

I. Mademlis, Prof. Ioannis Pitas Aristotle University of Thessaloniki pitas@csd.auth.gr www.multidrone.eu Version 2.5





- 36 UAV video datasets have been created/assembled within the MULTIDRONE H2020 project.
- Approximate total size: 260 GBs. **To be found in:**

http://www.aiia.csd.auth.gr/LAB_PROJECTS/MULTIDRONE/AUTH_MULTIDRONE_Dataset.html

https://multidrone.eu/

- They are partitioned into two subsets: datasets provided by AUTH and datasets provided by RAI or Deutsche Welle.
- In order to access the desired dataset subset, please complete and sign the corresponding license agreement found in the web page, i.e., from AUTH or from RAI. Subsequently, email it to Prof. loannis Pitas (Aristotle University of Thessaloniki, AUTH) or to Dr. Alberto Messina (RAI), so as to receive the relevant FTP credentials for downloading.

NOTE: For datasets assembled from Youtube videos, only links to the videos and the relevant annotation files, if any, are provided.
 Artificial Intelligence & Information Analysis Lab



- If one uses any part of these datasets in his/her work, he is kindly asked to cite the following papers:
 - I. Mademlis, V. Mygdalis, N.Nikolaidis, M. Montagnuolo, F. Negro, A. Messina and I.Pitas, "High-Level Multiple-UAV Cinematography Tools for Covering Outdoor Events", IEEE Transactions on Broadcasting, vol. 65, no. 3, pp. 627-635, 2019.
 - Mademlis, N.Nikolaidis, A.Tefas, I.Pitas, T. Wagner and A. Messina, "Autonomous UAV Cinematography: A Tutorial and a Formalized Shot-Type Taxonomy", ACM Computing Surveys, vol. 52, issue 5, pp. 105:1-105:33, 2019.





AUTH datasets





DCROWD_VID

 A dataset for visual human crowd detection was assembled from Youtube videos, licensed mainly under Standard Youtube License.

 It is a collection of 53 videos selected by querying the Youtube search engine with specific keywords describing crowded events (e.g., parade, festival, marathon, protests).

 Non-crowded videos have also been gathered by searching for unspecified drone videos.





• SHOT_TYPES

- A dataset containing 46 professional and semi-professional UAV videos was assembled from Youtube material.
- Care was taken to include as many UAV framing shot types and UAV/camera motion types as possible, based on the UAV shot type taxonomy defined in the context of the MULTIDRONE project.





Annotations_boats_ Raw

A dataset for boat detection/tracking was assembled, consisting of 13 Youtube videos (resolution: 1280 × 720) at 25 frames per second. Annotations are not exhaustive, i.e. there may be unannotated objects in the given image frames.







Annotations_Bicycles_Raw

- A dataset for bicycle detection/tracking was assembled, consisting of 7 Youtube videos (resolution: 1280×1080) at 25 frames per second.
- Annotations are not exhaustive, i.e., there may be unannotated objects in the given video frames.





Benchmark_RAI

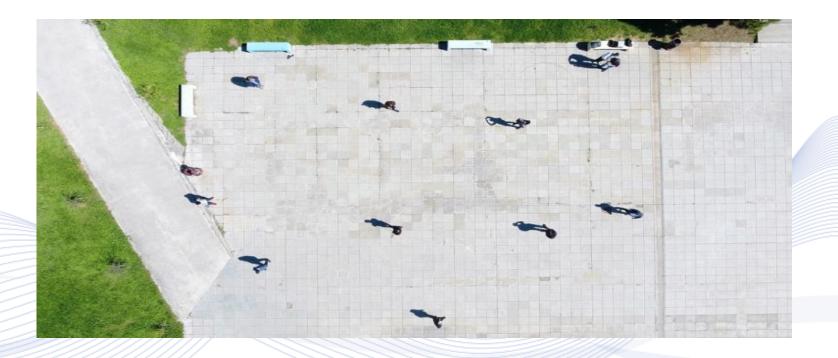
- A dataset for bicycle detection/tracking was prepared by processing/editing and annotating material made available by RAI under the "Giro 2017" MULTIDRONE dataset.
- It is a dataset consisting of two videos (resolutions: 768×432 and 960×540) at 25 frames per second. The videos are from Giro d'Italia TV coverage provided by RAI.
- Annotations are exhaustive, i.e., all objects of a certain class present in a given image are covered by an annotation.





person_detection_UA V

- A visual person detection dataset has been prepared, consisting of two UHD videos at 25 frames per second.
- The camera was mounted on a DJI Phantom IV UAV and pointed towards the ground.
- Total video duration is 4 minutes and 20 seconds.



Artificial Intelligence & Information Analysis Lab



AUTHDroneSunday_ VID

- A dataset for visual human crowd detection was collected, in the form of 6 videos shot inside the AUTH Campus using a DJI Phantom IV UAV.
- UHD 20160p, 25 frames per second.
- Total video duration is 4 minutes and 20 seconds.



Artificial Intelligence & Information Analysis Lab



UAV_Detection

- A dataset was prepared by AUTH for visual drone detection.
- 12 Full HD videos filmed using two cameras.
- The drone is shot against various backgrounds
- Total video duration is 31 minutes.







• UAV_Detection_2

- A dataset for drone detection was collected using one camera held by a person on the ground, within AUTH campus.
- 11 Full HD videos were produced, which contain shots of a DJI Phantom IV
- Total duration of this dataset is 15 minutes, or about 22K frames at 25fps.

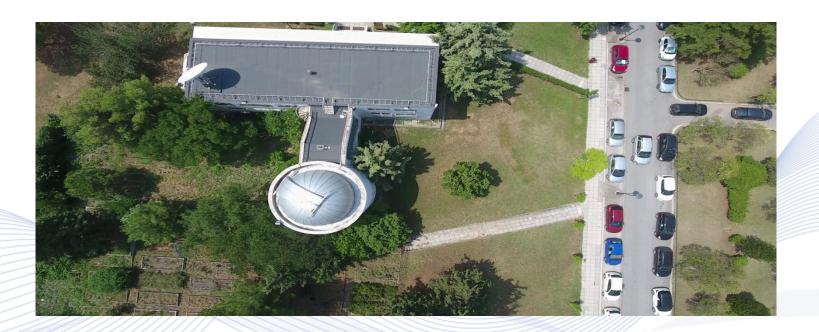


Artificial Intelligence & Information Analysis Lab



• Landing_sites

- A dataset of videos depicting potential UAV landing sites has also been captured.
- It consists of 2 videos captured by a DJI Phantom IV within AUTH campus, containing potential landing sites around a point of interest (POI), or generally in the university campus.







AUTHObservatory_VI D

- A dataset named "AUTHObservatory_VID" was also collected by AUTH for building/Point-of-Interest detection purposes.
- It consists in two videos, containing the building of the AUTH observatory with the telescope dome.
- UHD 2160p, 25 frames per second.







face_deid_UAV

- A dataset for face deidentification consists of one 3840 × 2160 video, which was shot by flying a DJI Phantom IV.
- The total video duration is 45 seconds with a framerate of 25 fps.
- Each face in the 1124 extracted frames is annotated with a bounding box.







- face_deidentification_UAV_mu lt_views
- A multiview dataset for face de-identification purposes was collected by a DJI Phantom IV UAV and consists of one 4096 × 2160 video.
- The subjects were recorded from multiple viewpoints while walking-by.
- Total duration: 2 minutes and 23 seconds, at a framerate of 25 fps.



Artificial Intelligence & Information Analysis Lab



- Annotations_eights_D W_raw
- A dataset for boat detection/tracking was created, using footage from DW, consisting of 3 videos (resolution: 1280 x 720) subsamplbed at 25 frames per second.
- An annotation file is included along with each video file.



Artificial Intelligence & Information Analysis Lab



Aerial_Crowd_Auth

- An aerial crowd detection dataset was created by annotating videos captured by two different RGB cameras (olympus, sony) placed ~10m over the ground, recording a human crowd from different viewing angles.
- The videos were partially annotated, resulting in 563 1920x1080 RGB images along with their segmentation maps, which consist of two classes ('crowd', 'non-crowd').
- The segmentation maps are available as .png images, where pixels belonging in the 'crowd' class are in red color, while 'non-crowd' class pixels are in black
 Adjorial Intelligence & Information Analysis Lab





Multiview_Synthetic_UAV

- This dataset contains sequences
 generated by simulating 3
 camera-equiped UAVs flying
 simultaneously under specific
 UAV/camera motion types
 (CMTs) and framing shot types
 (FSTs), while filming a simulated
 bicycle race.
- Each sequence may include up to 10 cyclists, differing only in the color of their jerseys. The UAVs
 fly in a 3-UAV ORBIT setup, a 2-UAV CHASE plus 1-UAV VTS and a 3-UAV TRACK setups, according to the MULTIDRONE UAV shot type taxonomy.

The Unreal Engine 4-based AirSim simulation environment was employed for constructing the sequences.

The evaluation dataset contains more than 90000 video frames, at a resolution of 640 x 360 pixels and a framerate of 25 FPS, while each video is more than 6.5 minutes long.

Temporally synchronized groundtruth is provided for all camera parameters, 3D target positions and the corresponding 2D bounding boxes across all sequences.



Artificial Intelligence & Information Analysis Lab



• UAV_Parkour

 A UAV dataset for parkour athlete detection was assembled from 8 Youtube videos, depicting both male and female athletes performing pakour at different landscapes, under differ lighting conditions.





Final_bicycles

 A dataset for bicycle detection/tracking was created, consisting of 6 HD videos, at 50 or 25 frames per second. An annotation file is included along with each video file.





• Final_boats

 A dataset for rowing boat detection/tracking was created, consisting of 5 HD videos, at 50 or 25 frames per second. An annotation file is included along with each video file.





Final_single_boats

 A dataset for single boat detection/tracking was created, consisting of 5 HD videos, at 50 or 25 frames per second. An annotation file is included along with each video file.





UAV_BothKamp

- A UAV dataset for parkour athlete detection was created by annotating the footage acquired during MULTIDRONE experimental media production.
- It consists of 6 videos (1920 1080) at 50 frames per second.
- The annotations provided are not exhaustive, i.e., there may be unannotated objects in some video frames.





UAV_Crowd_Seville

- A UAV crowd detection dataset was created by annotating videos captured by three different operating UAV cameras.
- The videos were partially annotated, resulting in 603 1920x1080 RGB images along with their segmentation maps, which consist of two classes (crowd, non-crowd).
- The ground-truth segmentation maps are available as .png images, where pixels belonging in the 'crowd' class are in red color, while 'non-crowd' class pixels are in black color.





Multiview_Boats_Bothcamp

- This dataset depicts a sample of a rowing race with three row boats in Bothkamp, Germany (September 2019).
- The footage was captured at 50 FPS and at a resolution of 1920x1080, using two UAVs filming simultaneously from different positions and view angles.
- The footage from the two UAVs is contained in two separate RGB video files, losslessly compressed with the Lagarith codec. No annotation is currently available.





RAI/DW datasets





• IGA_2017

 The footage was taken during the International Horticultural Exhibition (IGA) in Berlin, June 2017. The drone used was a Mavic Pro, parts of the footage has been published on Deutsche Welle's Internet format 'Daily Drone' <u>https://www.youtube.com/watch?v=MBgjr3ua554</u>.





WUENSDORF_2017

 The footage showing a former Soviet base in the Federal State of Brandenburg, Germany, was shot with one Inspire 2 and one Mavic Pro (July 2017). The footage was used to create another Daily Drone clip https://www.youtube.com/watch?v=IIwQmGsXTNs. It shows the remains of the Soviet barracks and a Lenin statue.





• MUENCHEBERG_2017

 The footage was taken in Muencheberg, Brandenburg, Germany, in October 2017, using one Inspire 2 and one Mavic Pro. In total, 29 clips were produced focusing on the MULTIDRONE Camera Motion Types taxonomy. The dataset includes the clips and the associated flight records.





• MUENCHEBERG_2018

- One Inspire 2, one Mavic Pro and one Phantom 4 were used by a Deutsche Welle team to film a group of cyclists simulating a bicycle race, in Muencheberg, Brandenburg, Germany, during May 2018.
- The shoot was accompanied by colleagues from the University of Bristol who created simulations of such a bike race prior to the actual shooting.
- The parameters of these simulations such as flight altitude, camera angle, etc., were used during the recording of the race. Flight records are provided.





• NAUEN_2018

- One football player and one cyclist were filmed with one Inspire 2 and one Mavic Pro in Nauen, Germany, during April 2018.
- The shooting focussed on a subset of the UAV Camera Motion Types identified in the MULTIDRONE UAV shot type taxonomy (Lateral Tracking Shot, Vertical Tracking Shot, Pedestal/Elevator Shot With Target, Chase/Follow Shot, Orbit).
- The dataset contains 19 clips and their associated flight records.





• GIRO_2017

• This dataset consists of 9 clips taken form 2017 edition of the Giro d'Italia at 1920x1080 resolution and MP4 format at 25 frames per second.





• GIRO_2018

• This dataset consists of 26 clips taken form 2018 edition of the Giro d'Italia at 1920x1080 resolution and MP4 format at 25 frames per second.





ARCHIVE_2018

- This dataset consists in 36 clips taken from RAI archives and depicting various shots of bikers, football players, boat racers and other additional outdoor sports (ski, sailing).
- Resolution is varying from 720x576 to 19020x1080 depending on the stored copy in the archive.





• METEORA_2018

- This dataset contains UAV footage filmed for Deutsche Welle's "Euromaxx Lifestyle in Europe", in the mountains of Meteora, Greece, in August 2018.
- The footage mainly depicts rock climbing and it was shot using two drone models (a Mavic Air and an Inspire 2), as well as a variety of different shot perspectives, movements and angles.





• WANNSEE_2018

- This dataset contains UAV footage filmed by a Deutsche Welle team during the live rowing regatta "Rund um Wannsee" of 2018, one of the longest races in the world, set in the southwest of Berlin.
- Two drone teams were set along the track, a third drone was used for aerial overview and two additional standard camera teams covered the rest.





WANNSEE_2018_Test

- This dataset contains UAV footage filmed by a Deutsche Welle team before the live rowing regatta "Rund um Wannsee" of 2018.
- Three drones were employed (an Inspire, a Mavic Air and a Mavic Pro), with flight records provided.





CYCLISTS_2019

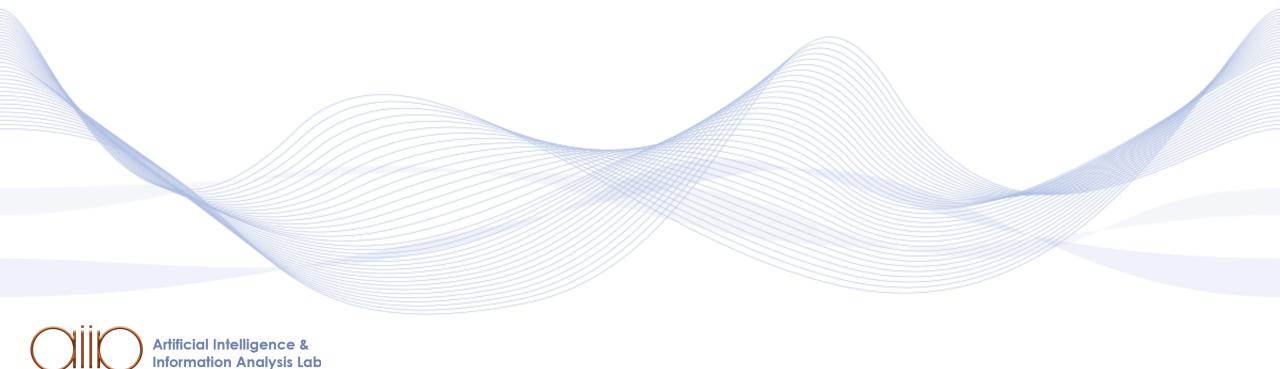
- This is a dataset depicting a bicycle race training session in northern Italy (May 2019).
- The footage was filmed by RAI, using a DJI Phantom UAV flying above the bikers.





YOUTUBE_Drone_Footage

• This is a dataset consisting in the list of links of roughly 10 hours of drone footage on YouTube on soccer, rowing and cycling.





Q & A

Thank you very much for your attention!

Contact: Prof. I. Pitas pitas@csd.auth.gr

