IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 1, May 2024

Review on Human Detection in Catastrophe Scenarios Using UAV and Deep Learning Techniques

Mrs. Farjana Farvin Sahapudeen[1], Divya Praba V[2], Nuha G[3]

Assistant Professor, Department of Computer Science and Engineering[1]

Students, Department of Computer Science and Engineering[2,3]

Anjalai Ammal Mahalingam Engineering College, Thiruvarur, Tamil Nadu, India

Abstract*:* *In catastrophe scenarios, such as natural disasters or humanitarian crises, rapid and accurate detection of human presence is critical for effective response and rescue operations. Unmanned Aerial Vehicles (UAVs) equipped with computer vision algorithms have emerged as a promising solution for remote sensing and human detection in such scenarios. This survey paper provides an overview of the state-of-the-art techniques for human detection using UAVs, with a focus on the integration of You Only Look Once (YOLO) object detection and DeepSORT (Deep Simple Online and Realtime Tracking) tracking algorithms.*

Keywords: Artificial Intelligence, Deep Learning, Human Detection, You Only Look Once (YOLOv8), DeepSORT.

I. INTRODUCTION

Around the world, countless tragedies have occurred. Human lives have been affected in a variety of ways, from short-term wounds to long-term effects. These disasters have caused great suffering in many lives. Natural disasters like floods, earthquakes, hurricanes, wildfires, etc. can result in fatalities, lifelong injuries, damage to one's home, deprivation of basic necessities, and a host of other unmentioned losses.

In order to assist those in need, the disaster rescue operation needs a large number of organizations and individuals with professional training. It entails preserving life, giving emergency medical attention, and makingsure the sufferers are safe. Professionals such as firefighters, paramedics, search and rescue teams, and volunteersare on hand to assist anyone who were hit by falling debris or found themselves in dangerous situations as a resultof the accident.

Conventional search and rescue operations frequently encounter major obstacles when attempting to enter isolated or dangerous regions, which causes delays in finding survivors and offering aid in a timely manner.But new developments in unmanned aerial vehicle (UAV) technology, along with the strength of deep learning methods, present a potential way to improve human detection in disaster situations. Through the utilization of deep learning algorithms and the aerial perspective offered by unmanned aerial vehicles (UAVs), it is feasible topromptly recognise and localise people who require aid in a variety of environments and circumstances.

This study presents a unique method that combines two cutting-edge models: YOLOv8 for real-time object detection [1] and Deep SORT for robust object tracking. Object detection is a crucial task in computer vision that involves identifying and localizing objects within an image or video frame. Over the years, object detection techniques have evolved from traditional methods to deep learning-based approaches, which have significantly improved accuracy and efficiency

II. LITERATURE REVIEW

A literature review is a survey of scholarly sources on a specific topic. It provides an overview of currentknowledge, allowing you to identify relevant theories, methods, and gaps in the existing research that you can later apply to your paper, thesis, or dissertation topic.

DOI: 10.48175/IJARSCT-18090

ISSN 2581-9429 IJARSCT

In this article [8] ,Unmanned aerial vehicles (UAVs) domain has seen rapid developments in recent years. As the number of UAVs increases and as the missions involving UAVs vary, new research issues surface. An overview of the existing research areas in the UAV domain has been presented including the nature of the work categorized under different groups. These research areas are divided into two main streams: Technological and operational research areas. The research areas in technology are divided into onboard and ground technologies. The research areas in operations are divided into organization level, brigade level, user level, standards and certifications, regulations and legal, moral, and ethical issues. This overview is intended to serve as a starting point for fellow researchers new to the domain, to help researchers in positioning their research, identifying related research areas, and focusing on the right issues.

This paper [9] is about Multiple people tracking. Multiple people tracking is an important task for surveillance. Recently, tracking by detection methods had emerged as immediate effect of deep learning remarkable achievements in object detection. In this paper, we use Faster-RCNN for detection and compare two methods for object association. The first method is simple Euclidean distance and the second is more complicatedSiamese neural network. The experiment result show that simple Euclidean distance gives promising result as object association method, but it depends heavily on the robustness of detection process on individual frames.

Convolutional neural network (CNN) [10], a class of artificial neural networks that has become dominant in various computer vision tasks, is attracting interest across a variety of domains, including radiology. CNN is designed to automatically and adaptively learn spatial hierarchies of features through backpropagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. This review article offers a perspective on the basic concepts of CNN and its application to various radiological tasks, and discusses its challenges and future directions in the field of radiology. Two challenges in applying CNN to radiological tasks, small dataset and overfitting, will also be covered in this article, as well as techniques to minimize them. Being familiar with the concepts and advantages, as well as limitations, of CNN is essential to leverage its potential in diagnostic radiology, with the goal of augmenting the performance of radiologists and improving patient care.

Deep learning (DL), a branch of machine learning (ML) and artificial intelligence (AI) is nowadays considered as a core technology of today's Fourth Industrial Revolution (4IR or Industry 4.0). Due to its learningcapabilities from data, DL technology originated from artificial neural network (ANN), has become a hot topic inthe context of computing, and is widely applied in various application areas like healthcare, visual recognition, text analytics, cybersecurity, and many more. However, building an appropriate DL model is a challenging task, due to the dynamic nature and variations in real-world problems and data. Moreover, the lack of core understanding turns DL methods into black-box machines that hamper development at the standard level. This article [11] presents a structured and comprehensive view on DL techniques including a taxonomy considering various types of real-world tasks like supervised or unsupervised. In our taxonomy, we take into account deep networks for supervised or discriminative learning, unsupervised or generative learning as well as hybrid learning and relevant others.

Drone examination has been overall quickly embraced by NDMM (natural disaster mitigation and management) division to survey the state of impacted regions. The proposed model is used YOLOv3 (you only look once) algorithm for the detection and recognition of actions. In this study [12], it provide the fundamental ideas underlying an object detection model. To find the most effective model for human recognition and detection, we trained the YOLOv3 algorithm on the image dataset and evaluated its performance. The proposed work shows that existing models are inadequate for critical applications like search and rescue, which convinces us to proposea model raised by a pyramidal component extracting SSD in human localization and action recognition.

Drones are unmanned aerial vehicles that can be remotely operated to perform a variety of tasks. They have been used in search and rescue operations since the early 2000s and have proven to be invaluable tools for quickly locating missing persons in difficult terrain and environment. In certain cases, automated human detectionon drone camera feed can help the responder to locate the victims more effectively. In this work [13], they proposethe use of a deep learning method called You Only Look Once version 5, or YOLOv5. The YOLOv5 model is trained using data collected during a simulation of search and rescue operations, where mannequins were used torepresent human victims

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, May 2024

III. METHODOLGY

Human Detection using YOLOv8:

YOLOv8 is a powerful deep learning model that has demonstrated exceptional performance in real-timehuman detection and people counting tasks. Researchers have leveraged the capabilities of YOLOv8 to enhance human detection and people counting in various applications such as surveillance, crowd management, and trafficmonitoring [2]

The YOLO (You Only Look Once) methodology, which YOLOv8 is based on, is known for its ability to recognize objects in real-time. It is particularly useful for detecting people in crowded scenes, where individualsmay overlap and appear in different sizes and shapes due to perspective effects.

Studies have shown that the YOLOv8 [3] algorithm can achieve high Mean Average Precision (MAP) scores, indicating accurate identification of people in still images and video frames. The model's ability to simultaneously detect and count multiple people makes it a valuable tool for applications that require precise people counting, such as retail analytics and event management

Furthermore, the YOLOv8 architecture incorporates advanced techniques like anchor boxes and non- maximum suppression, which help to enhance the overall accuracy of the human detection process. The model'sefficient single-stage processing and optimized network design also enable real-time performance, even on standard hardware.

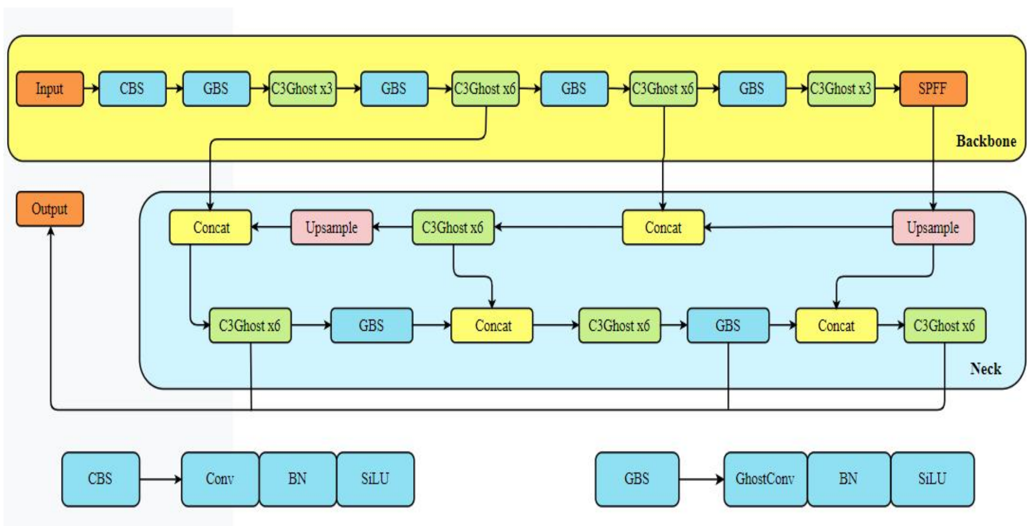


Figure 1. Schematic structure diagram of YOLOv8

Object tracking with Deep SORT:

Deep SORT (Deep Simple Online and Realtime Tracking) is a state-of-the-art algorithm designed for object tracking in videos. Traditional tracking algorithms rely on handcrafted features and mathematical models to estimate object motion and appearance. Examples include the Kalman filter, CamShift, and Mean Shift algorithms. These methods often struggle with complex scenes, occlusions, and variations in object appearance.

Combining deep learning with traditional tracking techniques, DeepSORT provides robust and accurate tracking of objects across frames. The algorithm first utilizes a deep neural network, often based on the popular YOLO (You Only Look Once) architecture, to detect objects in each frame [4]. These detections are then fe2.d into the SORT (Simple Online and Realtime Tracking) algorithm, which associates detections across frames using a combination of appearance features and motion information. This association process is further enhanced by DeepSORT, which leverages deep learning embeddings to improve the accuracy of object matching and tracking[5], especially in crowded or occluded scenes.

DeepSORT relies on a detection model, such as YOLO or Faster R-CNN [6], to detect objects in each frame of a video sequence. Detected objects are passed through a feature extraction network, typically a deep convolutional neural network (CNN), to extract appearance features. DeepSORT associates detected objects with existing tracks by comparing appearance features and predicting motion using methods like the Kalman filter [7]. The algorithm maintains

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, May 2024

a set of active tracks and updates their states based on detections and predictions. It handles track initialization, termination, and handling of occlusions.

Evaluation

The unmanned aerial vehicle (UAV) picture collection is used to train the proposed system, which combines YOLOv8 with deep SORT. We used CVAT (Computer Vision Annotation Tools) to annotate the gathered information because it requires a lot of work to construct class and labels. With the help of this open- source application, computer vision tasks including object identification, image classification, and picture segmentation may be annotated in photos and videos. Unlike object detection models, it is a tool primarily focused on annotation rather than model inference, hence it lacks explicit performance indicators in terms of accuracy or speed.

Evaluation MetricsAccuracy:

Accuracy represents the proportion of correctly classified samples out of the total number of samples inthe dataset.

$$\text{Accuracy} = \textit{Number of Correct Predictions} / \quad \textit{Total Number of Predictions}$$

F1 Score:

The F1 score is the harmonic mean of precision and recall, providing a single score that balances both metrics.

Precision measures the proportion of true positive predictions out of all positive predictions, while recall measures the proportion of true positive predictions out of all actual positives.

$$F1 = 2 \times \frac{\text{Precision} + \text{Recall}}{\text{Precision} + \text{Recall}}$$

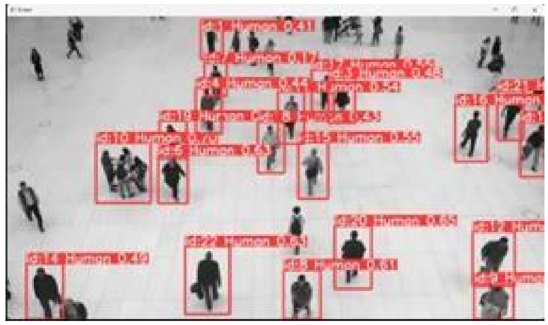
$$\text{Precision} = \frac{TP}{TP+FP} \qquad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \qquad (2)$$

In the above equation (1),(2) TP,FP,FN respectively represent True Positive, False Positive and False Negatives.

Figure 2 Human Detection

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, May 2024



Pros, Cons and Applications

Pros:

- High Performance: YOLOv8 combined with DeepSORT achieved high performance in tracking and time violation detection in parking areas, showcasing the effectiveness of this combination.
- Improved Tracking Accuracy: Integrating DeepSORT with YOLOv8 resulted in enhanced tracking accuracy and robustness, demonstrating the effectiveness of this combination
- Efficient Object Detection: YOLOv8 is particularly effective at handling challenges like occlusions,varying scales, and complex backgrounds, making it suitable for real-time car detection applications
- Seamless Integration: YOLOv8 and DeepSORT work seamlessly together, with YOLOv8 providingprecise detection and DeepSORT offering effective tracking across video sequences

Cons:

- YOLOv8, despite its efficiency in object detection, may face challenges in accurately detecting humansdue to their relatively small size compared to other objects, potentially leading to lower detection accuracy
- DeepSORT, while effective in tracking objects across frames, may struggle with human detection in scenarios involving occlusion or crowded environments, which can lead to tracking errors and identity switches
- Applications:
- Vehicle Detection and Tracking: YOLOv8 and DeepSORT are utilized for car detection and tracking in intelligent traffic systems, demonstrating effectiveness in handling challenges like occlusions and complex backgrounds.
- Object Tracking in Videos: Research proposes real-time and recorded video-based object detection and tracking using YOLOv8 and DeepSORT, highlighting their critical capabilities for computer vision systems
- Real-Time Object Tracking: YOLOv8 and DeepSORT are combined to implement multiple objects tracking in Python, showcasing their powerful combination for object tracking applications.

IV. CONCLUSION

The crucial role that cutting-edge technology and techniques play in improving search and rescue operations is highlighted by the discovery of human remains in catastrophe situations.

The advancement of human detection methods, including convolutional neural network The identification of human remains in catastrophe . Finding catastrophe victims more quickly and accurately hasbeen made possible by advancements in human detection technologies, such as hybrid algorithms, convolutionalneural networks, and thermal imaging. When paired with multi-sensor technology such as robots and drones, thesemethods provide a complete way to find and rescue persons who are buried under rubble during disasters. The need of real-time data gathering and analysis for prompt and accurate human identification in difficult circumstances is highlighted by the necessity for integrated sensor technologies, machine learning, and 5G networks

REFERENCES

[1] 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, 32(9):1627–1645, 2010.

[2] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, pages 886–893. IEEE, 2005.
[3] Chengji Liu ,Yufan Tao ,Jiawei Liang ,Kai Li1 ,Yihang Chen, ”Object Detection Based on YOLO Network” 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC 2018)
[4] Zhang, X.; Zhou, X.; Lin, M.; Sun, J. ShuffleNet: An extremely efficient convolutional neural network for mobile devices. arXiv 2017, arXiv:1707.01083v2.
[5] Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 580–587
[6] N. Shlezinger, Y. C. Eldar and S. P. Boyd, "Model-Based Deep Learning: On the Intersection of Deep Learning and Optimization," in IEEE Access, vol. 10, pp. 115384-115398, 2022, doi: 10.1109/ACCESS.2022.3218802
[7] M. Al-Qatf, Y. Lasheng, M. Al-Habib and K. Al-Sabahi, "Deep Learning Approach Combining Sparse Autoencoder With SVM for Network Intrusion Detection," in IEEE Access, vol. 6, pp. 52843-52856, 2018, doi: 10.1109/ACCESS.2018.2869577
[8] Demir, K. A., Cicibas, H., & Arica, N. (2015). Unmanned Aerial Vehicle Domain: Areas of Research. Defence Science Journal, 65(4), 319-329. https://doi.org/10.14429/dsj.65.8631
[9] Dina Chahyati, Mohamad Ivan Fanany, Aniati Murni Arymurthy,Tracking People by Detection Using CNN Features,Procedia Computer Science,Volume 124,2017,Pages 167-172,ISSN 1877-0509, https://doi.org/10.1016/j.procs.2017.12.143
[10] Yamashita, R., Nishio, M., Do, R.K.G. et al. Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9
[11] Sarker, I.H. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. SN COMPUT. SCI. 2, 420 (2021). https://doi.org/10.1007/s42979-021-00815-1
[12] B. Valarmathi, Jain Kshitij, Rajpurohit Dimple, N. Srinivasa Gupta, Y. Harold Robinson, G. Arulkumaran, Tadesse Mulu, "Human Detection and Action Recognition for Search and Rescue in Disasters Using YOLOv3 Algorithm", Journal of Electrical and Computer Engineering, vol. 2023, Article ID 5419384, 19 pages, 2023. https://doi.org/10.1155/2023/5419384
[13] Zaman, Fadhlan & Tahir, Noorita & Yusoff, Yusnani & M Thamrin, Norashikin & Hasmi, Ahmad Hafizam. (2023). Human Detection from Drone using You Only Look Once (YOLOv5) for Search and Rescue Operation. Journal of Advanced Research in Applied Sciences and Engineering Technology. 30. 222-235. 10.37934/araset.30.3.222235.