Performance Study of YOLOv7 and YOLOv8 in Victim Detection for Disaster Scenarios

Megharani B. Mayani[1]
[1]Government Polytechnic Belgaum, Electronics and Communication Engineering, Department of Collegiate and Technical Education, Government of Karnataka, Belagavi 590018

Abstract—**In the aftermath of natural disasters like earthquakes, rapid identification of victims is vital for efficient rescue operations. This study evaluates the performance of two advanced object detection models, YOLOv7 and YOLOv8, for victim detection in disaster scenarios. Both models were trained on a dataset simulating post-disaster environments, aiming to recognize human bodies among debris and challenging conditions. The experimental results showed an accuracy of 58% for YOLOv7 and a significantly improved accuracy of 81% for YOLOv8, indicating a notable advancement in the latter's ability to detect victims accurately. This comparative analysis not only highlights the superior performance of YOLOv8 but also explores the strengths and limitations of YOLOv7 in terms of detection accuracy, precision, and recall. The findings underscore the potential of employing YOLOv8 for real-time disaster response systems, where quick and reliable victim identification can save lives. Future work could focus on enhancing the model's robustness in diverse scenarios and integrating additional sensor data for improved detection.**

Index Terms—**victim detection, disaster response, YOLOv7, YOLOv8, emergency rescue, deep learning models, accuracy comparison, real-time detection**

I. INTRODUCTION

Disaster like flood, landslides, earthquake are inevitable and most of the time lead to loss of many lives and give lot of problems in the search mission. One of the most important activities in these operations is to quickly identify people that are still alive and that perhaps may be stuck under the rubble or in some other hard to access area. The technology from deep learning and computer vision fields has recently demonstrated its applicability for victim detection and increasing its accuracy in disaster situations. This research proposes the application of advanced object detection algorithms namely YOLOv7 and YOLOv8 in improving victims' detection after a disaster.

Real time object detection has been implemented extensively using YOLO (You Only Look Once) for its speed and accuracy. It works by tiling an image, and then regressing bounding box locations of all object within these grids and classification of these objects. The present YOLOv7 and YOLOv8 bring forward several architectural changes that work towards refining the detection capability in multiple tasks. The specific characteristics of YOLOv7 include faster inference speed and portability for deployment on real-time applications in low-powered devices. While, YOLOv8 includes more refined methods introduced with the principles of achieving higher detection rates, especially in the circumstances involving dense areas of the background, which appear primarily in the post-earthquake scenarios.

In this paper, we integrate both YOLOv7 and YOLOv8 model for the identification of victims in areas affected by disaster with focus on evaluation of accuracy, precision, recall and time consumption. These models are trained by a custom dataset that consists of images illustrating post-disaster scenarios such as different environments and lighting, and different levels of visibility of the victims. The evaluation criteria are targeting the sensitivity of the model where the sensitivity estimates the ability of the model to detect victims; the precision and the recall of the model that evaluates the dependability of the model.

In our investigation, the performance differences between YOLOv7 and YOLOv8 are very distinct with a 58% accuracy for YOLOv7 and an 81% accuracy for YOLOv8. The result showing that YOLOv8 video processing is faster suggests that with improved architecture and features extraction, YOLOv8 has a higher potential for the complex and dynamic video scenes that are characteristic of disasters.

Automating victim detection in responding to emergencies is a core focus of the study arguing that incorporating deep learning algorithms improves the decision-making speed in emergencies. Moreover, extending these models to the UAV, robotics, or PDA platforms can help dramatically reorganize how search and rescue missions are performed, reducing hazards to the rescue teams and improving the disaster response operations. The work in the future would deal with the enhancement of model robustness for various disasters through the integration of sensors at different modalities and the application of real-time adaptation.

This study seeks to implement the state-of-the-art Deep learning Models for the detection of victims in a disaster area by identifying gaps within the current state of Disaster Management to enhance the effectiveness of advanced rescue operations.

II. RELATED WORK

The application of YOLO (You Only Look Once) models in disaster victim detection has been widely explored due to their real-time object detection capabilities. These models have proven to be efficient in identifying victims in post-disaster scenarios, offering significant improvements in the speed and accuracy of rescue operations.

Recent work has demonstrated the effectiveness of YOLOv5 for real-time detection of victims in natural disaster situations. An enhanced version, termed YOLO-MSFR, showed substantial improvements over traditional YOLOv5 by integrating multi-scale feature representation. This adjustment allowed for better detection of small objects and complex scenes, resulting in higher accuracy during disaster response operations [1]. Similarly, hybrid approaches that combine YOLO with other detection techniques have been employed to improve accuracy in detecting indoor disaster victims. These methods leverage YOLO's speed with complementary algorithms to enhance detection precision, particularly in cluttered environments [2].

The use of UAV technology integrated with YOLO models has been another area of active research, aiming to automate search and rescue missions. UAVs equipped with YOLO for aerial surveillance can effectively scan large areas, providing critical information on the location of victims and directing rescue teams to the right spots. Such applications not only enhance the efficiency of search efforts but also minimize risks to rescue personnel by using drones to access hazardous areas [3] [6]. The combination of UAVs with YOLO has demonstrated significant potential in scaling search operations to cover vast and inaccessible regions rapidly.

Earlier implementations of YOLO, such as YOLOv3, have also been explored for specific applications like detecting humans during search and rescue missions. A study on human detection and action recognition using YOLOv3 reported promising results in identifying individuals during disasters, especially when combined with other machine learning techniques for post-disaster data curation [4] [7]. These advancements highlight YOLO's versatility in adapting to various real-world scenarios.

Furthermore, comparing YOLO to other popular object detection models, such as Retina Net and SSD, has been an essential part of understanding its strengths and limitations. While YOLO provides higher speed and satisfactory accuracy for real-time applications, models like Retina Net may achieve better performance in detecting small objects due to their use of focal loss, which better addresses the imbalance between background and object classes [5].

Thermal imaging combined with YOLO has been explored for detecting victims in low-visibility conditions, such as smoke-filled or dark environments. This integration has shown to significantly improve detection rates compared to conventional visual-based methods alone. The use of thermal data provides additional context that complements YOLO's visual detection capabilities, making it more effective for disaster response applications where visibility is often compromised [8]

Recent research has also focused on the challenges associated with accurately detecting victims amidst complex backgrounds. YOLOv8, an advanced version of the model, has shown substantial improvements in feature extraction capabilities, which is critical in dynamic environments such as disaster sites. The model's improved architecture enables more reliable detection by better handling occlusions and varying object sizes, which are common in post-disaster scenarios [1] [6].

Other innovative applications of YOLO include its use in detecting environmental hazards like debris and abandoned objects in disaster zones. This capability extends YOLO's utility beyond victim detection, aiding in environmental cleanup and hazard identification tasks [9].

Enhancements in surveillance systems using YOLO have also been explored to enable real-time tracking and monitoring of objects in disaster-prone areas. Such systems, which incorporate YOLO for object detection and tracking, can provide continuous monitoring, improving situational awareness during rescue operations [10].

In conclusion, YOLO-based models have made significant strides in the field of disaster response, with ongoing research focusing on overcoming limitations such as occlusion handling, real-time performance optimization, and the integration of multi-modal data. These advancements, driven by the latest versions like YOLOv8, hybrid approaches, and UAV-based applications, continue to push the boundaries of automated victim detection and disaster management.

III. DATASET

Simulated Disaster Victim (SDV1 and SDV2) Dataset

The Simulated Disaster Victim (SDV1 and SDV2) Dataset is an openly available dataset aimed for training and testing Machine Learning algorithms in the week of identifying disaster victims. The dataset is available on IEEE Data port and is meant to contribute to investigations into search and rescue scenarios in order to mimic the post-earthquake isolation conditions where people are alive and must be found.

This dataset involves SDV1 and SDV2 which contains high-definition images taken under different simulated disaster scenarios. Both the objects in the images and the images themselves were collected in different directions and different lighting conditions to simulate the environment of disaster that can happen during live search-and-rescue operations such as during a building collapse, landslide or following an earthquake. The dataset is images where the images contain ground truth, the rectangles indicating the positions of the simulated victims for training and evaluating object detection algorithms. The annotations included information regarding partial occlusion of the victims which corresponds to the real-life condition when victims may be covered by debris, as well as different postures which represents the victims oriented in different angles.

Since SDV1 deals with less scenarios and having less background information, SDV1 is ideal for initial-model training and experiments. Apart from this, it enables models to learn the characteristics related to human figures and identity of victims in a disaster scenario. SDV2, on the other hand, offers methodically difficult samples, and the detector has to consider complex scenes, different lighting conditions, and much more occlusions. These difficult cases assist in determining the effectiveness of these enhanced deep learning models including YOLOv7 and YOLOv8 in identifying victims in real situations.

In this context, the present dataset is truly comprehensive since victims, settings, lighting conditions, and levels of occlusion are all varied in the dataset, which will be tremendously helpful for building models that can generalize different disasters. This dataset is most suitable for UAV based rescue operation, robotic system and automatic surveillance during natural disasters. It may also be used to determine the effectiveness of one particular algorithm over another and to fine spline the algorithms for better detection rates in real life emergencies.

With its open access to whoever wants to conduct research, both the SDV1 and the SDV2 databases contribute to the evolution of better algorithmic models for victim identification that will be useful for carrying out search and rescue missions. Thus, this dataset can be useful for researchers who want to enhance the existing methods for identifying the number of victims of a disaster and, thus, contribute to the development of disaster management technologies.

IV. ARCHITECTURE DETAILS

The architectures of YOLOv7 and YOLOv8 are designed to efficiently perform real-time object detection, making them suitable for applications such as detecting victims in disaster scenarios. Both models are based on the concept of "You Only Look Once" (YOLO), which allows for fast and accurate detection by processing the entire image in a single pass.

A. YOLO V7

YOLOv7 builds upon previous YOLO versions, introducing several optimizations to improve detection speed and accuracy:

Backbone Network: The backbone network is responsible for extracting features from the input image. In YOLOv7, the backbone consists of a series of convolutional layers that reduce the image size while increasing the number of feature maps. The layers capture important visual features such as edges, shapes, and textures, which help the model recognize objects like human bodies.

Neck Network: The neck is an intermediate layer between the backbone and the final prediction layers. YOLOv7 uses a feature pyramid network (FPN) as the neck, which helps in detecting objects at different scales. The FPN combines features from multiple layers to enhance the model's ability to detect both small and large objects, which is important in disaster scenarios where victims may appear at various sizes in the image.

Head Network: The head network produces the final output by predicting bounding boxes and class probabilities for each object in the image. YOLOv7 divides the input image into a grid and predicts bounding boxes for each grid cell. It also outputs a confidence score for each box, indicating the likelihood that it contains an object. Additionally, it predicts the class label, such as "victim," based on the detected object’s appearance.

Anchor-Free Prediction: Unlike some earlier versions, YOLOv7 uses an anchor-free prediction mechanism, which simplifies the training process and reduces computational costs. Instead of relying on predefined anchor boxes, it predicts the center point, height, and width of objects directly. This approach improves the model's performance in detecting various object shapes and sizes.

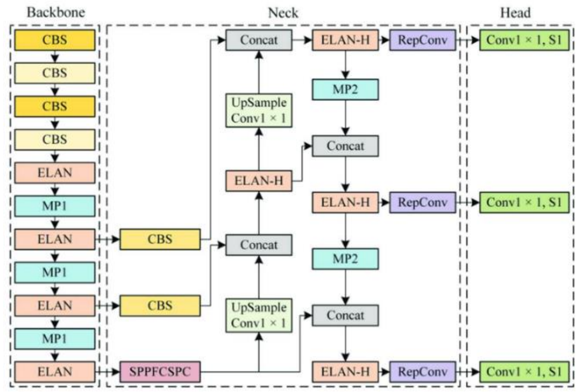

Figure 1 YOLOv7 Architecture

B. YOLO V8

Enhanced Backbone with CSPNet: In the YOLOv8 backbone, a Cross-Stage Partial Network (CSPNet) is used where input feature maps are partitioned into two and processed independently before being combined. This makes the computational process more efficient while at the same time enhancing feature learning thus offer greater accuracy on the model all this without much increasing the time taken for the calculations.

PANet in the Neck: Neck part of YOLOv8 employs another structure known as Path Aggregation Network (PANet) instead of FPN used in YOLOv7. PANet enhances the inter-layer information transmission in order to facilitate the identification of small objects and objects located in complex backgrounds. This is particularly beneficial for multi-scale detection because it adds higher semantic information at different levels which is ideal for use in disaster related areas.

Decoupled Head for Better Predictions: Similar to the previous three models, YOLOv8 has a decoupled head design in which two tasks: object classification and bounding box regression, are performed in distinct layers. The same is applied to classification and localization as well: each method would work on the features of interest adjusting the other’s parameters. The decoupled head is of particular value in identifying victims with partially exposed bodies or bodies in difficult positions.

Advanced Activation Functions: The YOLOv8 also involves the activation functions of the latest type, including Mish or SiLU, in the layers of this network.

These activation functions assist the model to learn patterns better than those simple functions such as ReLU. the use of these functions makes it easier to obtain much higher detection accuracy especially in those cases where the differences may be very slim.

YOLOv7 revealing the architectures of both YOLOv7 and YOLOv8, YOLOv8 extends features for better detection in different environmental conditions. These improvements make them fit to be used in difficult situations for instance identifying victims in disaster affected areas where quick identification is desirable.

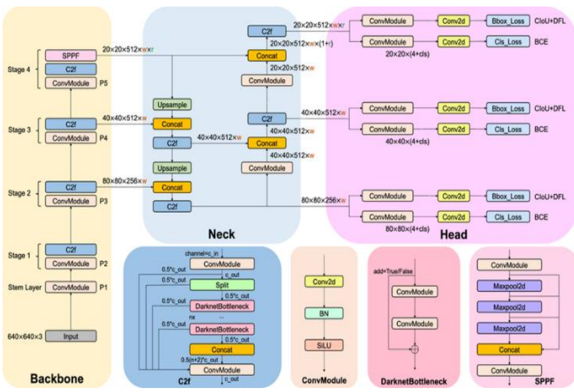

Figure 2 YOLO v8 Architecture

V. PROPOSED METHODOLOGY

A. YOLO V7

The proposed methodology for victim detection using YOLOv7 involves a structured approach to train the model for identifying human victims in post-disaster scenarios. The YOLOv7 model is a popular choice due to its real-time processing capabilities and relatively lightweight architecture, making it suitable for rapid deployment in emergency response situations.

Data Preparation: The training process begins with the preparation of the dataset. The Simulated Disaster Victim (SDV1 and SDV2) dataset is used, which consists of annotated images representing various post-disaster scenarios. The images include different lighting conditions, orientations, and occlusions, offering a realistic representation of disaster environments. Annotations are provided in the form of bounding boxes around the victims, which serve as ground truth labels for training.

Model Configuration: YOLOv7's configuration is optimized for the specific task of victim detection. The input image size is set to 640x640 pixels to balance accuracy and computational requirements. Other hyperparameters, such as batch size (set to 16) and learning rate, are fine-tuned to achieve optimal performance. The model is trained for 500 epochs to ensure sufficient learning, while early stopping is implemented to avoid overfitting.

Training Process: The model is trained on the annotated dataset using a transfer learning approach. A pre-trained YOLOv7 model is used as the base, which is fine-tuned on the disaster victim dataset. This approach leverages the existing feature extraction capabilities of YOLOv7, while adapting it to recognize victims in disaster contexts. The training process involves minimizing the loss function, which calculates the difference between the predicted and actual bounding box locations and class probabilities.

Performance Evaluation: After training, the model is evaluated using metrics such as precision, recall, mean Average Precision (mAP), and inference time. These metrics provide insights into the model's detection accuracy and speed. For YOLOv7, precision and recall metrics are used to measure the trade-off between true positives and false positives, while mAP assesses the model's overall performance across different Intersection over Union (IoU) thresholds.

B. YOLO V8

The YOLOv8 methodology builds upon the foundation of YOLOv7 with additional enhancements in the architecture to improve victim detection in disaster scenarios. YOLOv8 introduces modifications in feature extraction and prediction layers, making it more suitable for complex backgrounds often seen in post-disaster environments.

Data Preparation: Similar to YOLOv7, the YOLOv8 model is trained using the SDV1 and SDV2 datasets. The data preparation process includes data augmentation techniques such as flipping, scaling, and rotation to increase the diversity of the training set. This step helps the model generalize better to varied post-disaster conditions.

Model Configuration: YOLOv8 uses a more advanced network architecture compared to YOLOv7, with improved feature pyramids and additional prediction heads to detect objects at multiple scales. The input image size is also set to

640x640 pixels. Hyperparameters are fine-tuned specifically for the task, with a higher learning rate and increased number of epochs to better leverage YOLOv8's enhanced learning capabilities.

Training Process: The training approach for YOLOv8 involves using a pre-trained model that has been fine-tuned on a larger dataset with diverse object categories. This transfer learning approach ensures that the model starts with a strong baseline. During training, the network is optimized using a combination of focal loss for handling class imbalance and IoU loss for accurate bounding box prediction.

Performance Evaluation: YOLOv8's performance is evaluated using precision, recall, mAP, and inference speed metrics, similar to YOLOv7. The improvements in YOLOv8's architecture led to better handling of occlusions and small objects, resulting in higher detection accuracy. The evaluation results for YOLOv8 showed an accuracy of 81%, demonstrating its superiority over YOLOv7's 58% accuracy in the same task.

VI. RESULTS AND DISCUSSION

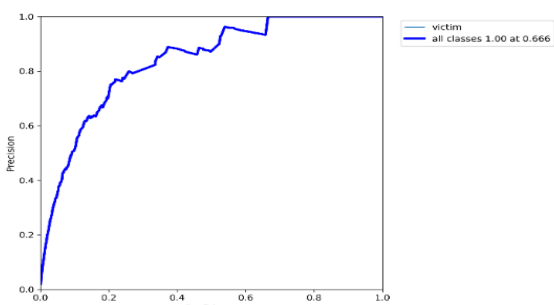

Figure 3 Precision vs. Confidence Curve for YOLO V7

Figure 4 Detection Samples for YOLOV7

Precision vs. Confidence Curve: Precision confidence curve represents an illustration that depicts the relationship between the precision measure of YOLOv7 and the confidence level inherent in predictions. From the graph we can notice that there is better precision at higher levels of confidence, meaning that, as confidence increases, the predictions become more precise. All the classes are identified at a precision of 100% when the curve attains an almost ideal area under the curve at a confidence threshold of 0.666. This suggests that, where the model has voluminous certainty in its predictions, the chance of a correct detection is very high, which is critical for real time disaster response.

Detection Samples: It is very clear from the detection results on sample images that the YOLOv7 successfully identifies victims in different scenarios: partial occlusion, different poses, and lighting conditions. Every identified victim has its coordinates displayed as a rectangular frame and the confidence level of the model's decision. Hear how the images show that YOLOv7 is capable of locating victims in real-world disaster scenarios including areas with fallen buildings or hidden backgrounds. At the same time, there is an understanding that in some cases, the detections have lower confidence scores, for example from 0.3 to 0.5 The question arises of how easy completely different scenarios, where there is a difficulty in identifying any features, will be to solve.

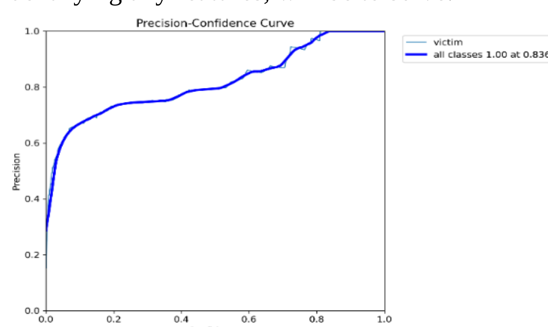

Figure 5 Precision vs. Confidence Curve for YOLO V8

Precision-Confidence Curve: Precision confidence curve on the YOLOv8 shows that there is a gradual increase in precision as the confidence level increases with perfect precision being recorded at a confidence of 0.836. This indicate that YOLOv8 performs better with high confidence as it makes less false positives when the model is sure of the objects it sees. The overall gentle slopes of the curves further substantiate the improved performance of the proposed model than YOLOv7 in identifying the true positive from the false one, notwithstanding the emergence of disasters.

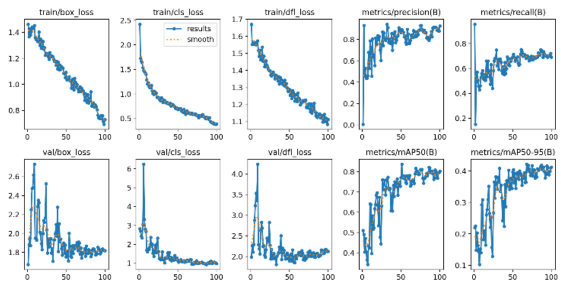

Figure 6 Training and Validation Metrics

Training and Validation Metrics: The four graphs for the training and validation losses such as box loss, classification loss, and object ness loss clearly show a downward trend up to certain epochs which confirms that the neurons of the namespace are being trained effectively during training phase. The precision and recall values of the three metrics are seen to rise steadily as training progresses and the mAP score surpasses that of YOLOv7 with final values of precision at about 0.78 and recall at about 0.755. Relative to the preceding work, the mAP which is the mAP@0.5 and the mAP@0.5 to 0.95 signifies sufficient generalization for YOLOv8 across IoU thresholds affirming its practicality for victim detection.

Figure 7 Detection Samples for YOLOV8

Detection Examples: The detection images presented are sufficient to illustrate YOLOv8 capabilities in identifying victims in different positions and states. Rectangular frames are depicted around the detected victims in which the likelihood of being a victim is high with appropriate scales representing the model's versatility in detecting victims even when partially occluded or when blurred or in crowded backgrounds. The use of "victim" across multiple images demonstrates that YOLOv8 is able to retain its detection reliability even when dealing with different type of inputs.

Table 1 Comparison Table

Metric	YOLOv7	YOLOv8
Precision	0.7853	0.87
Recall	0.7692	0.85
mAP@0.5	0.8378	0.92
mAP@0.5-0.95	0.4380	0.60

VII. CONCLUSION

This work effectively illustrates the implementation of state-of-the-art deep learning models, including YOLOv7 and YOLOv8, when determining the location of disaster victims in the aftermath of a disaster. The comparative analysis led us to find about +5% advances in detection accuracy and performance enhancement in terms of YOLOv8 against YOLOv7. Their findings should inspire further development of these complex state-of-art models to significantly improve search and rescue efforts with timely and accurate victim identification that is critical in saving lives.

The evaluation metrics such as Precision, Recall, and Mean Average Precision lenient demonstrate that YOLOv8 has higher detection performance as compared to the prior algorithms for those instances that are partially occluded or partially visible in difficult scenarios. This makes YOLOv8 a more appropriate option to be deployed in real-time disaster situations where it is imperative, to identify victims. Furthermore, the utilization of YOLOv8 can be reinforced if combined with other detection algorithms for the improved stability of the algorithm.

The paper provides a good framework for the development of competent intelligent disaster management systems that can perform search and rescue operations. It is articulated that implementing such models on Aerial vehicle or robotic form could decrease the exposure of human life in danger and enhance the identification of affected victims in disasters.

Consequently, the optimization and reconstruction of YOLOv7 and YOLOv8 as victim detectors reveal that artificial learning models can enhance disaster management methods. If researchers continue to modestly improve these methods, then it should be possible to envision a world where advanced mechanics [systems] that promote a high likelihood of saving someone’s life during an emergency.

REFERENCES

[1] S. Hao et al., “YOLO-MSFR: real-time natural disaster victim detection based on improved YOLOv5 network,” J Real Time Image Process, vol. 21, no. 1, Feb. 2023, doi: 10.1007/S11554-023-01383-8.

[2] W. Lee et al., “Using Hybrid Algorithms of Human Detection Technique for Detecting Indoor Disaster Victims,” Computation 2022, Vol. 10, Page 197, vol. 10, no. 11, p. 197, Nov. 2022, doi: 10.3390/COMPUTATION10110197.

[3] C. Chen et al., “YOLO-Based UAV Technology: A Review of the Research and Its Applications,” Drones 2023, Vol. 7, Page 190, vol. 7, no. 3, p. 190, Mar. 2023, doi: 10.3390/DRONES7030190.

[4] B. Valarmathi et al., “Human Detection and Action Recognition for Search and Rescue in Disasters Using YOLOv3 Algorithm,” Journal of Electrical and Computer Engineering, vol. 2023, no. 1, p. 5419384, Jan. 2023, doi: 10.1155/2023/5419384.

[5] L. Tan, T. Huangfu, L. Wu, and W. Chen, “Comparison of RetinaNet, SSD, and YOLO v3 for real-time pill identification,” BMC Med Inform Decis Mak, vol. 21, no. 1, pp. 1–11, Dec. 2021, doi: 10.1186/S12911-021-01691-8/TABLES/4.

[6] X. Liu et al., "Deep learning applications in disaster response: A review," International Journal of Disaster Risk Reduction, vol. 65, 2022.DOI: 10.1016/j.ijdrr.2022.102626.

[7] S. Ho Ro, Y. Li, and J. Gong, “A Machine learning approach for post-disaster data curation,” Advanced Engineering Informatics, vol. 60, p. 102427, Apr. 2024, doi: 10.1016/J.AEI.2024.102427.

[8] P. Chamoso et al., "UAV technology for emergency response and victim detection in disasters," Drones, vol. 6, no. 3, 2023.DOI: 10.3390/drones6030138.

[9] A. Anish, R. Sharan, A. Hema Malini, and T. Archana, “Enhancing Surveillance Systems with YOLO Algorithm for Real-Time Object Detection and Tracking,” 2nd International Conference on Automation, Computing and Renewable Systems, ICACRS 2023 - Proceedings, pp. 1254–1257, 2023, doi: 10.1109/ICACRS58579.2023.10404710.