2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) | 979-8-3503-8247-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/UPCON59197.2023.10434788

UAV based Human Detection for Search and Rescue Operations in Flood

Samta Gaur
Computer Science and Engineering Department
Motilal Nehru National Institute of Technology Allahabad
Prayagraj, India - 211004
samta.2022p14@mnnit.ac.in and

J. Sathish Kumar
Computer Science and Engineering Department
Motilal Nehru National Institute of Technology Allahabad
Prayagraj, India - 211004
sathish613@mnnit.ac.in

***Abstract*—During disastrous events like floods, the urgency to rescue people facing immediate danger poses a significant challenge. Traditional search operations led by humans often struggle due to dangerous conditions and limited resources. As a response, a new application has emerged: employing Unmanned Aerial Vehicle (UAV) for rescue missions, utilizing image processing to locate individuals in distress. However, existing application relies on deep learning based approach for image analysis suffer from slow human detection, a crucial drawback given the time sensitive nature of such scenarios. To tackle this issue, we have proposed a novel approach using You Look Only Once version 4 (YOLOv4) for rapid human detection. YOLOv4 is renowned for its speed, surpassing traditional methods by swiftly identifying humans within flooded environments. To train and validate our approach, we have used Common Objects in Context (COCO) dataset (pre-trained or labeled). The COCO dataset's large size and wide variety (real-world scenarios, including images from natural scenes, indoor environments, and crowded public spaces) of images make it a valuable asset for training and testing human detection models. Through exhaustive experimentation, our method showcases an impressive average accuracy of 79.46% in accurately identifying humans amidst flood-ravaged surroundings. By harnessing the efficiency of YOLOv4 and the diversity of the COCO dataset, our approach holds the potential to significantly enhance the efficiency of UAV-assisted search and rescue operations during critical moments, ultimately resulting in the preservation of more lives in the face of disasters. As well as the human detection information collected from the model facilitates the rescue team, assisting in tasks such as distributing food, providing first aid, and formulating rescue strategies.**

Keywords—Human detection, Unmanned Aerial Vechicle (UAV), You Look Only Once version 4 (YOLOv4), Darknet, flood.

I. Introduction

Floods are a major natural disaster commonly experienced in monsoon countries like India [1]. They result in extensive damage to both properties and human lives, leading to devastating consequences that affect communities, economies, and ecosystems. The impact of floods goes beyond the immediate physical damage, and detecting individuals in flooded environments becomes crucial for effective rescue management, disaster response, flood assessment, and overall disaster preparedness.

During a flood, people may become stranded, trapped, or in need of urgent rescue [17]. Effective and timely detection of individuals in these situations is crucial for disaster response teams to plan and execute rescue operations efficiently. Lives can be saved, and the suffering of affected communities can be mitigated through accurate and swift human detection.

The use of advanced technologies like drone imaging plays a vital role in flood assessment and disaster preparedness. Drones equipped with high-resolution cameras can navigate through the flood-affected areas, capturing images and providing crucial visual data to rescue teams. Machine learning algorithms and computer vision techniques are applied to these drone images to identify and count individuals present in the flooded regions.

The benefits of human detection in flooded environments are multi-faceted. Firstly, it allows rescue teams to prioritize their efforts by identifying areas with the highest concentration of people in distress. This way, they can allocate resources effectively and rescue those who need immediate assistance. Secondly, it aids in assessing the scale of the disaster and estimating the number of people affected, enabling authorities to coordinate relief efforts accordingly.

Traditionally, human detection in flood scenarios has relied on manual search and rescue operations, which can be slow, labour-intensive, and dangerous for rescue teams. In such situations, time is of the essence, and efficient human detection methods are essential to minimize casualties and aid timely rescue efforts.

Deep learning has revolutionized computer vision tasks, including object detection, by enabling the development of advanced algorithms capable of learning intricate patterns and representations from vast amounts of data [18]. Among the prominent deep learning models, the You Only Look Once version 4 (YOLOv4) approach has gained popularity for its real-time capabilities and high accuracy [19]. By integrating YOLO-based human detection algorithms with UAVs, the system can efficiently survey large areas, provide real-time analysis, and assist rescue teams in identifying human presence in flood-affected regions.

This paper's contributions are as follows:

- Application of Open-Source Drones for practical Search and Rescue (SAR) operations during floods.
- Swift and effective real-time surveillance of affected regions.
- Detection of stranded individuals through the implementation of the YOLOv4 object detection algorithm.
- Experimental analysis of a flood-prone area of Uttar Pradesh, offers insights into the drone's time efficiency in area assessment.

The II section delves into the realm of related studies, followed by the III section which presents the details of the proposed work. The IV section is dedicated to the discussion of the experimental setup, while the subsequent V section showcases the obtained results. Ultimately, the VI section summarizes the obtained findings, offering a conclusive endpoint to the discourse. Furthermore, it hints at prospective avenues for future investigation and exploration.

II. Related Work

In this section we have discussed about various human detection approaches performed by previous researches. Dhanushree M et al. [1], used data augmentation techniques to simulate diverse weather conditions experienced during floods, including rain, clouds, and fog. The augmented datasets are then utilized to evaluate the performance of the proposed HOG-based Robust Human Object Detection (HOG-based RHOD) algorithm. Subhashree Rath et al. [2], proposes a flood alert system using IoT and Machine Learning. The system monitors flood-causing factors and sends quick alerts via SMS or emails when water levels are critical. It predicts human presence in flooded areas using image processing from satellite, drones, and CCTV cameras. Emmanuel Okafor et al. [3] explored a unique data augmentation method known as ”rotation matrix images”. This technique involved using a grid of cells, each containing a rotated image and a natural background. They also investigated incorporating color constancy versions of the images to enhance CNN performance for classification purposes.Vallimeena P et al. [4] conducted a study to estimate flood water levels using crowd-sourced images of people. The experimental outcomes successfully determined flood water levels by analyzing the average height of individuals. The categorization of people was achieved through gender and age identification. Hua-Tsung Chen et al. [7] sought to augment the training dataset for deep learning algorithms by transforming regular images into aeriallike representations. They utilized methods such as border padding, perspective transformation, and image rotation, which demonstrated promising results when incorporated alongside established datasets. In their research, Keerthana et al. [8] employed the You Only Look Once (YOLO) deep learning network to identify survivors of flood incidents. Degrossi et al. [9] proposed a crowdsourcing-based approach for flood risk management volunteer information. Payal Mittal et al. [10] presented varied object detection results from UAV images, utilizing deep learning techniques. Mbaitiga Zacharie et al. [11] introduced a rapid human object detection method using HSV color mapping. Chaudhary et al. [12] estimated water levels from social media flood scene images. Aarthy et al. [13] extracted flood-related Twitter posts for water level and flood risk estimation. These studies underscore the importance of human object detection in flooded settings for accurate water level assessment and rescue efforts. From the above related works we have found that the approaches takes more time during human detection. To reduce this drawback we have proposed a model which uses YOLOv4 trained on COCO dataset which has been discussed in the next section.

III. Proposed Work

This section will delve into the methodology outlined in this paper for constructing a model designed to swiftly detect humans in flooded regions. Prior to delving into the model details, an overview of the dataset employed to train the model will be presented.

A. Dataset

In the proposed work we have used Common Objects in Context (COCO) dataset for training our model [15]. The COCO dataset is a widely used large-scale dataset for object detection, segmentation, and captioning tasks in computer vision. It was introduced by Microsoft in collaboration with several academic institutions. The COCO dataset contains more than 200,000 images, covering 80 different object categories, making it diverse and comprehensive.

The COCO dataset includes a specific category for human detection, making it an essential resource for training and evaluating human detection models. It contains a diverse set of images that depict various real-world scenarios, including images from natural scenes, indoor environments, and crowded public spaces.

For human detection, each image in the COCO dataset is annotated with bounding boxes that precisely outline the location of humans present in the image. These bounding boxes provide essential information for object detection algorithms to identify and localize human instances accurately. These labels provide a standardized way to identify and categorize objects within images and used for human detection with the proposed model. The COCO dataset's large size and wide variety of images make it a valuable asset for training and testing human detection models.

B. Model

In flooded environments, a fast human detection model is crucial for swift response. This model enables the rescue team to quickly identify people, aiding in food distribution, first aid, and rescue planning. This section introduces our rapid human detection model. This human detection process relies on the You Only Look Once version 4 (YOLOv4) algorithm, a cutting-edge real-time object detection technique in computer vision. YOLOv4, introduced in 2020, is a onestage object detection algorithm that combines the strengths of its predecessors (YOLOv1 to YOLOv3) [19]. It achieves an optimal balance between detection speed and accuracy, setting a new standard in real-time, high-precision object detection.

YOLOv4's principle is to simultaneously predict bounding boxes and class probabilities for multiple objects in a single neural network forward pass, making it significantly faster than traditional multi-stage object detection methods. The input image is divided into a grid, with each grid cell predicting bounding boxes and class probabilities. Non-maximum suppression is then applied to retain only the most confident and accurate predictions while removing duplicates.YOLOv4 incorporates several improvements to enhance both detection accuracy and speed.

To implement and train YOLOv4 models, the project utilizes the Darknet framework, an open-source deep learning platform written in C and CUDA. The Darknet framework is optimized for efficient computation on both CPUs and GPUs, making it a perfect match for the demanding requirements of YOLOv4's real-world object detection. The Darknet framework uses the class labels or vector size of the COCO Dataset for training and validation purpose. The vector size is of COCO data is subjected to the CSPdarknet for taining the model.

The model architecture, depicted in Figure 1, comprises three key components: Backbone, Neck, and Prediction. [14] YOLOv4 offers multiple choices for backbone architectures, including CSPDarknet53 and CSPResNeXt50. By incorporating the ResNet structure, YOLOv3 introduced the residual module to create Darknet53. In YOLOv4, building upon the effectiveness of Cross-Stage Partial Network (CSPNet), CSPDarkNet53 was constructed [20]. In this paper we are using CSPDarknet as backbone. CSPDarknet is an open source convolutional network developed on C++. It has total 162 layers and detections are made at 139, 150 and 161 layer. These backbones capture hierarchical features from input images, enhancing the model's feature representation capabilities. Within the residual module, higher-level feature information is produced, effectively implementing residual learning while reducing model parameters and enhancing feature learning.

The Neck section encompasses SPPNet and Path Aggregation Network (PANet). SPPNet convolves the feature layer thrice and applies maximum pooling with various kernel sizes. Concatenating these pooled results and subjecting them to convolution enhances the network's receptive field. PANet operates post-Backbone and SPPNet, up-sampling the feature layers, fusing them through concatenation, and further downsampling. This process repeats for more feature fusion (five times).

In the Prediction module, predictions are made using network-extracted features which is validated by the COCO dataset. For instance, considering a 13 × 13 grid, each grid contains three preset prior frames. Network predictions adjust these frames' positions, which are then refined through nonmaximum suppression (NMS) for the final predictions. The conceptual framework of our proposed work is shown in the Figure 2.

The decision to use YOLOv4 was driven by its efficiency, speed, and accuracy, making it an ideal choice for realtime object detection tasks in various applications, including surveillance, autonomous vehicles, and robotics. Its ability to swiftly and accurately identify humans from drone imagery makes it an invaluable tool for rescue and disaster response operations.

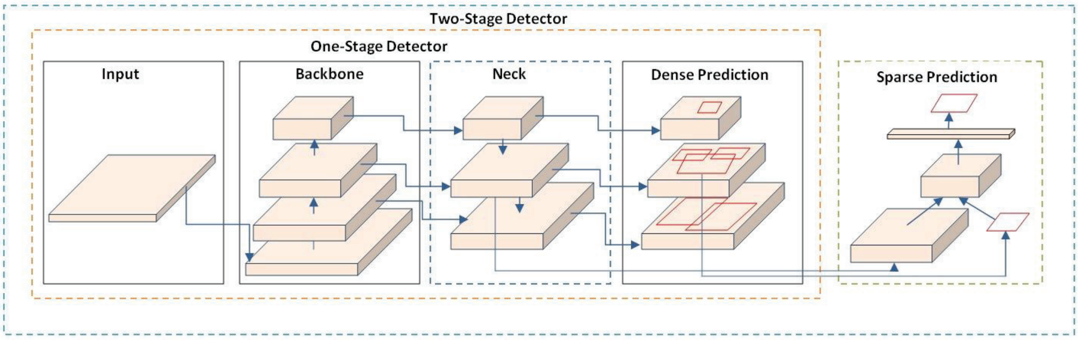

Fig. 1. YOLOv4 Architecture

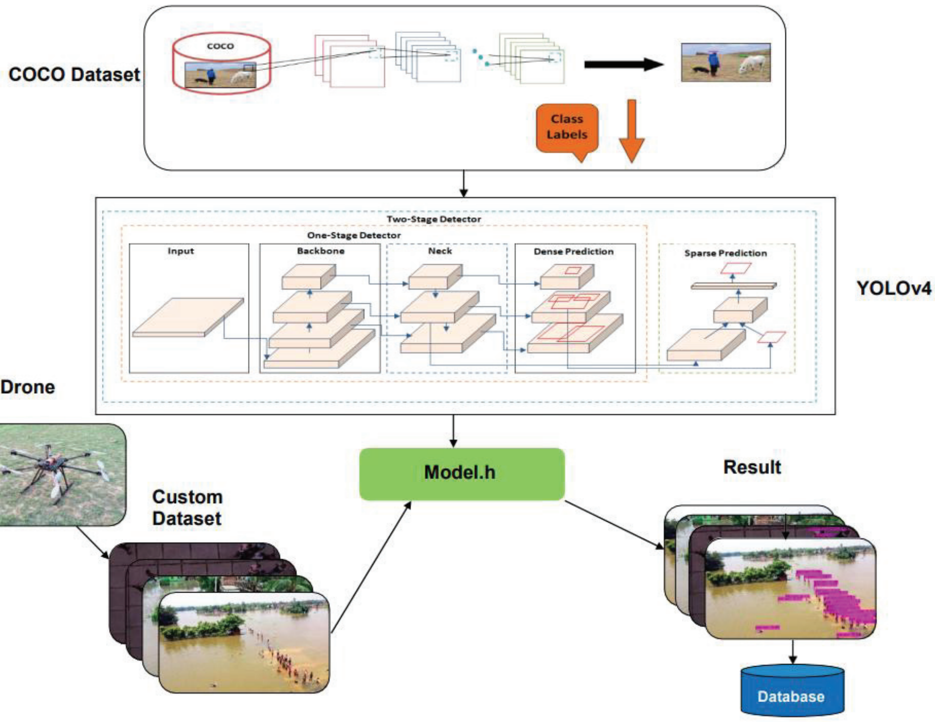

Fig. 2. Conceptual framework

C. Human Detection

At the ground control station, human detection is performed on a PC. Equipped with a mounted camera, the hexacopter captures images essential for detecting stranded individuals. The hardware setup involves a Raspberry Pi 4 connected to a UAV (hexacopter), powered by a LiPo battery. Once connected to Wi-Fi and GPS, the system displays the live video feed on a webpage, initiating custom data collection.

To validate our approach, we conducted tests using a drone to capture data from different areas. The captured images include scenarios such as people on roads and individuals in flood-affected regions. After establishing a link between the hexacopter and the ground control station, the video feed is acquired and subsequently subjected to analysis. To enable detection, individual frames are extracted from the video. These extracted images are then subjected to thorough testing and evaluation. For testing, each image is fed into the saved ”model.h” file. Following this step, the image's validation is performed in conjunction with the class labels found in the COCO dataset. This series of steps ultimately yields the final result.

Subsequently, the YOLOv4 algorithm classifies the detected objects, with a specific focus on identifying humans. If the system successfully identifies an object as a human, the output data can be stored as a database in the computational system.

Overall, the aim of the paper is to harness the capabilities of YOLOv4 and real-world camera feed analysis to swiftly and accurately identify humans in flood affected areas. This technology has the potential to significantly contribute to timely and effective disaster management and rescue operations revolves around utilizing real-world camera feeds for human detection.

IV. Experimental Setup

Certainly, in this section, we will delve into the tools and technologies employed for the human detection model.

A. System Hardware

The system uses a hexacopter as shown in figure 3. The hexacopter uses ZD850 Hexa-Rotor frame. The Hexacopter Frame is constructed entirely from 3K TWILL carbon fiber, ensuring its durability and strength. The frame's arms are made using 20mm carbon fiber tubes, effectively preventing arm breakage at the motor mounts even during hard landings. With its six arms, this Hexacopter Frame is designed to achieve flight stability and reliable performance in the air. The frame is specifically designed to support six Emax MT2213 935KV brushless motors, which are effortlessly controlled by Emax Bullet Series 35A ESCs. For maneuvering the drone, the FlySky FS-i6S 2.4GHz 10CH AFHDS 2A RC Transmitter is utilized in combination with the FS-iA10B 10CH Receiver.The drone is equipped with a high-capacity Orange 5200mAh 4S 40C/80C (14.8V) Lithium Polymer Battery, pack to power its operations. The on-board flight controller, Pixhawk provides advanced stability and flight control algorithms, ensuring precise and stable flight maneuvers.

In terms of hardware, a hexacopter is the central component, carrying out missions over disaster-stricken areas. The hexacopter is equipped with a camera that provides the live video feed required for detection. An open-source APM 2.8 flight controller manages the hexacopter's functionalities, powered by a LiPo battery for flight endurance. Live coordinates are obtained from an external GPS module, vital for effective rescue operations. The Ground Control Station involves image processing, with the Mission Planner software monitoring the hexacopter's performance. Radio Telemetry, operating at 433 MHz and 100mW, facilitates data transfer for drone monitoring. This telemetry set enables real-world observation of the hexacopter's performance and position. Output video signals are processed to detect stranded humans using realworld object detection techniques.

Fig. 3. Hexacopter used for capturing images.

B. Google Colab

Google Colab (Colaboratory) is a cloud-based platform by Google that provides a free environment for writing and running Python code [16]. It offers Jupyter Notebook integration, access to GPUs and TPUs, easy sharing and collaboration, and comes with pre-installed libraries. It's suitable for tasks like coding, data analysis, and machine learning, especially for those without powerful hardware.

V. Results

A. Case:1 Testing human detection using drone images.

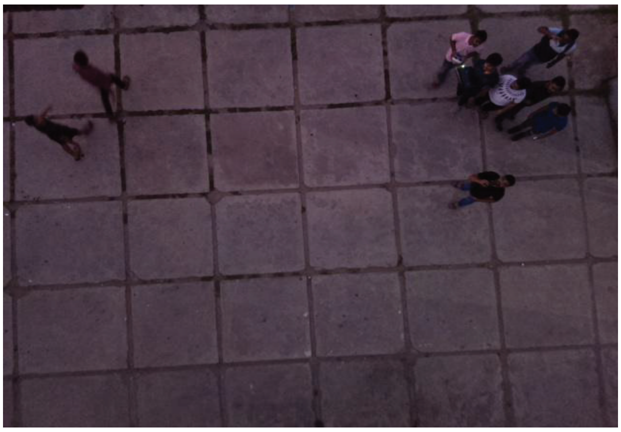

Fig. 4. Trial 1 input drone image

In the initial test, we analyzed the drone-captured image illustrated in Figure 4. During this trial, our model accurately identified 7 out of the 9 individuals present, as demonstrated in Figure 5. This yielded a detection accuracy of approximately 77.77 percent.

Fig. 5. Trial 1 output drone image.

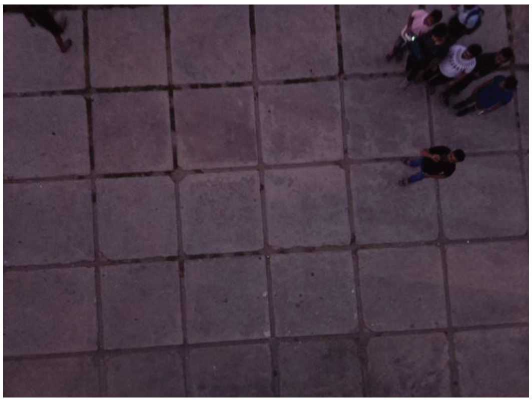

Fig. 6. Trial 2 input drone image.

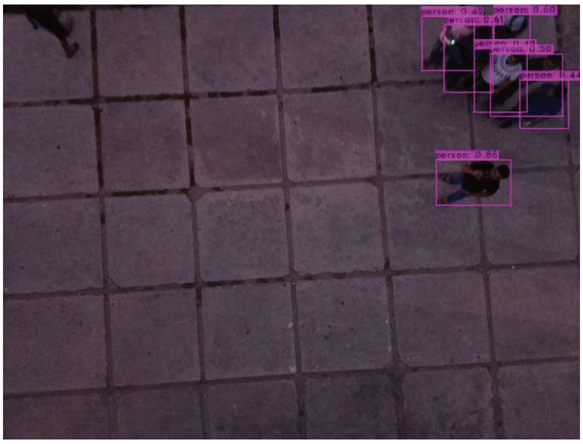

Fig. 7. Trial 2 output drone image

Subsequently, in the second test iteration, we once again examined a drone image as depicted in Figure 6. Notably, our model successfully detected all 7 individuals shown in Figure 7, resulting in a remarkable accuracy of 100 percent for this particular trial.

B. Case:2 Testing human detection using flooded images.

In the second set of tests, our focus shifted towards analyzing images depicting flooded areas, as visually represented in Figures 8 and 10. This specific context introduced unique challenges due to the dynamic and complex nature of flood environments.

In the first trial conducted within this flooded scenario shown in Figure 8, the model showcased its capability by successfully detecting and localizing 40 individuals out of a total of 51 present within the imagery as depicted in Figure 9. This impressive performance resulted in an accuracy rate of approximately 78.43 percent.

In the subsequent trial, the input image is shown in Figure 10. The scenario involved a reduced number of individuals, with only 5 persons present within the flooded area depicted in the Figure 11. Here, the model's prowess persisted, as it managed to identify and pinpoint 4 out of the 5 individuals accurately. This remarkable achievement translated to an accuracy rate of 80 percent in this specific trial.

Table 1 presents the outcomes of testing on a custom dataset acquired through UAV (Unmanned Aerial Vehicle). This specific dataset encompasses 20 images featuring human objects and an additional 50 images showcasing non-human entities. Among the set of 20 images featuring humans, a total of 112 humans were initially present. Remarkably, the detection process successfully identifies 89 out of these 112 humans, translating to an accuracy rate of 79.46 percent.

Fig. 8. Trial 1 input flood image.

Fig. 9. Trial 1 output flood image.

Fig. 10. Trial 2 input flood image.

Fig. 11. Trial 2 output flood image

TABLE I. TEST RESULTS ON CUSTOM DATASET

Trials	Total No. of Images	Total No. of Humans present	Total No. of Humans detected	Average Accuracy
20	20	112	89	79.46

VI. CONCLUSION

A UAV-based approach for human detection for SAR operations in flood affected areas is presented in this paper. The YOLOv4 algorithm have applied to a custom dataset generated from on-board camera footage captured by the drone. The project uses open-source hardware that makes it economical and easy to build. Utilizing the CSPDarknet architecture, the model achieves the remarkable feat of detecting humans in just 55.0456 milliseconds. This efficiency is attributed to its object detection prowess across three distinct layers: 139, 150, and 161. In the future, we intend to augment the current approach by incorporating additional features to identify humans with fewer body parts in flooded images. Additionally, we will mount GoPro or DJI Osmo capture high-quality aerial footage, which will widen the search area by flying the drone at higher altitudes. With these enhancements, the drone's full autonomy will lead to improved datasets and faster, highly accurate human detection capabilities. The system will be poised to revolutionize SAR operations, providing a more effective and efficient approach to disaster response and rescue missions.

ACKNOWLEDGEMENT

The work presented here is performed as part of the research project entitled "Internet of Things (IoT) and Machine Learning (ML) based Smart UAV-Assisted Emergency System for Disaster Management", that is supported and funded by the Council of Science and Technology (CST), Government of UP, India (Project and Sanction No. 1199 and CST/D/1219).

REFERENCES

[1] CM, Bhatt. "Robust human detection system in flood related images with data augmentation." Multimedia Tools and Applications, Vol 82. No 7, pp: 10661-10679, 2023.

[2] Rath, Subhashree, et al. "IoT and ML based Flood Alert and Human Detection System." In the proceedings of 4th International Conference on Smart Systems and Inventive Technology (ICSSIT), IEEE, 2022, pp: 132-137.

[3] Okafor, E., Schomaker, L., and Wiering, M. A,. An analysis of rotation matrix and colour constancy data augmentation in classifying images of animals. Journal of Information and Telecommunication, Vol 2, No 4, pp: 465-491, 2018.

[4] Okafor, E., Schomaker, L., and Wiering, M. A, An analysis of rotation matrix and colour constancy data augmentation in classifying images of animals. Journal of Information and Telecommunication, Vol. 2, No 4, pp: 465-491, 2018.

[5] Vallimeena, P., Nair, B. B., and Rao, S. N., Machine vision based flood depth estimation using crowdsourced images of humans. In Proceedings of IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2018, pp: 1-4.

[6] Liu, T., Fu, H. Y., Wen, Q., Zhang, D. K., and Li, L. F. Extended faster R-CNN for long distance human detection: Finding pedestrians in UAV images. In the proceedings of IEEE International Conference on Consumer Electronics (ICCE), 2018, pp: 1-2.

[7] Nguyen, D. T., Li, W., and Ogunbona, P. O. Human detection from images and videos: A survey. Pattern Recognition, Vol 51, pp: 148-175, 2016.

[8] Chen, H. T., Liu, C. H., and Tsai, W. J. Data augmentation for cnn-based people detection in aerial images. In the proceedings of IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 2018, pp: 1-6.

[9] Degrossi, L. C., de Albuquerque, J. P., Fava, M. C., and Mendiondo, E. M. Flood Citizen Observatory: a crowdsourcing-based approach for flood risk management in Brazil. In the proceeding of SEKE, 2014, pp: 570-575.

[10] Mittal, P., Singh, R., and Sharma, A. Deep learning-based object detection in low-altitude UAV datasets: A survey. Image and Vision computing, Vol. 104, pp: 104046, 2020.

[11] Zacharie, M., Fuji, S., and Minori, S. (2018, October). Rapid human body detection in disaster sites using image processing from unmanned aerial vehicle (UAV) cameras. In the proceeding of International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), 2018, Vol. 3, pp: 230-235.

[12] Chaudhary, P., D'Aronco, S., Moy de Vitry, M., Leitão, J.P. and Wegner, J.D.,. Flood-water level estimation from social media images. ISPRS Journals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol 4 No 2/W5, pp.5-12, 2019 .

[13] Aarthy, K. P., Ramesh, K., Prasanna Venkatesh, V. P., Chitrakala, S., and Bhatt, S. C. (2022, February). Social Media Analysis for Flood Nuggets Extraction Using Relevant Post Filtration. In the proceedings of International Conference on Computational Intelligence and Data Engineering: ICCIDE, 2021, pp. 201-212.

[14] Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

[15] Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... and Zitnick, C. L. Microsoft COCO: Common objects in context. In the proceeding of Computer Vision–ECCV 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Part V 13 pp. 740-755.

[16] Bisong, E., and Bisong, E. Google colaboratory. Building machine learning and deep learning models on google cloud platform: a comprehensive guide for beginners, 2019, pp. 59-64.

[17] Ragini, J. R., Anand, P. R., and Bhaskar, V. Mining crisis information: A strategic approach for detection of people at risk through social media analysis. International journal of disaster risk reduction, 2018, Vol 27, pp 556-566.

[18] Sharma, V., and Mir, R. N. A comprehensive and systematic look up into deep learning based object detection techniques: A review. Computer Science Review, 2020, Vol 38, Article 100301.

[19] Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

[20] Yu, J., and Zhang, W. Face mask wearing detection algorithm based on improved YOLO-v4. Sensors, Vol 21 No. 9, pp. 3263, 2021.