclusively on humans to undertake the task of monitoring and recognizing violent or dangerous situations is inadequate (ineffective). Typically, this role has required security personnel to constantly watch a large number of monitors. As a result of human weariness (exhaustion), poor attention, and inexperience, occlusions, violent behavior (including violence) events can be missed. Therefore, to assist security staff and ensure safety, automated surveillance systems that recognize specific pattern such as abnormalities in real time are critical.

_

Due to the repetitive nature of the task of video monitoring and the time necessary, real-time automatic detection of violence is essential to prevent such situations. However, occlusion, poor image resolution, visual noise, fast changes in events and/or in interactions, data encoding, illumination changes, and other factors make the process difficult. A methodology and good practice for distributed video surveillance must be generalized (based on previous works).
_

Timeframe

See your professor to determine the scheduling of your work. The content and deadlines are not final.

Mid-September – Debut October: 2 weeks of problem discovery.
Debut October – Debut November - 1 month of technology review.
Debut November – Debut December: 1 month of analysis of the technology revue and experimental (including methodology) development.
December: 1 month of material testing (Cisco camera).
January and February - 1-2 month of experimentation following the methodology' guidelines
Mid/End February - Proposal of a viable solution.
End February/ Marth - 1 month of writing (reports).

Evaluation criteria

How this first report will be marked

Research topic → 3 points (clarity of the aims and objectives)
Research question → 3 points (clarity, relevance, and discussions)
Literature review → 9 points (3 points per article, clarity, relevance, and discussions of each article)
Context → 3 points (clarity, relevance, and discussions)
Timeframe → 2 points (clarity and feasibility of the scheduling)