Project Report

Capacity building for object detection and tracking in UAV videos using deep learning

Principal Investigator : Dr. K. Vani
Address : Department of Information Science and Technology,
College of Engineering,
Anna University,
Chennai - 600 025
Email ID : vanirrk@gmail.com, vani@annauniv.edu

Co-investigator- I : Dr. S. Sanjeevi
Address : Department of Geology,
College of Engineering,
Anna University,
Chennai - 600 025
Email ID : ggjeevi@gmail.com, sanjeevigeo@gmail.com

Co-investigator- II : Dr. Chao-Hung Lin
Address : Department of Geomatics,
National Cheng-Kung University
No. 1, University Road, Tainan City 701, Taiwan
Email ID: linhung@mail.ncku.edu.tw

Junior Research Fellow : A. Ancy Micheal
Address : Department of Information Science and Technology,
College of Engineering,
Anna University,
Email ID: ncysus17@gmail.com

Summary:

The main goals of the project are:

- To aid students and researchers in exploring the possibilities of Deep Learning with open-source tool for object detection and tracking using UAV videos.
- To provide novel methodologies for object detection and tracking using deep learning methodologies with UAV videos.

The novel methodologies developed during the project period are:

- Deep learning based single object detection and tracking with UAV videos
- Semi-automatic image annotation to reduce manual labor and time
- Deep learning based multiple object detection and tracking with UAV videos

During the project period, a tool named DL-ODT-UAV(Deep Learning-based Object Detection and Tracking with UAV data) has been developed, which enables the students and researchers to widen the applicability of UAV and deep learning.

Outcome of the Project:

DL-ODT-UAV Tool : Tool for Deep Learning based Object Detection and Tracking with UAV data

This tool aims to provide a resource for students, researchers and trainers in Deep Learning methodologies with UAV data. Since deep learning has been playing a vital role in object detection and tracking during the last decade, the need to educate the students with deep learning methodologies is undeniably in-demand. The DL-ODT-UAV tool is designed for:

- Single Object Detection and Tracking
- Multiple Object Detection

Outcomes of DL-ODT-UAV Tool :

- The first screen of the tool showcases a study resource providing basic knowledge about UAV, deep learning, object detection and tracking (Figure 1). The next screen provides the steps for the user (Figure 2).

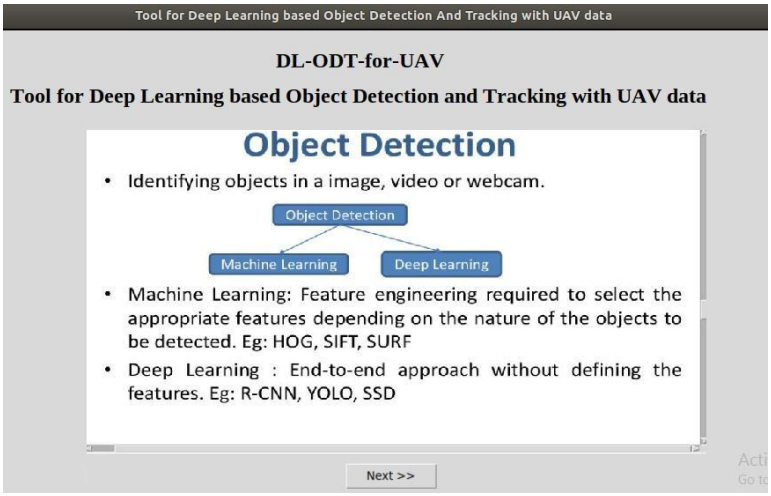


Figure 1. First Screen of DL-ODT-UAV tool

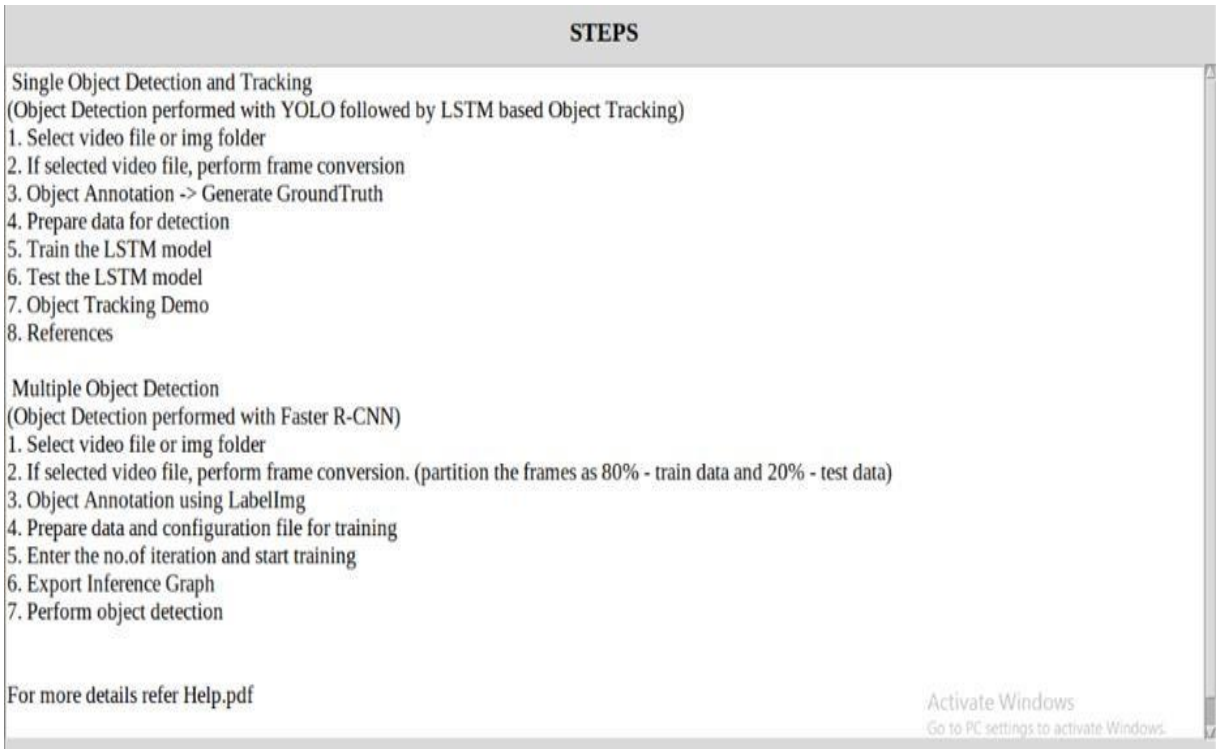


Figure 2. Manual for the User

➢ The following screen provides the type selection options for the user – Choose Single object detection and tracking/ Multiple object detection (Figure 3).

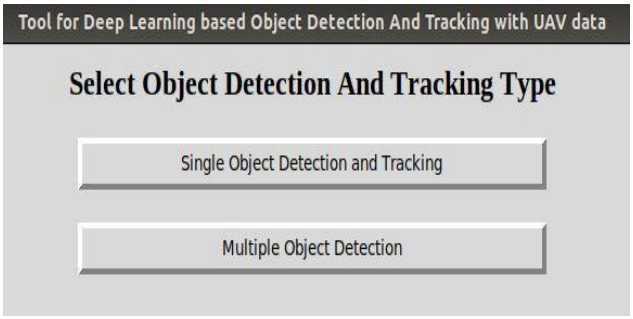


Figure 3. Type selection for the user

➢ **Type Selection 1: Single Object Detection and Tracking:** In this module, the dataset is loaded by the user. An open source BBox-Label-Tool is integrated for object annotation purpose. The annotation is restricted to single object for user convenience. Semi-automatic annotation has been implemented based on YOLO(You Look Only Once) object detector to reduce the annotation difficulty of the user. YOLO is a single neural network which predicts bounding boxes and class probabilities directly from full images in single evaluation. The generated groundtruth is fed into the YOLO object detector for object detection. The detected locations and the visual features are trained with LSTM(Long Short-Term Memory) for object tracking. LSTM is a recurrent neural network composed of a cell, an

input gate, an output gate and a forget gate. LSTM exploits the spatio-temporal information represented by the locations and visual features. The option of varying the training iteration is provided for the user. Finally, single object tracking is performed with LSTM.

The steps involved in the single object tracking module are as follows:

- **Upload Video or Images or Select Existing Dataset**: If video is given as the input, it is converted to frames for annotation. The access to the previous and the next screen are provided with <<Prev and Next>> buttons (Figure 4).

Figure 4: Loading the dataset

- **Single Object Annotation:** The user draws the bounding box on the object (Figure 5). *Generate Groundtruth* option generates the bounding box location in the text file.

Figure 5. Single object annotator

- **Object detection with YOLO:** The generated ground truth is fed into the YOLO detector for object detection (Figure 6).

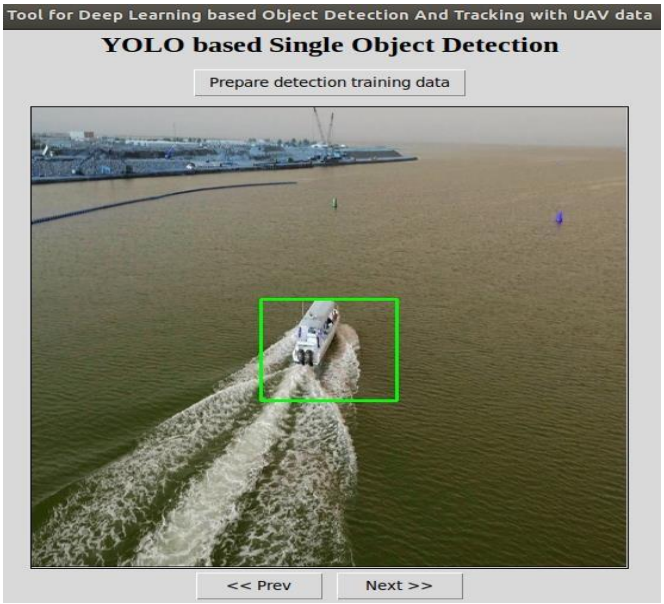


Figure 6. Single object detection

- **Object Tracking with LSTM:** This screen obtains the no. of iterations from the user and performs training and testing followed by the tracking demo (Figure 7). The trajectory of the object is generated while tracking (Figure 8).

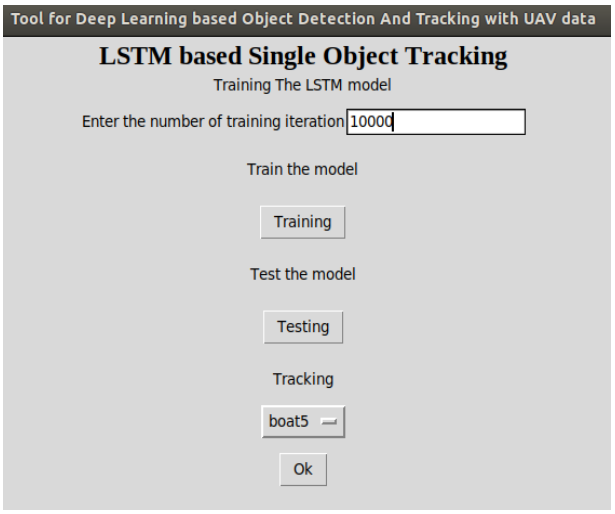


Figure 7. Training the LSTM tracker

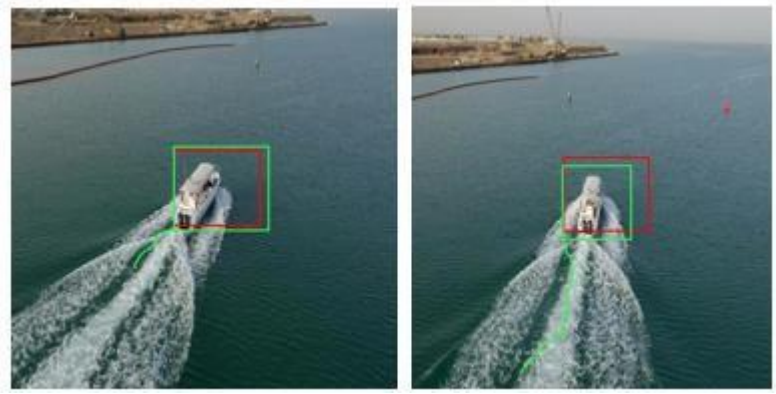


Figure 8. LSTM-based single object tracking

- **Type Selection 2 : Multiple Object Detection :** In multiple object detection, the dataset is loaded by the user. An opensource tool LabelImg is integrated to annotate multiple objects. The groundtruths are obtained and the data is prepared for training the Faster R-CNN object detector. Faster R-CNN is a detection pipeline which uses region proposal network as a region proposal algorithm and Fast R-CNN as a detector network. The training iteration is obtained from the user and multiple object detection is performed with the trained checkpoints. The steps involved in multiple object detection are as follows:

- **Upload Video or Images:** If video is uploaded, it is converted to frames. The obtained video or image folder is divided as training data and testing data (Figure 9).

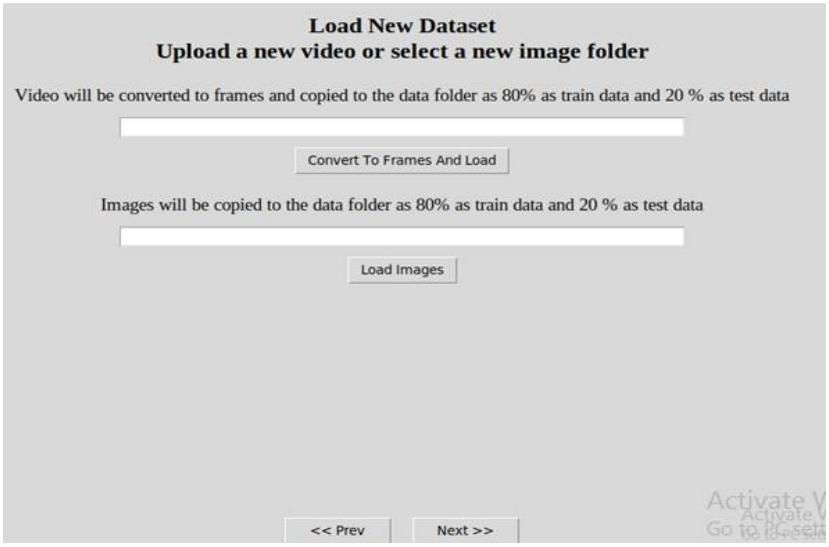


Figure 9. Loading data for multiple object detector

- **Multiple Object Annotation:** The user draws the bounding boxes on the objects (Figure 10 and 11).

Figure 10. Annotating the training and testing images

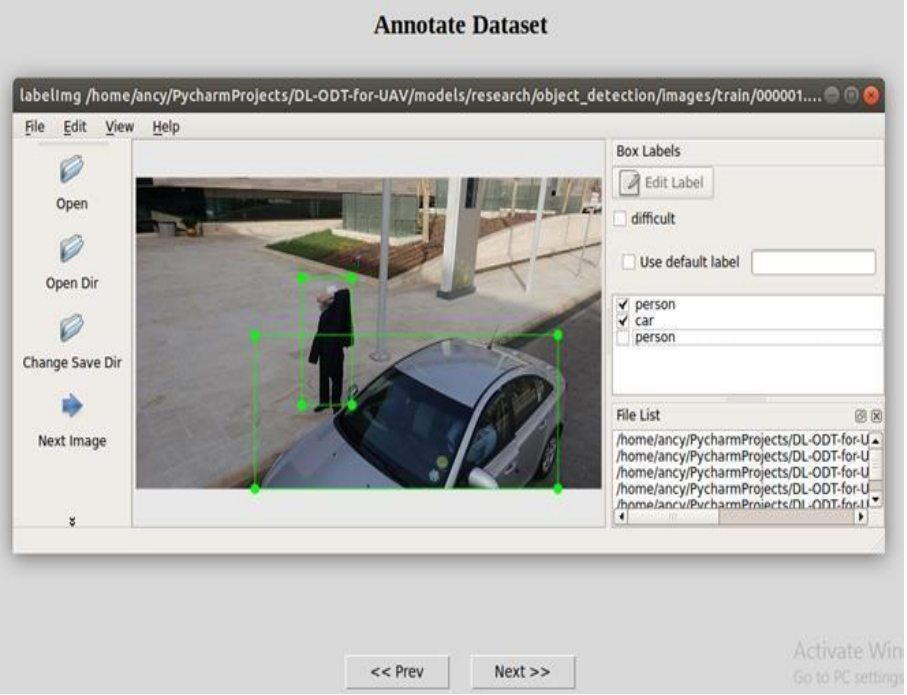
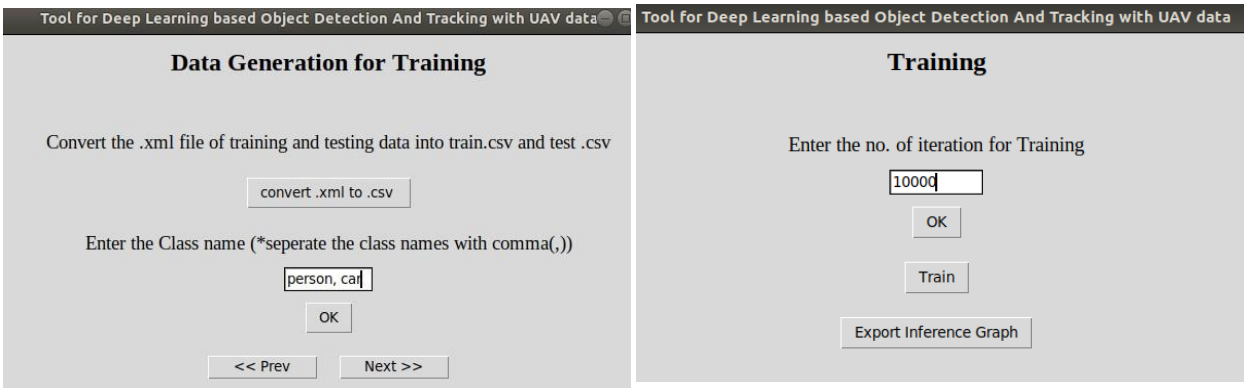


Figure 11. Multiple object annotation

- **Generate data for training:** The annotated files are converted to compatible format(.csv files). The class names are obtained from the user (Figure 12a). The number of iterations for training Faster R-CNN is obtained from the user (Figure 12b). The checkpoints are generated for every 500 iteration. After training, the inference graph is exported(.pb file).



(a) (b)

Figure 12. (a) Data generation for object detector training (b)Training Faster R-CNN

- **Multiple object detection:** With the last generated checkpoint, object detection is performed. The user provides the image for object detection. Multiple object detection is performed with Faster R-CNN as shown in Figure 13.

Figure 13. Multiple object detection

The detected objects are bounded by bounding box with a text description of the class name.

Publications and Products:

- A. Ancy Micheal., K.Vani., "Automatic object tracking in optimized UAV video", The Journal of Supercomputing, Vol.75, Issue.8, pp 4986–4999, Published August 2019 (Presented at 1st International Symposium on Artificial Intelligence and Computer Vision(26.09.2018-28.09.2018), College of Engineering, Anna University, Chennai - 600025. The paper along with 17 best papers, was forwarded to The Journal of Supercomputing).

- A paper entitled "Object Detection and Tracking with UAV Data using Deep Learning", was presented at First International Conference on Unmanned Aerial System in Geomatics, IIT Roorkee, UP, India during 06.04.2019 to 07.04.2019.

- Presented the concepts of the project and the methodology during a talk on ” Recent Trends on UAV- Geoinformatics” by the Principal Investigator Dr. K.Vani at First International Conference on Unmanned Aerial System in Geomatics, IIT Roorkee, UP, India on 07.04.2019.

- A paper entitled "Object Detection and Tracking with UAV Data using Deep Learning", has been accepted for publication in Journal of the Indian Society of Remote Sensing (JISRS).

- A paper entitled, "A tool to enhance the capacity for deep learning based object detection and tracking with UAV data", has been submitted for XXIV ISPRS Congress, International Society for Photogrammetry and Remote Sensing, 31.08.2020 – 02.09.2020, Nice, France. The paper has been accepted for publication in the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. The publication fee has been waived under ISPRS travel grant.

- A paper entitled "Semi-Automatic Image Annotation for Videos", was presented at International Conference on Integration of Advanced Technologies for Industry 4.0, KCG college of Technology, Chennai, during the period of 12th - 13th June 2020.

- A. Ancy Micheal., K. Vani., S. Sanjeevi., Chao-Hung Lin., "Semi-Automatic Image Annotation for Videos", International Journal of Advanced Science and Technology, 29(10s), pp 6872 – 6878, 2020.

- A novel methodology for multiple object tracking with UAV videos using deep learning methodologies has been implemented. Journal paper submission for this work is initiated.

Amount Sanctioned : 5000CHF (Rs. 355,397)

Acknowledgement:

We thank ISPRS for funding the project under ISPRS Scientific Initiatives 2019. We thank Dr. A. Senthil Kumar, President of ISPRS Technical Commission V, for guiding us during the project period. We thank Anna University for permitting us to execute the project.