

# Reconhecimento de Ações Humanas a partir de Dispositivos Aéreos Não-Tripulados (UAVs)

*Proposta de Projeto*

Orientador: Hugo Proença

## 1 Objectives

Video-based surveillance refers *the act of 'watching a person or a place, esp. a person believed to be involved with criminal activity or a place where criminals gather'* [1]. Over the years, this technology has been used in far more applications than its roots in crime detection, such as traffic control and management of physical infrastructures.

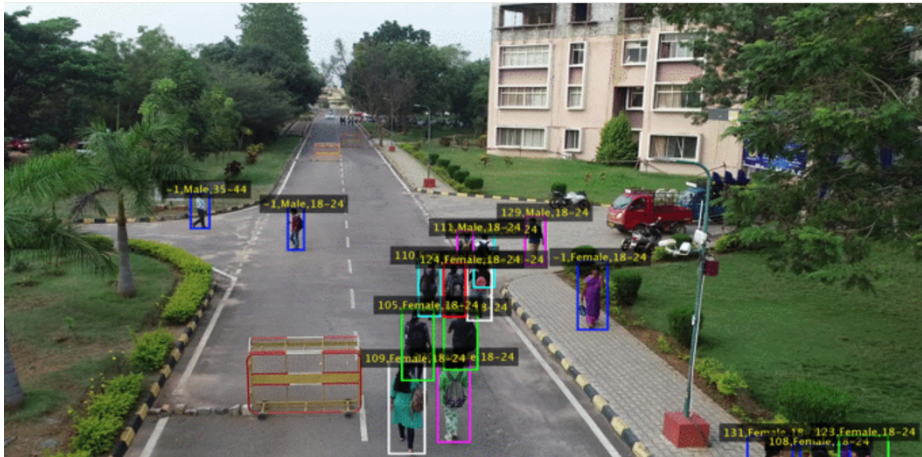
Over the recent years, unmanned aerial vehicles (UAVs, Fig. 1) have been regarded as a potential solution to surveil public spaces, providing a cheap way for data collection, while covering large and difficult-to-reach areas. This kind of solutions can be particularly useful to detect, track and identify objects of interest in crowds, for security/safety purposes[1].



**Figure 1:** Illustration of the kind of Unmanned Aerial Vehicles (UAVs), available at the SOCIA – Soft Computing and Image Analysis Lab. for research/development purposes (DJI Phantom 4).

In particular, this project aims at the development of a framework to automatically infer basic human actions in outdoor surveillance scenes, using video data captured by an UAV. The proof of concept system should be able to:

- Detect and recognize basic human movement types (such as “walking”, “running”, “laying”);
- Detect subjects using specific kind of vehicles (e.g., “skates”, “bicycles”,...)
- Characterize the observed types of movements for each human in the scene, in terms of the “velocity” and “direction” features.



**Figure 2:** Example of one frame of a video sequence captured by an unmanned aerial vehicle (UAV).

At the end, the develop framework should be able to answer to queries, such as:

- “Raise an alarm in case of subjects running in this street”;
- “Raise an alarm in case of subjects riding bikes in this area”;
- “Raise an alarm in case of subjects moving in the wrong direction in this street”.

The developed framework is expected to be built upon deep learning frameworks, and should incorporate at least four modules: 1) “human detection”; 2) “human tracking”; 3) “pedestrian classification”; and 4) “movement characterization”.

## 2 Work Plan

**T1:** Analysis of the existing solutions for pedestrian detection, tracking and movement characterization.

**T2:** Implementation of the chosen approach;

**T3:** Debugging, tests and optimization;

**T4:** Report writing.

## 3 Academic Pre-requisites

- Interest in Machine Learning, Computer Vision and Artificial Intelligence domains;

- Solid programming competence, preferably with expertise in Python language.

## 4 Expected Results

- One computational prototype;
- One technical report describing the work carried out.

## 5 Contacts

Hugo Proença ([hugomcp@di.ubi.pt](mailto:hugomcp@di.ubi.pt))

## References

[1] SV Aruna Kumar, Ehsan Yaghoubi, Abhijit Das, B.S. Harish, Hugo Proença. The P-DESTRE: A Fully Annotated Dataset for Pedestrian Detection, Tracking and Short/Long-term Re-Identification from Aerial Devices. IEEE Transactions on Information Forensics and Security, vol. 16, pag. 1696-1708, doi: 10.1109/TIFS.2020.3040881, 2021.