RECENT ADVANCES IN ARTIFICIAL INTELLIGENCE AND COMPUTER VISION FOR UNMANNED AERIAL VEHICLES

Juan Carrillo, Katherine Borda

Global Spatial Technology Solutions GSTS

IGARSS 2021 - 2021 IEEE International Geoscience and Remote Sensing Symposium | 978-1-6654-0369-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/IGARSS47720.2021.9553713

ABSTRACT

In recent years we have seen an accelerated development in the technologies enabling advanced navigation and control of Unmanned Aerial Vehicles (UAVs). Such progress has been fueled by the combination of improved hardware as well as breakthrough advances in Artificial Intelligence methods to accomplish tasks that were previously though extremely difficult to automate. In this document we present a brief summary of some of the most representative methods focused on enabling advanced perception, collision avoidance, flight planning and control, as well as some key industry applications of these capabilities.

Index Terms— UAVs, Artificial Intelligence, Computer Vision, Deep Learning

1. INTRODUCTION

Completing a comprehensive review of the state of the art Artificial Intelligence methods in this field would require an extensive effort; therefore, in this brief paper our goal is rather to give a glimpse of recent ideas and why they represent a valuable contribution. Due to the complexity and variety of operation scenarios, hardware settings, industry applications and algorithms, we simplify our review in three major categories. The first presents advances in object detection and tracking, two key tasks for reliable perception. The following section highlights recent methods in the context of flight planning, control, and collision avoidance. And the third section summarizes interesting applications of Artificial Intelligence methods for a variety of industry use cases. We finish by presenting a few concluding remarks.

2. OBJECT DETECTION AND TRACKING

Computer vision methods for object detection and tracking have rapidly evolved along with the Graphic Processing Units (GPUs) used to handle efficiently massive amounts of data in parallel. Moreover, with the release of large public datasets [1] for benchmark purposes researchers can focus more on the design of advanced algorithms instead of data collection and preparation efforts. We select six papers dealing with key challenges in object detection and tracking, and highlight the key ideas behind their contributions.

2.1. Small-Object Detection in UAV-Captured Images via Multi-Branch Parallel Feature Pyramid Networks [2]

Introduces a Deep Learning model titled *Multi-branch Parallel Feature Pyramid Networks (MPFPN)* with the goal of improving feature extraction and object detection performance for objects comprising only a few pixels in the image. Also uses a spatial attention mechanism to mitigate the effects of background and irrelevant areas in the input image.

2.2. Offloading Optimization in Edge Computing for Deep Learning Enabled Target Tracking by Internet-of-UAVs [3]

Proposes a hybrid architecture for processing images on board UAVs. The first layers of a feature extraction model run onboard and the remaining part of the backbone and high level features are transmitted to a near by edge computing node.

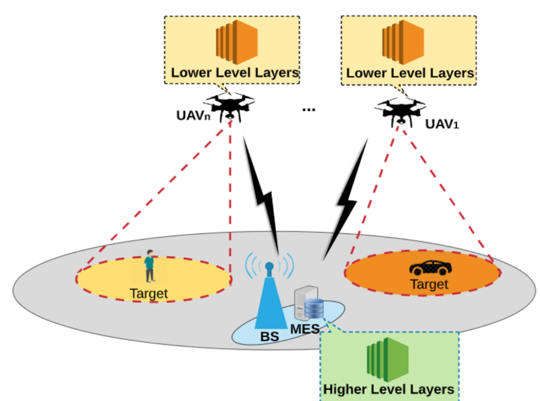

Fig. 1. Offloading Optimization in Edge Computing [3]

2.3. Computer Vision for Small UAS Onboard Pedestrian Detection [4]

Presents a single stage Deep Model inspired by RetinaNet to detect pedestrians appearing as very small objects in UAV images. Also performs tracking and highlights the challenges involved in processing images when both the sensor and objects are moving simultaneously.

2.4. Counting Vehicles with Deep Learning in Onboard UAV Imagery [5]

Introduces a combination of an efficient single stage detection model based on the YOLO architecture for the purpose of counting vehicles in very cluttered scenes with many vehicles. The authors implement their approach using an NVIDIA Jetson TX2 board, showcasing the potential for onboard processing.

2.5. Delving Into Robust Object Detection From Unmanned Aerial Vehicles: A Deep Nuisance Disentanglement Approach [6]

In this paper the authors make a comprehensive analysis of the factors that introduce high variability and therefore increased difficulty in object detection from UAVs. Such factors include the flight altitude, angle of view, and weather conditions. The paper introduces a disentanglement approach to handle variability in input features and applies adversarial training to cope with domain adaptation.

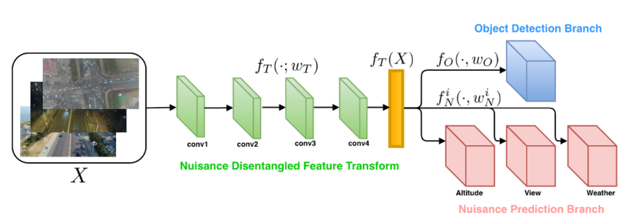

Fig. 2. Object Detection with disentanglement approach [6]

2.6. SlimYOLOv3: Narrower, Faster and Better for Real-Time UAV Applications [7]

The authors design an efficient model for real time object detection by applying a channel pruning approach that reduces the number of parameters while keeping a target performance. More specifically, they manage to reduce the number of parameters and computing operations (FLOPs) by up to 90% without a significant degradation in performance.

3. FLIGHT PLANNING, CONTROL, AND COLLISION AVOIDANCE

The recent progress in advanced flight planning and collision avoidance systems is powered by the use of Deep Learning models, more specifically Deep Convolutional Neural Networks that extract visual cues and translate raw image data into high level instructions to plan feasible flight trajectories, effectively control the UAVs, and enable collision avoidance capabilities.

3.1. Catch Me, If You Can! A Mediated Perception Approach Towards Fully Autonomous Drone Racing [8]

The authors present the design and implementation details of a system that scores among the best entrants in a simulated drones racing competition. Their system includes a perception, mapping, and planning stack powered by machine learning. The data inputs come from a monocular onboard camera within the simulation and the whole system runs efficiently in real time.

3.2. Deep Learning Based Hand Gesture Recognition and UAV Flight Controls [9]

The authors train a Deep Learning model to automatically interpret hand gestures that are then translated as commands for the control of a UAV. The paper includes details about the key controls necessary to maneuver a UAV. Moreover, they evaluate the performance of their system in a drone simulator and a physical drone.

3.3. An Open Source and Open Hardware Deep Learning-Powered Visual Navigation Engine for Autonomous Nano-UAVs [10]

The authors design a nano UAV capable of automated visual navigation and collision avoidance, with all perception, planning, and control subsystems running on board the vehicle with a power budget of less than 10 Watts and weighting less than 27 grams. Among their key contributions are an efficient deployment of highly optimized Deep Learning methods and the use of customized processing chips.

Fig. 3. Visual Navigation Engine for Autonomous Nano-UAVs [10].

3.4. UAV environmental perception and autonomous obstacle avoidance: A deep learning and depth camera combined solution[11]

This paper makes use of a Deep Network to process images coming from a RGB-Depth camera to enable an effective collision avoidance system onboard UAVs for low altitude flight in farms. Furthermore, the proposed model classifies nearby objects depending on their size and proximity, the results are used to inform an integrated collision avoidance system.

4. INDUSTRY APPLICATIONS

Multiple case studies benefit from automated analysis of data on board UAVs. From risk management to urban planning and agriculture, just to name a few, these use cases make use of UAVs for rapid data collection and decision making. When on board systems ingest and process data automatically, the downstream stages that make use of such data can operate faster and deliver actionable insights rapidly.

4.1. A YOLOv3-based Learning Strategy for Real-time UAV-based Forest Fire Detection[12]

The authors implement a Deep Learning model based on the YOLOv3 architecture to detect forest fires in real time. Their deployment is capable of running at 30 frames per second, offering an interesting potential for real time operations. Moreover, their system works over images collected at multiple conditions of flying altitude and angles of observation.

Fig. 4. Real-time UAV-based Forest Fire Detection [12]

4.2. Deep learning-based multi-feature semantic segmentation in building extraction from images of UAV photogrammetry[13]

This paper presents a comprehensive description of a Deep Learning model to produce segmentation masks of buildings on images collected by UAVs. The key contributions include a data fusion approach to incorporate a digital elevation model as well as infrared image channels. They also present extensive experimental evaluation over diverse scenes.

4.3. Applying Deep Learning to Automate UAV-Based Detection of Scatterable Landmines[14]

The authors introduce an innovative application of UAVs and automated image processing to detect landmines in Afghanistan. Their system uses images collected by multispectral and infrared sensors to maximise the discrimination capability. Extensive evaluation is performed across a variety of terrains and the model yields competitive performance even under cluttered backgrounds.

Fig. 5. UAV-Based Detection of Scatterable Landmines[14]

4.4. Towards Real-Time Building Damage Mapping with Low-Cost UAV Solutions[15]

This paper introduces a real time system to complete automated assessments of buildings and other types of infrastructure after damage originated by undesired circumstances. Their system incorporated techniques from aerial photogrammetry and geographic information systems to provide an accurate location of the studied areas.

5. CONCLUSIONS

We review a representative sample of papers introducing state of the art methods for image processing, automated navigation, flight planning, and control of UAVs. Most of these studies make use of Deep Learning techniques. In addition, we also cover some publications presenting advances in specific applications such as risk management and mapping.

6. REFERENCES

[1] Pengfei Zhu, Longyin Wen, Dawei Du, Xiao Bian, Qinghua Hu, and Haibin Ling, "Vision meets drones: Past, present and future," *arXiv preprint arXiv:2001.06303*, 2020.

[2] Y. Liu, F. Yang, and P. Hu, "Small-object detection in uav-captured images via multi-branch parallel feature pyramid networks," *IEEE Access*, vol. 8, pp. 145740–145750, 2020.

[3] B. Yang, X. Cao, C. Yuen, and L. Qian, "Offloading optimization in edge computing for deep learning enabled target tracking by internet-of-uavs," *IEEE Internet of Things Journal*, pp. 1–1, 2020.

[4] Xuxi Yang, Michael Murphy, Marc W Brittain, and Peng Wei, "Computer vision for small uas onboard pedestrian detection," in *AIAA AVIATION 2020 FORUM*, 2020, p. 3270.

[5] G. Amato, L. Ciampi, F. Falchi, and C. Gennaro, "Counting vehicles with deep learning in onboard uav imagery," in *2019 IEEE Symposium on Computers and Communications (ISCC)*, 2019, pp. 1–6.

[6] Zhenyu Wu, Karthik Suresh, Priya Narayanan, Hongyu Xu, Heesung Kwon, and Zhangyang Wang, "Delving into robust object detection from unmanned aerial vehicles: A deep nuisance disentanglement approach," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1201–1210.

[7] Pengyi Zhang, Yunxin Zhong, and Xiaoqiong Li, "Slimyolov3: Narrower, faster and better for real-time uav applications," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019, pp. 0–0.

[8] Florian Ölsner and Stefan Milz, "Catch me, if you can! a mediated perception approach towards fully autonomous drone racing," in *NeurIPS 2019 Competition and Demonstration Track*. PMLR, 2020, pp. 90–99.

[9] Bin Hu and Jiacun Wang, "Deep learning based hand gesture recognition and uav flight controls," *International Journal of Automation and Computing*, vol. 17, no. 1, pp. 17–29, 2020.

[10] D. Palossi, F. Conti, and L. Benini, "An open source and open hardware deep learning-powered visual navigation engine for autonomous nano-uavs," in *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, 2019, pp. 604–611.

[11] Dashuai Wang, Wei Li, Xiaoguang Liu, Nan Li, and Chunlong Zhang, "Uav environmental perception and autonomous obstacle avoidance: A deep learning and depth camera combined solution," *Computers and Electronics in Agriculture*, vol. 175, pp. 105523, 2020.

[12] Z. Jiao, Y. Zhang, L. Mu, J. Xin, S. Jiao, H. Liu, and D. Liu, "A yolov3-based learning strategy for real-time uav-based forest fire detection," in *2020 Chinese Control And Decision Conference (CCDC)*, 2020, pp. 4963–4967.

[13] Wuttichai Boonpook, Yumin Tan, and Bo Xu, "Deep learning-based multi-feature semantic segmentation in building extraction from images of uav photogrammetry," *International Journal of Remote Sensing*, vol. 42, no. 1, pp. 1–19, 2021.

[14] Jasper Baur, Gabriel Steinberg, Alex Nikulin, Kenneth Chiu, and Timothy S de Smet, "Applying deep learning to automate uav-based detection of scatterable landmines," *Remote Sensing*, vol. 12, no. 5, pp. 859, 2020.

[15] Francesco Nex, Diogo Duarte, Anne Steenbeek, and Norman Kerle, "Towards real-time building damage mapping with low-cost uav solutions," *Remote sensing*, vol. 11, no. 3, pp. 287, 2019.