A Brief Survey of UAV Applications in Video Surveillance

Naqqash Dilshad, JaeYoung Hwang, JaeSeung Song*
Department of Information Security
Sejong University
Seoul, South Korea
{dilshadnaqqash, forest62590}@sju.ac.kr, jssong@sejong.ac.kr

NakMyoung Sung
Autonomous IoT Research Center
Korea Electronics Technology Institute
Gyeonggi-do, South Korea
nmsung@keti.re.kr

***Abstract*—As we are moving rapidly towards smart urbanization, more and more IoT devices such as vision sensors are deployed, yielding the need of continuous surveillance via automated techniques. A handsome amount of research is going on to provide secure surveillance using CCTV cameras, but these cameras pose limited coverage, lack location sharing and tracking abilities. In contrast, vision sensors installed on drones are flexible and scalable with broader surveillance coverage. But at the same time vision sensors' data from drones pose several challenges such as limited resources, shuddering camera affects, signals' manipulation and disturbance. The drones' surveillance lacks the researchers' attention and therefore, we accumulate the related papers in the literature review and discuss their functional approaches comprehensively. The articles' focused topics in terms of surveillance using drones include object detection, tracking and activity recognition. This article throws light on the research challenges and deep insight of the methods used in the cited articles and also proposes a new architecture for Flex-drone i.e., context-aware intelligent Unmanned Aerial Vehicle (UAV) in smart city environment.**

***Index Terms*—UAV, 5G-IoT, AI, Video Surveillance**

I. Introduction

SECURITY is of great concern in this modern era of computing and scientific developments. Earlier it was a human intensive task but due to recent technological advancements it has became autonomous. Utilizing drones for video surveillance with latest computer vision and machine learning techniques can significantly reduce security risk in critical scenarios by overcoming limited coverage and easing access for difficult to reach areas. For security personnel it is always hard to cover the whole area by patrolling in a metropolitan. With a bird's eye view, the coverage area increases by up to 200 percent and also enhancing the alert system by generating signals for violent activities [1]. The future of UAV applications in daily life is near, ranging from delivering products door to door, prevention of crime by surveillance, improving journalism in crowded areas, employing rescue operation in hard to reach areas. In short UAV application possibilities are limitless and it will grow with the technological advancements.

Although, in object detection and tracking research is in play since decades, but still lots of challenges are there to be addressed properly. Adding more, the inter and intra human movement classes inconsistency is a known and major flaw in the area of human activity recognition. Challenges such as image noise, complex object movement, inconsistency in illumination, blockage of objects, their complex structures, dynamic point of view as well as variation of complex background and elevation of camera and human physical appearance still exists in object detection related literature [2]. In addition to the above challenges, object tracking with UAV poses other complications. As the video is captured from a substantial altitude, the target objects (pedestrians and automobiles) are comparably smaller due to which most of the current algorithms and paradigms are ineffective.

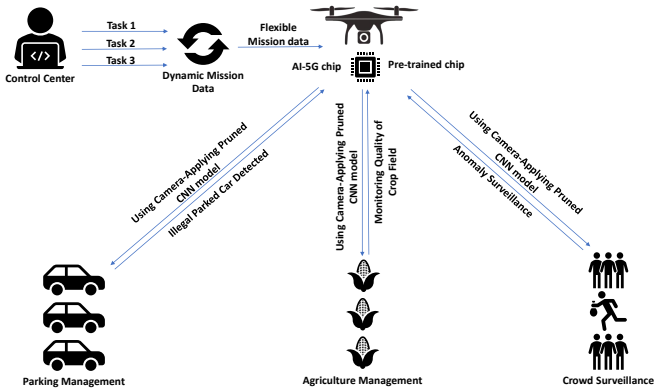

Fig. 1. High Level Architecture for Context Aware Intelligent UAV

Related literature suggests that very less work has been done in the field of video surveillance using drones, so a brief survey will be an eye-catching article for readers which can pave a path for research by embedding AI engine in UAV and also utilizing 5G chipset for fast and reliable communication. This intelligent silicon must be resource friendly i.e., easy on the battery while processing more instructions per transistor per unit time. In Fig. 1 different use cases are described using AI-5G technology for UAV. Scenario 1 indicates car parking management, scenario 2 shows agriculture management while

in 3 crowd surveillance has been presented. Moreover, the paper is divided into the following sections, Section 2 briefly introduces the concept of drone video surveillance in object detection and activity recognition. The last section provides conclusion and offers a viewpoint of future research direction.

II. Literature Review

A. Object Detection

Object detection is one of the most focused and challenging issue in the field of computer vision applications due to complications in real time tracking of objects. Which can be the result of various extrinsic and intrinsic aspects like camera motion, deformation, occlusion and motion blur. Object detection has many applications in different areas like automatic detection of vehicle number plates, people counting to generate crowd statistics, facial recognition at security checkpoints. Similarly, it has also been utilized by using UAV for detection/ tracking purpose, reconnaissance missions, search-and-rescue operations and security surveillance. UAV is more cost efficient and accurate in terms of photographic details than satellites because a drone can capture images with extremely high level of spatial features [3]. Earlier statistical based models were only used for object detection and object recognition but later on deep learning (DL) based models were introduced because statistical models are weak and their results are solely based on probability, due to which DL models has an upper hand over statistical methods.

Multiple object tracking includes approximating the trajectory of various objects simultaneously over a period of time in a sequence of frames. This can be achieved by online and offline method. Online method is useful for real-time object tracking due to which previous and current frames are accessible for the tracker, which needs substantial computational power for more complex algorithms. In such situation tracking by detection method has been proposed comprising of YOLOv3 and Retina-Net. Which uses an object detection technique to start, update or terminate a tracker. SORT is an example of such technique, it is based on Kalman-Filter and Hungarian algorithm to handle motion prediction and data association. On the other hand, it skips the salient features in the association matrix which is in turn overcame by integrating object appearance data in an association matrix. Although the online method is speedy but lacks certain level of accuracy which can be overcame by training the tracker in offline mode for optimization purpose of its parameters [4].

B. Activity Recognition

Human behavior analysis is a hot topic of discussion among many other topics of interest. This includes number of topics ranging from detection, tracking and analyzing human behavior. Going further, human activity recognition is considered as an important research paradigm in robotics, human-computer interaction, sports analysis and motion of players, video games for player characters. Specifically, human activity recognition by UAV has got the attention of many researchers, due to its advantages to overcome the problem of static coverage and the ability to access difficult areas. Allowing the drone to cover the issue of occlusion which ensures the surveillance of open places (e.g., festivals, open air concerts, forests).

Due to numerous issues in this field like dynamic and complex background along with variable point of view from a high altitude is a thorny problem. To address this, new approach has been proposed which works in two phases i.e., an offline phase which generates human/non-human activity models using convolutional neural network and an inference phase, which enables the already generated model to perform human detection followed by activity recognition. Furthermore, the proposed model doesn't require an annotated dataset and human activity classification has been performed in two scenarios i.e., an instant video frame classification and an entire video sequence classification. For video surveillance systems, instant video classification is more suitable so that suspicious activity can be reported in due time, while on the other hand entire video sequence classification is based on assigning a single activity label to the video sequence [5].

III. Conclusion

In this article, we presented a compact and brief literature about video surveillance using UAV, specifically for object detection and activity recognition. The challenges discussed in object detection include real-time tracking of objects which is always difficult due to many distortion factors. Similarly, in activity recognition, wider field of view, mobile objects and motion blur are some of the major issues discussed in activity recognition using UAV. This brief survey can be used for the ongoing UAV, AI-5G, IoT related projects. The scope of the research can be context-aware learning to human like scene understanding by utilizing a smart chipset in urban environment.

Acknowledgment

This work was supported by Institute for Information communications Technology Promotion (IITP) grant funded by the Korea government (MSIT) (No.2020-0-00959, Fast Intelligence Analysis HW/SW Engine Exploiting IoT Platform for Boosting On-device AI in 5G Environment). Corresponding author: Prof. Song JaeSeung.

References

[1] F. U. M. Ullah, A. Ullah, K. Muhammad, I. U. Haq, and S. W. Baik, "Violence detection using spatiotemporal features with 3d convolutional neural network," *Sensors*, vol. 19, no. 11, p. 2472, 2019.

[2] L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikäinen, "Deep learning for generic object detection: A survey," *International journal of computer vision*, vol. 128, no. 2, pp. 261–318, 2020.

[3] Y. Bazi and F. Melgani, "Convolutional svm networks for object detection in uav imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 6, pp. 3107–3118, 2018.

[4] S. Kapania, D. Saini, S. Goyal, N. Thakur, R. Jain, and P. Nagrath, "Multi object tracking with uavs using deep sort and yolov3 retinanet detection framework," in *Proceedings of the 1st ACM Workshop on Autonomous and Intelligent Mobile Systems*, 2020, pp. 1–6.

[5] H. Mliki, F. Bouhlel, and M. Hammami, "Human activity recognition from uav-captured video sequences," *Pattern Recognition*, vol. 100, p. 107140, 2020.