Toward AI-Assisted UAV for Human Detection in Search and Rescue Missions

Mostafa Rizk[§†‡], Fatima Slim[§], Jamal Charara[§]
[§] Lebanese University, Department of Physics and Electronics, Lebanon
[†] Lebanese International University, Department of Computer and Communication Engineering, Lebanon
[‡] IMT-Atlantique, CNRS Lab-STICC, UBL, F-29238 Brest, France

***Abstract*—Search and rescue missions during and after disasters require all efforts and high financial expenses. Rapid locating of wounded and lost individuals contributes in directing the rescuers and medical teams. This may increase the probability of saving human lives and plays a significant role in reducing expenses. Currently, the use of unmanned aerial vehicles (UAVs) or drones for remote surveillance and reconnaissance is becoming increasingly popular. On the other hand, emergent artificial intelligence (AI) algorithms based on convolution neural networks (CNN) reveal the ability of real-time detection. Combining the high-performance detection and classification capabilities provided by emergent AI techniques with the exploratory abilities of UAVs allows the UAVs to process the captured sequence of images and report back results in real-time. The evolution of AI-assisted UAVs enables the detection of wounded and trapped persons while flying and allows proper and fast transfer of information to ground stations to lead the rescuers and medical teams to victims' locations. In this paper, we explore augmenting UAVs with processing units executing emergent AI-based detectors. The proposed system can detect humans in real time and send the corresponding coordinates to the ground station.**

***Index Terms*—UAV, Human detection, Search and rescue missions, artificial intelligence, YOLO**

I. Introduction

Nowadays, government authorities and non-governmental organizations seek to increase their capabilities to rescue wounded individuals during crises. One of the most dangerous threats to an injured person is the delay in response of the medical staff. The time since injury is important because it influences medical management. The longer the delay between injury and medical action, the greater risk of complications [1]. In general, the delay in providing proper medical treatment refers to the lack of information about the position of the injured individuals and their current medical situation. During rescue and search missions, significant efforts, financial expenses, and time are invested to save human lives. Despite the fact that the UN-affiliated International Search and Rescue Advisory Group (INSARAG) has set various standards on rescue and saving operations [2], records show that hundreds of first responders are involved at random to save a single human life without any direction or guidance. This situation influences the efficiency and efficacy of the operations and may have bad effects on the wounded, trapped or lost individuals.

One promising method to ensure successful and efficient search and rescue missions is the use of geospatial imaging. In this context, unmanned aerial vehicles (UAVs) have been proposed to provide geospatial imaging. A UAV is an aircraft that is controlled remotely without the need of an on-board human pilot. Also, current UAVs can navigate in fully or partially autonomous modes [3]. THe main use of UAVs have been focused in military applications. Nowadays, UAVs are being used in commercial, scientific, recreational, security, agricultural, power line inspection and other civilian applications [4] [5] [6]. This rapid expanding to civilian fields other than surveillance and aerial photography refers to the ability of UAVs to carry multitude equipment and to navigate for long time periods. The employment of UAVs equipped with electro-optical sensors, real-time processing modules and advanced communication systems demonstrates a novel low-cost solution to enhance the current capabilities of government authorities and rescue agencies in detecting and locating wounded and lost individuals during and after disasters. The use of UAVs in search and rescue missions has a great impact in decreasing efforts, financial expenses and time invested to save human lives.

In computer vision, several techniques based on feature extraction. Histogram of oriented gradients (HOG) [7] and scale-invariant feature transform (SIFT) [8]are examples of the techniques that have been widely used for object detection. Currently, artificial intelligence (AI) is the key-trend for object detection. Neural networks (NN) have been exploited to implement computer vision algorithms using machine learning concept. Several algorithms based on NN have been proposed targeting object detection where objects are classified and then localized within the image by drawing a bounding box around the object of interest. Recently, convolutional neural networks (CNN) have been deployed to classify images. Image classification using CNN results in assigning a label to an input image from a fixed set of predefined classes.

Mainly, an image is labeled according to the most dominating object in the image. However, images could include multiple objects. Hence, one label is not sufficient. In addition, detection imposes localizing the objects within the image. A spatial region (coordinates and size) of each object is required. Several techniques have been proposed to make use of the CNN in object detection. These techniques offer detecting of multiple objects in images while generating the information of their bounding boxes. Object detection methods are classified into two categories: (1) one stage models and (2) two stage

This work was supported by a project funded by the Lebanese University.

models. In two-stage models, region proposals are determined in the first stage and image classification in the second stage. In [9], sliding window approach has been utilized in region proposal stage. Later, classifiers are applied on windows at different locations with different scales. In [10], region-convolutional neural network (R-CNN) has been introduced to detect objects in images. R-CNN method generates about 2000 region proposals in the image using selective search method [11]. This technique reduces the number of regions needed when adopting sliding window methods. However, after running classification in all regions, post-processing is needed to refine the bounding boxes, eliminate duplicates and adjust the detection scores. These operations increase the complexity of the method. R-CNN suffers from space and time cost during training. Also, a limitation of slow detection is imposed. Other enhanced versions of R-CNN have been introduced such as fast R-CNN and faster R-CNN [12] which form a leap towards real-time detection.

In [13], You Only Look Once (YOLO) has been introduced. Instead of exploiting available classifiers to detect objects in various locations in an image, object detection is treated as a single regression issue in the YOLO technique, which goes straight from image pixels to bounding box coordinates and class probabilities [13]. It does not iterate the classification in different regions of an image. In deed, YOLO forms a unified model of all phases on a neural network. The input image including the object(s) to be detected, passes forward through a single neural network of multiple layers. The method directly produces predictive vectors corresponding to each object occurring in the input image including the corresponding coordinates of its bounding box and class probabilities. Since it has been released [13], several versions of YOLO have been introduced. In [14], YOLOv2 has been proposed an improved model of YOLO by replacing the backbone network with a simpler Darknet-19 and by eliminating the fully connected layers at the end. In [15], the authors have introduced YOLOv3, which is another improved detector based on YOLO. The feature extractor in YOLOv3 adopts a hybrid architecture based on Darknet-53 and Residual networks (ResNet). Also, YOLOv3 makes detections at 3 different scales to enable detection of small objects. YOLOv3 is a fully convolutional neural network made up of 1×1 and 3×3 convolution filters. It maintains a fast processing speed while providing remarkable accuracy for objects of varying sizes. Fig. 1 shows the network architecture of YOLOv3 [16]. YOLOv3 has been verified in different works targeting various applications. The obtained results show that YOLOv3 insures real- time detection while achieving remarkable accuracy and precision.

In our work, we experimented YOLOv3 to detect humans in aerial images. The obtained results show its relevance in detecting human bodies in real-time while providing acceptable accuracy. Accordingly, we propose an airborne system that alerts the rescue teams. Once a human body is detected, the system sends a notification message to the ground station. The message includes the coordinates of the location and the number of detected humans. Fig. 2 depicts the overview of the

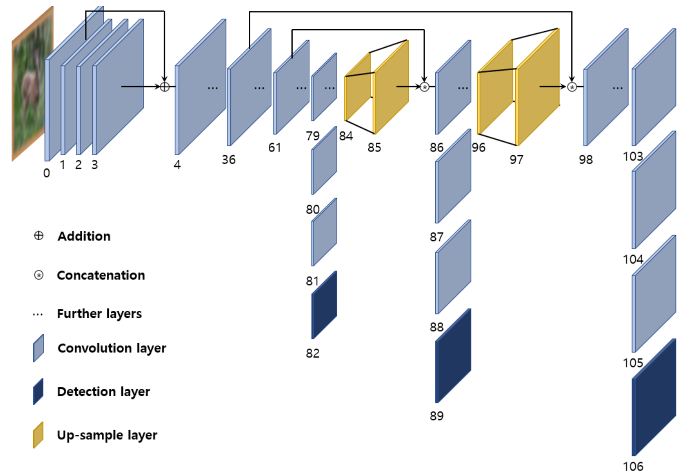

Fig. 1. The network architecture of YOLOv3 [16]

Fig. 2. Overview of the proposed system

proposed system. The remainder part of this paper is organized as follows. Next section states the motivations about using UAVs in search and rescue missions. Section II provides a brief survey about related research works. Section IV shows the proposed method, the conducted experiments and the obtained results. Finally, the conclusion is presented in Section.

II. Motivations

Search and rescue missions depending on human rescuers may not be sufficient in many cases due to several reasons. First, the wounded persons are mainly not precisely located. Also, the number of these persons is unknown for the rescue team as well as their current medical situation and required medical tools for treatment. On the other hand, covering a huge affected area requires additional time for human rescuers especially for regions with poor weather and uneven surfaces. Also, the environment of the affected area forms a serious threat for the rescuers such as the presence of toxic gases, fires, radiation, high temperature, biohazards, etc. This endangers the ground rescuers' lives and causes mental and physical fatigue [17].

Compared to other vehicles, UAVs are characterized by their cost-effectiveness, ease to use and maintain, affordability and availability in the markets in addition to their ease and fast accessibility to the disaster regions [18]. The cost of acquiring manned aerial vehicles is prohibitively high. Manned aerial vehicles operation necessitates high-trained personnel and their use requires costly preventive and corrective maintenance in order to ensure the safety of the aerial vehicle and the pilot. Furthermore, very severe restrictions are required for take-off and landing areas, which are frequently located far from the search and rescue area. Ground vehicles have been proposed to be used in search and rescue missions [19] [20] [21]. Powering the ground vehicles adds an overhead in terms of the size, weight and cost of batteries. The usage of ground vehicles is related to the characteristics of the region as the inclination and land type affect the movement of ground robots, which makes ground robots not guaranteed to move outside urban areas. Also, ground vehicles suffers from limited mobility and face difficulties to overcoming obstacles and barriers. This makes ground robots not guaranteed to move outside urban areas. Also, it requires hard real-time controlling by the operator as a small delay in transmission of control commands causes the robot to drift, which may lead to damaging the robot or its surrounding. In addition, the human operator should be provided with real-time video for the robot surrounding environment, which forms a bottleneck in terms of transmission power and bandwidth.

In addition, the use of UAVs has been proved to be efficient for medical missions. Many research works have addressed using UAVs to deliver medication to rural regions and transport medical products across long distances, including blood derivatives and pharmaceuticals human blood samples [18] [22], microbiological specimens and biological samples [23] [24] and defibrillators to out-of-hospital heart attack victims [25]. The use of UAVs equipped with electro-optical sensors and processing modules could help to overcome the current vulnerabilities of traditional man-dependent methods and that of manned aerial vehicle or ground robots. In fact, it could play an advanced role in providing real-time, data-on demand and reliable information, which pace of collection and processing is critical in rescue and search missions.

On the other hand, human detection using AI techniques has been addressed in several works. In [26], the authors have proposed a new human detection approach based on Fourier and Histogram of Orient Gradient (HOG) descriptors using CUDA. The support vector machine classifier (SVM) has been used to integrate both types of features descriptors to achieve best performance. In [27], the authors have proposed the use of CNN-based algorithm to detect and localize human position inside surveillance video footage. The experimental results show high accuracy level at 76.4%. The processing time for one image is 0.145 seconds. Thus, it will take at least 2 seconds of processing time for 15 fps to detect humans on a second video footage. The authors in [28] have proposed an combination between: Intensity histogram, Local Binary Pattern (LBP), and HOG to detect human and animal. The suggested method in [28] develops a rapid deep learning classification system using a deep convolutional neural network (DCNN).

III. RELATED WORK

Many research works have been conducted to prove the feasibility of using UAVs equipped with sensing techniques in rescue and search operations. In [29], a self-adaptive image-matching system has been presented for processing in real-time UAV videos for fast natural disaster response. In [30], the authors have proposed a method for gathering information of damaged area after a natural disaster. The method compromises real-time capturing of images, labeling their position and altitude, and sending the image with the flight parameters to the ground control station (GCS) to construct a three-dimensional hazard map. In [31], a brief analysis of the wilderness search and rescue problem with emphasis on practical aspects of visual-based aerial search has been presented. As part of this analysis, the authors have presented and analyzed a generalized contour search algorithm. In [32], the authors have presented a sensors-suite that integrates video and still visible spectrum cameras deployed on helicopter UAV. Human body detection adopting positioning algorithms through visible and infrared imagery has been presented in [33]. In addition, vision-based human detection algorithm from UAV imagery has been explored in [34]. In [35], the task of automatically finding people lying on the ground in images taken from the on-board camera of a UAV has been addressed. The authors have evaluated various state-of-the-art visual people detection methods in the context of vision based victim detection from an UAV. The well-known cascaded Haar classifiers have been explored in [36] for real-time detection on UAV imagery for human detection. In this work human detection has been proven using a combination of thermal and color images. Human body detection and geo-localization using color and thermal images for UAV search and rescue missions have been addressed. In [37], a quadcopter equipped with light-weight microwave detector has been introduced to detect living humans trapped under rubble. In [38], multiple airborne vehicles and ground platforms h ave been used to complete the EU-ICARUS project, which has been hailed as one of the most reliable geospatial data delivery systems for rescue operations. In [39], the authors have proposed a camera-based position-detection system for SAR missions. A fully autonomous fixed-wing UAV has been built using an all-in-one camera-based target recognition and positioning system. In [40], the authors have demonstrated using UAVs equipped with vision cameras to assist in avalanche search and rescue operations. Machine learning techniques have been used in the procedure. To extract discriminative features, the photos of the avalanche debris collected by the UAV have been first processed with a pre-trained CNN. In [41], a human body detection algorithm based on color and depth data captured from on-board sensors has been presented. Also, the authors have proposed a computational model for tracking numerous persons with invariance to scale and rotation of the point of view with respect to the target. In [42], a fully autonomous rescue UAV with on-board real-time human detection have

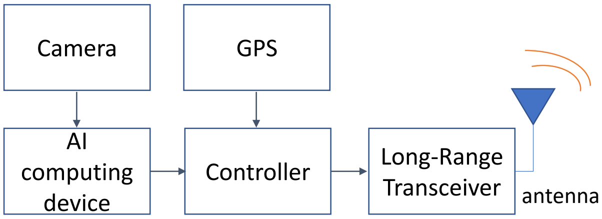

Fig. 3. The block diagram of the proposed system

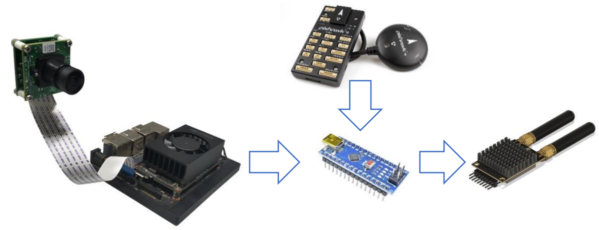

Fig. 4. System architecture

been proposed. The authors have relied on deep learning techniques to achieve detection of open water swimmers. In [43], the authors have introduced an autonomous human-following quadcopter based on human detection and tracking techniques. The system utilizes the CNN-based MobileNetv2-SSDLite model in order to detect a human. The used model was originally trained on Common object in context dataset (COCO) [44]. The feature extractor layers of the model were kept and the detection layers were fine-tuned to enhance the human detection.

IV. Proposed method

The block diagram of the proposed system is shown in Fig. 3. As shown in the figure, the proposed system is made up of five primary components These embedded components and their corresponding functionality are outlined in the following subsections.

A digital camera has been utilized to capture the video scenes under the drone. In order to ensure that it is directed towards the ground a downward gimbal mount is used. The used camera is connected to the computing device through universal serial bus (USB) port. Hence, the captured video scenes are directly processed. In addition, the camera stores the recorded video on its internal memory card.

In this work, Jetson Xavier NX from Nvidia has been adopted. It is characterized by its low weight and small size which are two essential characteristics for airborne devices. Jetson Xavier NX is considered a perfect platform for high-performance AI systems like autonomous robots and smart cameras. YOLOv3 has been employed to detect humans in captured aerial images. Transfer learning has been adopted starting from a network pre-trained on COCO dataset. Then, custom training has been conducted targeting human bodies. This accelerates the training speed and ensure generalization accuracy on a related target domain while using limited labeled data.

The controller is responsible for receiving the information about the position, velocity and time (PVT) from the GPS receiver module and generate the notification message to be sent to the ground station via the long-range transceiver. The GPS receiver has been configured to generate PVT data in National Marine Electronics Association (NMEA) format [45]. The controller parses continuously the PVT data provided by the GPS receiver and stores the updated values locally. Once a human body is detected, the computing module sends a signal to the controller. The signal interrupts the controller, which enters a service routine dedicated for generating the notification packet, which includes the following flits: drone ID, date, time, latitude, longitude, and the number of detected human bodies with the corresponding coordinates of the bounding boxes and confidence ratio. Then, the controller delivers the packet to the long-range transceiver module to be sent to the ground station.

In this work, Arduino NANO has been adopted as a controller as it includes the required peripherals to communicate with the connected modules. Also, it is characterized by its small size and low power consumption. Concerning the GPS receiver module, we make use of the one associated with the drone autopilot. For the transceiver, RFD900 module has been utilized. It is characterized by low weight, small size packaging and adequate power consumption. Also, it supports transmission diversity with long line-of-sight range of 40km. Fig. 4 presents the proposed system architecture. In order to custom train YOLOv3 network, a dataset including 3500 images has been created. The images of the dataset show full human body, upper part of human body, lower part of human body or a specific part of a human body such as leg, hand, head, etc. All used images are of JPG format and are collected from internet sources. *LabelImg* has been used to label the objects in the collected images. The labeled data has been split into training set, validation set and test set. The training has been conducted using graphical processing units (GPUs) provided by Google Colaboratory (Colab). The trained model has been tested targeting 940 images showing a full human body or a part of a human body. It has been recorded that the trained model has succeeded to detect human body in 88.88% of images with an average of accuracy of 82.21%. In addition, the trained model has been tested using 3 video sequences. The videos (from internet sources) show humans in different positions in various environments (sea, forest from far distance and forest from near distance). The obtained results show the relevance of adopting YOLOv3 for detecting humans from different heights with in various environments. Fig. 5 shows samples of detection results in the 3 video sequences. In addition, the model has been tested using live video sequences captured at a height of 14 meters using a 4K digital camera characterized by 529 pixels per inch and pixel size of 1.4 μm. The video sequences show humans moving in a field between several trees. The trained model achieves 78.78% of average accuracy with zero false positive detections in 5925 cases.

Fig. 5. Samples of the obtained detection results in test videos

V. Conclusion

The major goal of the proposed work in this paper is to develop a revolutionary solution that will aid government officials in better detecting and locating wounded, lost, and trapped people during and after a crisis. Nowadays drone-based systems are the key trend for remote surveillance and reconnaissance. On the other hand, AI has a growing role in the domain of object detection. In this paper, we propose a system that employs drones equipped with advanced technological devices and embedding AI techniques to facilitate search and rescue missions. The system makes use of emergent real-time and high accuracy AI-based object detection method to detect and locate individuals. The architecture of the proposed system is demonstrated by showing the functionality of all constituting modules. The obtained detection results show that the proposed system meets with real-time requirement and provides acceptable accuracy in detecting individuals in diverse environments. Future work will focus on improving the accuracy and conducting studies concerning power consumption.

Acknowledgment

The authors would like to thank the University of Bahrain for organizing the DASA'21 conference and supporting the publication of this paper.

Fig. 6. Samples of the obtained detection results in live video sequences

References

[1] J. Hayward-Karlsson, S. Jeffery, A. Kerr, and H. Schmidt, *Hospitals for War-Wounded: A practical guide for setting up and running a surgical hospital in an area of armed conflict*. Geneva: International Committee of the Red Cross, 1998.

[2] "INSARAG web site," http://www.insarag.org/, accessed: 2021-10-20.

[3] M. Rizk, A. Mroue, M. Farran, and J. Charara, "Real-time SLAM based on image stitching for autonomous navigation of UAVs in GNSS-denied regions," in *Proc. IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, 2020, pp. 301–304.

[4] J. Everaerts, "The use of unmanned aerial vehicles UAVs for remote sensing and mapping," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, 01 2008.

[5] C. Martinez, C. Sampedro, A. Chauhan, and P. Campoy, "Towards autonomous detection and tracking of electric towers for aerial power line inspection," in *Proc. IEEE International Conference on Unmanned Aircraft Systems (ICUAS)*, 2014, pp. 284–295.

[6] M. Rizk, T. N. Al-Deen, H. Diab, A. M. Ahmad, and Z. El-Bazzal, "Proposition of online UAV-based pollutant detector and tracker for narrow-basin rivers: A case study on Litani River," in *Proc. IEEE International Conference on Computer and Applications (ICCA)*, 2018, pp. 189–195.

[7] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” *IEEE Computer society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 886–893, 2005.
[8] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.
[9] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.
[10] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014, pp. 580–587.
[11] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders, “Selective search for object recognition,” *International Journal of Computer Vision*, vol. 104, pp. 154–171, 2013.
[12] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards real-time object detection with region proposal networks,” in *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, Eds., vol. 28. Curran Associates, Inc., 2015.
[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, real-time object detection,” 2016.
[14] J. Redmon and A. Farhadi, “YOLO9000: Better, faster, stronger,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.
[15] ——, “YOLOv3: An incremental improvement,” 2018.
[16] S. Kim and H. Kim, “Zero-centered fixed-point quantization with iterative retraining for deep convolutional neural network-based object detectors,” *IEEE Access*, vol. 9, pp. 20 828–20 839, 2021.
[17] R. Murphy, S. Tadokoro, and A. Kleiner, *Disaster robotics*. Springer, 2016, pp. 1577–1604.
[18] C. A. Thiels, J. M. Aho, S. P. Zietlow, and D. H. Jenkins, “Use of unmanned aerial vehicles for medical product transport,” *Air Medical Journal*, vol. 34, no. 2, pp. 104 – 108, 2015.
[19] B. Doroodgar, Y. Liu, and G. Nejat, “A learning-based semi-autonomous controller for robotic exploration of unknown disaster scenes while searching for victims,” *IEEE Transactions on Cybernetics*, vol. 44, no. 12, pp. 2719–2732, 2014.
[20] L. Pineda, T. Takahashi, H. Jung, S. Zilberstein, and R. Grupen, “Continual planning for search and rescue robots,” in *Proc. IEEE-RAS International Conference on Humanoid Robots*, 2015, pp. 243–248.
[21] Y. Liu and G. Nejat, “Multirobot cooperative learning for semi-autonomous control in urban search and rescue applications,” *Journal of Field Robotics*, vol. 33, no. 4, p. 512–536, June 2016.
[22] T. Amukele, P. Ness, A. Tobian, J. Boyd, and J. Street, “Drone transportation of blood products,” *Transfusion*, vol. 57, 11 2016.
[23] T. K. Amukele, J. Street, K. Carroll, H. Miller, and S. X. Zhang, “Drone transport of microbes in blood and sputum laboratory specimens,” *Journal of Clinical Microbiology*, vol. 54, no. 10, pp. 2622–2625, 2016.
[24] T. K. Amukele, J. Hernandez, C. L. H. Snozek, R. G. Wyatt, M. Douglas, R. Amini, and J. Street, “Drone transport of chemistry and hematology samples over long distances,” *American journal of clinical pathology*, vol. 148, no. 5, p. 427—435, November 2017.
[25] A. Claesson, A. Bäckman, M. Ringh, L. Svensson, P. Nordberg, T. Djärv, and J. Hollenberg, “Time to delivery of an automated external defibrillator using a drone for simulated out-of-hospital cardiac arrests vs emergency medical services,” *JAMA*, vol. 317, p. 2332, 06 2017.
[26] H. Bahri, M. Chouchene, R. Khemiri, F. Sayadi, and M. Atri, “Fast moving human detection using fourier and hog descriptors based CUDA,” in *International Multi-Conference on Systems, Signals Devices (SSD)*, 2018, pp. 202–207.
[27] D. M. Dinama, Q. A’yun, A. D. Syahroni, I. Adji Sulistijono, and A. Risnumawan, “Human detection and tracking on surveillance video footage using convolutional neural networks,” in *Proc. International Electronics Symposium (IES)*, 2019, pp. 534–538.
[28] H. Yousif, J. Yuan, R. Kays, and Z. He, “Fast human-animal detection from highly cluttered camera-trap images using joint background modeling and deep learning classification,” in *Proc. IEEE International Symposium on Circuits and Systems (ISCAS)*, 2017, pp. 1–4.
[29] J. Wu and G. Zhou, “Real-time UAV video processing for quick-response to natural disaster,” in *Proc. IEEE International Symposium on Geoscience and Remote Sensing*, 2006, pp. 976–979.
[30] T. Suzuki, J. ichi Meguro, Y. Amano, T. Hashizume, R. Hirokawa, K. Tatsumi, K. Sato, and J. ichi Takiguchi, “Information collecting system based on aerial images obtained by a small UAV for disaster prevention,” in *Mechatronics, MEMS, and Smart Materials (ICMIT)*, M. Sasaki, G. C. Sang, Z. Li, R. Ikeura, H. Kim, and F. Xue, Eds., vol. 6794, International Society for Optics and Photonics. SPIE, 2008, pp. 538 – 543.
[31] M. Goodrich, B. Morse, D. Gerhardt, J. Cooper, M. Quigley, J. Adams, and C. Humphrey, “Supporting wilderness search and rescue using a camera-equipped mini UAV,” *Journal of Field Robotics*, vol. 25, pp. 89 – 110, 02 2008.
[32] A. Ahmed, M. Nagai, C. Tianen, and R. Shibasaki, “UAV based monitoring system and object detection technique development for a disaster area,” *International Archives of Photogrammetry, Remote Sensing and Spatial information Sciences*, vol. 37, pp. 373–377, 2008.
[33] P. Rudol and P. Doherty, “Human body detection and geolocalization for UAV search and rescue missions using color and thermal imagery,” in *Proc. IEEE Aerospace Conference*, 2008, pp. 1–8.
[34] M. Andriluka, S. Roth, and B. Schiele, “Pictorial structures revisited: People detection and articulated pose estimation,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, pp. 1014–1021.
[35] M. Andriluka, P. Schnitzspan, J. Meyer, S. Kohlbrecher, K. Petersen, O. von Stryk, S. Roth, and B. Schiele, “Vision based victim detection from unmanned aerial vehicles,” in *Proc. IEEE International Conference on Intelligent Robots and Systems (RSJ)*, 2010, pp. 1740–1747.
[36] A. Gaszczak, T. P. Breckon, and J. Han, “Real-time people and vehicle detection from UAV imagery,” in *Intelligent Robots and Computer Vision: Algorithms and Techniques*, vol. 7878, International Society for Optics and Photonics. SPIE, 2011, pp. 71–83.
[37] M. Donelli, “A rescue radar system for the detection of victims trapped under rubble based on the independent component analysis algorithm,” *Progress In Electromagnetics Research M*, vol. 19, 01 2011.
[38] G. De Cubber, D. Doroftei, D. Serrano, K. Chintamani, R. Sabino, and S. Ourevitch, “The EU-ICARUS project: Developing assistive robotic tools for search and rescue operations,” in *Proc. IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 2013, pp. 1–4.
[39] J. Sun, B. Li, Y. Jiang, and C.-y. Wen, “A camera-based target detection and positioning UAV system for search and rescue (SAR) purposes,” *Sensors*, vol. 16, p. 1778, 10 2016.
[40] M. B. Bejiga, A. Zeggada, A. Nouffidj, and F. Melgani, “A convolutional neural network approach for assisting avalanche search and rescue operations with UAV imagery,” *Remote Sensing*, vol. 9, no. 2, 2017.
[41] A. Al-Kaff, M. Gómez-Silva, F. M. Moreno, A. de la Escalera, and J. Armingol, “An appearance-based tracking algorithm for aerial search and rescue purposes,” *Sensors*, vol. 19, p. 652, 02 2019.
[42] E. Lygouras, N. Santavas, A. Taitzoglou, K. Tarchanidis, A. Mitropoulos, and A. Gasteratos, “Unsupervised human detection with an embedded vision system on a fully autonomous UAV for search and rescue operations,” *Sensors*, vol. 19, no. 16, 2019.
[43] M. Le and M. Le, “Human detection and tracking for autonomous human-following quadcopter,” in *Proc. IEEE International Conference on System Science and Engineering (ICSSE)*, 2019, pp. 6–11.
[44] “COCO - common objects in context web site,” https://cocodataset.org/, accessed: 2020-06-20.
[45] M. Rizk, A. Mroue, K. A. Ali, and J. Charara, “Development and implementation of an embedded low-cost GNSS receiver emulator for UAV testing,” in *Proc. IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, 2019, pp. 101–102.