Development of a Drone-Based Rescue Platform with Intelligent Human Detection and Multi-Payload Delivery Mechanism

Tran Quang Thuan
Faculty of Mechanical Engineering
Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Ho Chi Minh City
Vietnam

Van Thi Ngoc Han
Faculty of Mechanical Engineering
Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Ho Chi Minh City
Vietnam

Tra Quyen
IEC Quang Ngai International Education City Integrated School, Quang Ngai
Vietnam

Dang Phuoc Thinh
Faculty of Mechanical Engineering
Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Ho Chi Minh City
Vietnam

Phi Thanh Dat
Faculty of Mechanical Engineering
Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Ho Chi Minh City
Vietnam

Dang Anh Duy
Faculty of Mechanical Engineering
Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, Ho Chi Minh City
Vietnam

Floods often create emergencies in which access to victims is severely constrained by fast-moving water, deep terrain, and hazardous obstacles. Traditional rescue operations in such conditions are typically slow, resource-intensive, and expose rescuers to significant risks, which limits the efficiency of emergency response. This paper presents the conceptual design of an AI-enabled UAV platform for rapid rescue in flood conditions, particularly in hazardous and inaccessible areas. The system integrates a custom, reconfigurable payload mechanism comprising four articulated arms mounted beneath the drone, each with six degrees of freedom driven by RC servomotors. This configuration affords high adaptability, enabling secure transport and active release of diverse rescue supplies (e.g., medical boxes, essential goods, life buoys). In parallel, a deep learning–based human detection model built on YOLOv12 is trained on aerial images captured in flood scenarios to rapidly detect and localize victims from the UAV perspective. The detection outputs provide immediate target cues, allowing the drone to quickly acquire the victim's location and deliver the payload directly within reach, thereby reducing the cognitive and physical burden on both the operator and the victim. Mechanical simulations of the four-arm mechanism across multiple payload types, together with training outcomes of the YOLOv12 detector, indicate the feasibility of the proposed approach and its suitability for time-critical field conditions. Overall, the combination of AI-assisted victim localization and a flexible, active payload mechanism reduces human intervention, improves delivery accuracy, and supports a smart, efficient, and readily deployable solution for disaster-response operations.

Keywords: *Search and rescue, Emergency UAV system, Victim detection, Yolov12, Autonomous payload release.*

1. INTRODUCTION

Floods and storms are among the most destructive natural disasters, often causing severe loss of life, extensive property damage, and long-term socio-economic disruption. Beyond their immediate impact, these disasters create emergencies that demand rapid and effective rescue operations. However, conducting rescue missions under such extreme conditions poses numerous challenges and exposes rescuers to significant risks. In flood scenarios, strong water currents, deeply submerged terrain, and completely inundated infrastructure severely limit the ability of traditional rescue methods, such as motorboats, canoes, or ground vehicles, to reach affected individuals promptly. The presence of hazardous obstacles such as fallen trees, collapsed buildings, downed power lines, and floating debris further increases the danger for both victims and rescue teams. In particularly severe cases, adverse weather conditions, poor visibility, and high winds complicate navigation and maneuvering, heightening the risk of accidents or injuries to rescue personnel. Furthermore, in many flood-affected areas, transportation infrastructure and communication networks are damaged or destroyed, delaying the deployment of rescue resources and equipment. For remote or completely isolated regions, reaching victims within the "golden time", the critical period immediately following the disaster, becomes especially difficult, significantly reducing the chances of survival.

Unmanned Aerial Vehicles (UAVs), or drones, have emerged as transformative assets in disaster response due to their ability to operate over hazardous terrain, bypass destroyed infrastructure, and reach otherwise inaccessible locations within minimal time [1-5]. Their integration into search and rescue (SAR) operations and relief logistics has been driven by the urgent need for rapid victim detection, situation assessment, and delivery of essential supplies during emergencies [6-10].

Numerous studies have demonstrated the operational advantages of drones in locating victims across diverse disaster environments. For maritime emergencies, Claesson et al. [11] showed that drones reduced the

Received: July 2025, Accepted: August 2025
Correspondence to: Dang Anh Duy, Faculty of Mechanical Engineering, Ho Chi Minh City University (HCMUT), VNU-HCM, Ho Chi Minh City, Vietnam
E-mail: aduy@hcmut.edu.vn
doi: 10.5937/fme2504565Q

median search time for drowning victims from 4:34 minutes to 0:47 minutes, outperforming trained life–guard teams. In mountainous snow-covered terrain, Ka–raca et al. [12] demonstrated that the drone–snowmobile technique shortened the victim arrival time from 57.3 minutes to 8.9 minutes and expanded search coverage more than twenty-fold. Forest SAR has also benefited from drone innovations; Schedl et al. [13] employed airborne optical sectioning and thermal imaging to achieve 100% detection precision under dense canopy conditions. Urban and occlusion-challenged environ–ments have been addressed by Russell Bernal et al. [14], who adapted detection models with psychophysical loss functions to improve performance under high occlusion and low target resolution. Furthermore, Mishra et al. [15] developed a SAR-specific human action detection dataset, achieving 0.98 mAP with a deep learning model, while Meenakshi et al. [16] integrated reinfor–cement learning–based navigation with LiDAR, ther–mal, and RGB sensing to reach 95.6% victim detection accuracy in simulated disaster zones. Complementing visual and thermal sensing, Albanese et al. [17] introduced SARDO, a drone-based mobile phone localization system, locating devices within a few tens of meters in about 3 minutes per user without infrastructure support.

UAVs are equally effective in the distribution of essential items to disaster-stricken or isolated areas, where conventional transport is hindered. Hii et al. [18] confirmed that drone transport preserves the quality of sensitive medicines, such as insulin, even under varying temperatures and vibrations. In the context of traffic accident emergencies, Kristensen et al. [19] proposed the Rescue Emergency Drone (RED) for rapid site assessment and delivery of first aid, reducing response times compared to conventional ground units. Large-scale logistics optimization has also been explored: Liu and You's DroneGo system applied integer prog–ramming and Dijkstra's algorithm to determine the shortest medical delivery paths post-hurricane in Puerto Rico, significantly improving efficiency [20]; Rabta et al. [21] developed a mixed-integer linear programming model for last-mile relief delivery, considering payload and energy constraints alongside recharging infra–structure. Broader reviews, such as Patel et al. [22], highlight UAVs' role in bridging access gaps for hospitals, NGOs, and underserved communities in rural or rugged terrains.

Recent developments in AI, communications, and localization are expanding the operational scope of drones in SAR [23]. Ma et al. [24] proposed OWRT-DETR, a novel Transformer-based detection network tailored for small-object detection in open-water envi–ronments. By integrating modules such as cross-scale feature pyramid interaction and small-object enhan–cement, their method significantly improves feature representation and suppresses background confusion, achieving superior accuracy on public UAV datasets. Liu et al. [25] introduced the concept of an Internet of UAVs for smart cities, focusing on post-disaster search and rescue. They developed a mathematical model of UAV-based SAR missions with life-detection radars and applied an Improved Multi-Verse Optimizer (IMVO) to optimize 3D path planning. Simulation results demonstrated effective search coverage, reduced mission time, and increased survivor discovery rates. Similarly, Fu et al. [26] presented DLSW-YOLOv8n, a UAV-based detection framework that enhances small maritime target recognition. By combining deformable large kernel convolution, SPD-Conv, and an improved loss function, their approach boosts small target locali–zation and robustness, improving detection accuracy by 13.1% compared to YOLOv8n. Liu et al. [27] addressed the challenge of UAV perception under adverse weather by proposing WRRT-DETR and introducing the large-scale AWOD dataset. With modules that integrate local convolution, global attention, and frequency–spatial augmentation, their model demonstrated robustness against fog, glare, and low-light conditions, outper–forming state-of-the-art methods in maritime UAV detection. Velavan et al. [28] integrated YOLOv8-based survivor detection with onboard edge AI (Raspberry Pi 5 with Hailo AI kit) and 5G connectivity, enabling low-latency, high-reliability real-time streaming and control. Alnoman et al. [29] reviewed emerging AI and 6G-enabled localization solutions, including THz commu–nications, satellite-based non-terrestrial networks, and reconfigurable intelligent surfaces, which promise sub-meter victim localization in disaster contexts.

While drones have demonstrated substantial potential in search, rescue, and relief supply delivery, existing applications in flood and storm rescue scenarios remain limited. Most systems still rely heavily on manual control for navigation and payload release. In extreme flood conditions or when operating over long distances, human operators may struggle to maintain precise control due to signal delays, environmental hazards, or limited visibility. Furthermore, the payload delivery mechanisms commonly employed in such systems are often simple and fixed in structure, restricting their ability to handle diverse emergency situations. This limitation becomes particularly critical when victims are physically weakened or mentally distressed, making it difficult for them to retrieve the delivered items from the drone without direct assistance.

To address these challenges, this paper introduces a novel, reconfigurable four-robot arm payload carrying and release mechanism designed to enhance the adap–tability and effectiveness of drone-based rescue opera–tions. Each arm features six degrees of freedom, driven by individual RC servo motors, allowing the mechanism to dynamically adjust its configuration to accommodate a variety of rescue items such as medical kits, essential supplies, and life buoys. By employing YOLOv12 for victim detection, the drone can rapidly acquire the target location and easily deliver the payload directly within the victim's reach. Unlike prior studies that mainly focused on either UAV-based victim detection or simple payload attachments, this research uniquely integrates an intelligent detection framework with a flexible multi-arm payload mechanism, thereby minimizing human intervention, improving delivery accuracy, and incre–asing the overall success rate of rescue missions in time-critical and hazardous environments.

The remainder of this paper is organized as follows: Section 2 presents the design and development of the proposed UAV-based rescue system, detailing both the AI-powered victim detection module and the reconfigurable payload delivery mechanism. Section 3 describes the experimental setup and evaluation procedures. Section 4 discusses the results, and Section 5 concludes the paper with key findings and future research directions

2. MECHANICAL DESIGN OF RECONFIGURABLE PAYLOAD MECHANISM

The proposed payload delivery system is a modular unit that can be mounted or detached from various un–manned aerial vehicles (UAVs) without requiring structural modification to the host platform. It is built on a rigid square-shaped base plate, serving as the central interface between the drone and the payload handling mechanism. Four independent 6-DOF robotic arms are symmetrically attached to the corners of the base plate, with all joint axes arranged in parallel orientations to simplify kinematics, reduce control complexity, and enable precise planar positioning of the end segments relative to the payload.

Each arm can operate independently to hold an item or be coordinated with the others to firmly secure larger or irregularly shaped objects. This configuration also allows the arms to fold compactly when not in operation, minimizing aerodynamic drag and serving as stable landing legs during take-off and landing.

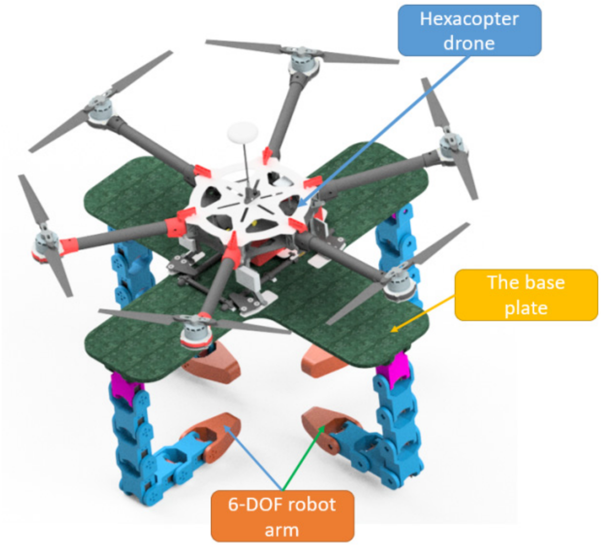

Figure 1. Complete Design of the Drone-Based Rescue Delivery System

The mechanical architecture is optimized for lightweight construction while maintaining sufficient rigidity to support various rescue payloads, including medical kits, emergency supplies, and life-saving flotation devices. The modular approach ensures adaptability to different drone sizes and configurations, making the system suitable for diverse rescue scenarios. Figure 1 shows the complete Drone-Based Rescue Delivery System, integrating a hexacopter with the four-arm payload mechanism for rapid and precise delivery in emergency operations. In this design, the four arms are folded inward at the middle joints, forming the landing legs of the drone when it touches down.

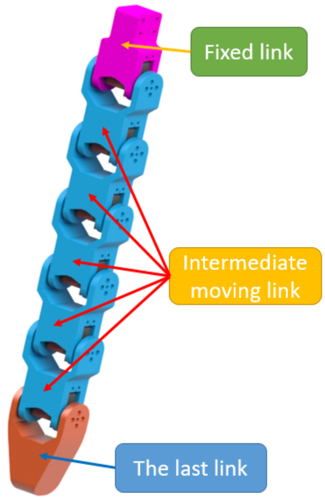

Figure 2. Structure of a single robotic arm

Each robotic arm in the payload mechanism is constructed with a fixed link and six serially connected moving links, as illustrated in Figure 2. The fixed link (highlighted in green) is rigidly attached to the base plate of the drone. The six moving links include five identical intermediate links (highlighted in blue) and one terminal link (highlighted in orange) that interfaces with the payload. All joints are actuated by RC Servo HiWonder HPS-2027 with parallel-axis alignment, enabling smooth, coordinated movement. The overall length of each arm reaches 570 mm when fully extended, providing sufficient reach to handle objects positioned below or around the drone. This configuration provides six degrees of freedom for each arm, allowing independent or cooperative operation to securely clamp rescue items.

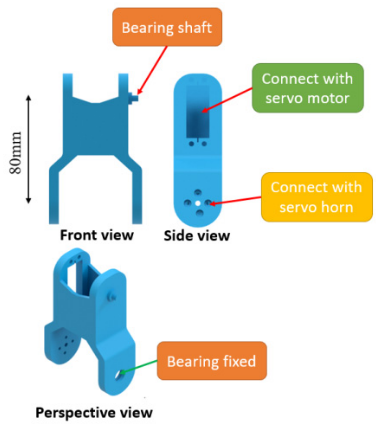

Figure 3. Design of the intermediate moving link of the 6-DOF robotic arm

The design of the intermediate moving link is illus–trated in Figure 3. One end features a rectangular slot for mounting the servo motor, ensuring precise align–ment and stable attachment. The opposite end is equi–pped with multiple holes for connecting to the servo horn of the subsequent link, allowing for secure trans–mission of motion between segments. The link incor–porates a bearing shaft on one side and a fixed bearing slot on the other, enabling smooth rotational movement while reducing mechanical wear. Each intermediate link has a total length of 80 mm, contributing to the overall reach and flexibility of the robotic arm.

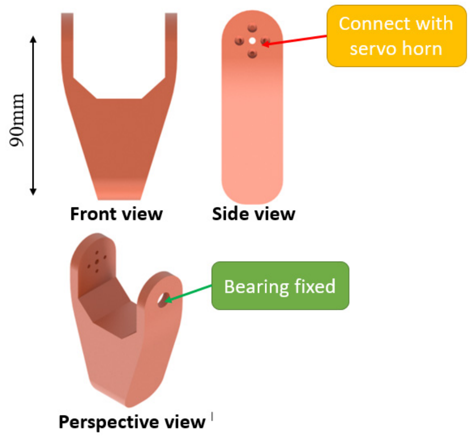

Figure 4. Design of the last link in the 6-DOF robotic arm

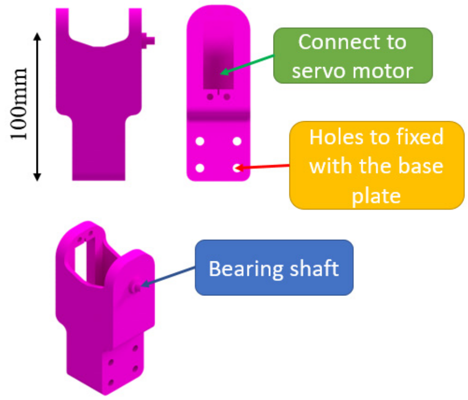

Figure 5. Design of the first link in the 6-DOF robotic arm

Figures 4 and 5 illustrate the designs of the first and last links of the 6-DOF robotic arm, respectively. The first link (Figure 5) has a total length of 100 mm and features four mounting holes at one end for securing it to the base plate, while the opposite end is designed to connect with the intermediate moving link via a servo motor interface.

The last link (Figure 4) measures 90 mm in length, with one end connected to the intermediate moving link through a servo horn, and the other end left free to directly interact with the payload or clamp in coordination with other arms. Both link types are equipped with bearing housings to ensure smooth joint rotation and to withstand operational loads during payload handling.

3. HUMAN DETECTION FROM AERIAL VIEW

2.1 Yolov12 Architecture

The YOLOv12 architecture is designed to achieve high detection accuracy while maintaining real-time inference speed, making it suitable for computationally constrained platforms such as drones. As shown in Figure 6, the network is composed of three main modules: Backbone, Neck, and Head.

a) Backbone

The backbone is responsible for feature extraction from the input image (640×640 in this configuration). YOLOv12 employs an enhanced version of the R-ELAN (Re-parameterized Efficient Layer Aggregation Network) structure, which improves feature repre–sentation by aggregating information from multiple receptive fields without significantly increasing compu–tational cost. Two key innovations in the backbone are:

- A2C2f module (Area Attention + Convolution): Introduces spatial area attention to emphasize informative regions while suppressing background noise, particularly useful for aerial-view human detection where targets are small and partially occluded.

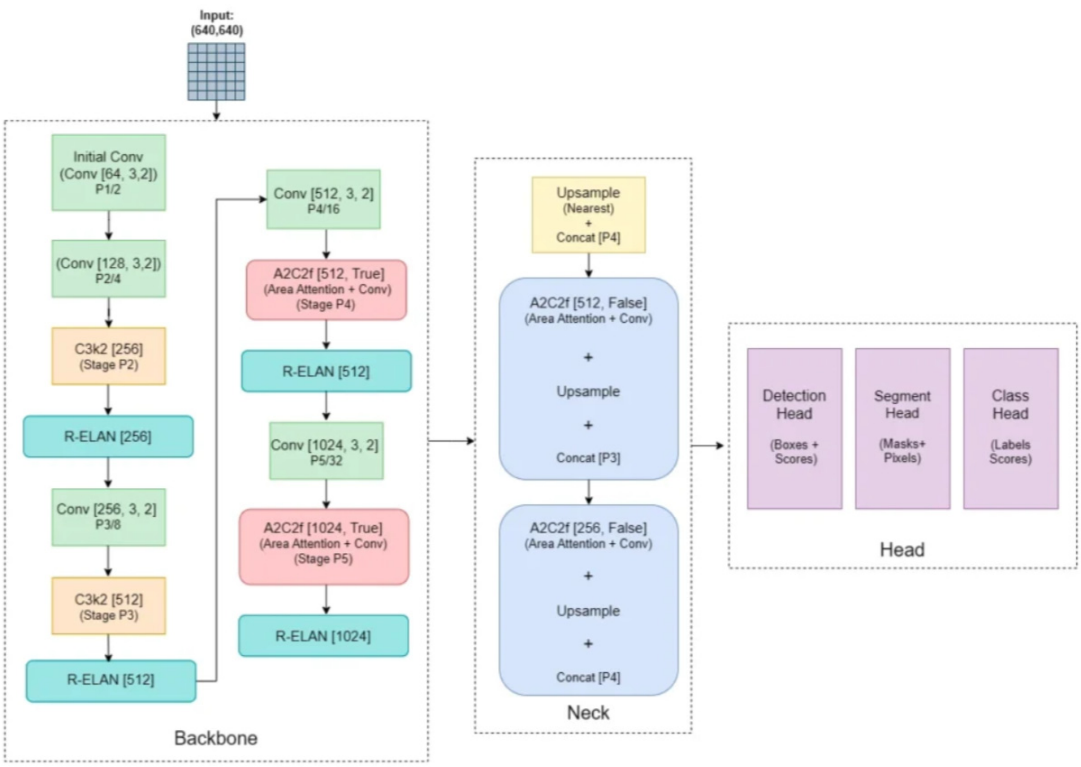

Figure 3. Fig. 6. The network architecture of YOLOv12

- Stage-wise feature scaling: The backbone prog–ressively increases the number of channels (from 64 to 1024) through convolution and C3k2 blocks, enabling multi-scale feature learning

b) Neck

The neck performs multi-scale feature fusion, combining high-resolution spatial features from earlier stages with semantically rich features from deeper layers. YOLOv12 integrates A2C2f modules with False attention mode in the neck, focusing on context refinement rather than re-weighting spatial regions. Feature maps from different pyramid levels (P3, P4, P5) are merged using up-sampling and concatenation, enhancing the detection of small objects.

c) Head

The head is a unified multi-task output layer comp–rising:

- Detection head: Predicts bounding box coordinates, objectness scores, and class probabilities.
- Segmentation head: Generates pixel-level masks for instance segmentation tasks (optional).
- Classification head: Outputs class scores for image-level classification.

This multi-head design allows YOLOv12 to be adapted for different vision tasks without retraining the entire backbone and neck.

Compared to YOLOv8 and YOLOv9, YOLOv12 introduces the Area Attention mechanism (A2C2f) and enhanced R-ELAN modules, which:

- Improve the detection of small and partially occluded objects.
- Provide better feature aggregation with minimal latency increase.
- Maintain high accuracy on complex backgrounds, as encountered in aerial rescue scenarios.

These architectural improvements make YOLOv12 well-suited for human detection from aerial views, where real-time performance and robustness against background clutter are critical.

2.2 Training Strategy

In this study, the YOLOv12 model is trained for the task of human detection from aerial imagery captured by drone-mounted cameras. The objective is to enable rapid victim localization during search and rescue operations, particularly in disaster scenarios such as floods and storms.

A custom dataset was developed, consisting of over 7,00 high-resolution aerial images collected from diverse environments, including urban areas, rural landscapes, and disaster-affected regions. This dataset contains numerous instances of humans in varying postures, scales, and occlusion conditions, with a significant subset depicting flood and storm victims. The diversity in background scenes and lighting conditions ensures that the model learns robust feature representations, making it applicable to real-world rescue missions where environmental variability is high.

All images were manually annotated using the Roboflow platform, with bounding boxes precisely drawn around each human instance visible from the aerial perspective. The annotated dataset was then randomly split into 75% for training and 25% for validation to ensure reliable performance evaluation while preventing overfitting.

Figure 7. Example of a flood rescue scenario included in the custom aerial-view dataset

To accelerate convergence and leverage prior knowledge, transfer learning was applied by initializing the YOLOv12 model with pre-trained weights from the MS COCO dataset. This allowed the network to retain general object detection capabilities while fine-tuning on the custom aerial-view dataset, improving perfor–mance on small and partially occluded human targets. Training was conducted on Google Colab with GPU acceleration to achieve a balance between computa–tional efficiency and accuracy.

The YOLOv12 model retained its default archi–tecture without altering the core layers. The pre-trained weights from the COCO dataset were loaded before fine-tuning on the custom dataset. The initial learning rate was set to 0.01, with a cosine decay scheduler applied to gradually reduce the rate, ensuring stable convergence. The SGD optimizer was used with a momentum of 0.937 and a weight decay of 0.0005 to minimize overfitting.

Training was performed for 100 epochs with a batch size of 16 and an input image size resized to 640 × 640 pixels. Data augmentation techniques, including random scaling, horizontal flipping, mosaic augmentation, and color jitter, were applied to enhance model robustness in diverse aerial-view conditions. The loss function combined bounding box regression loss, objectness loss, and classification loss for end-to-end optimization.

After training, the optimized model can be deployed on embedded platforms such as Raspberry Pi via an AI edge computing framework for real-time inference in field operations. The on-board camera streams video to the AI edge server, where OpenCV handles image pre-processing before feeding the frames into the YOLOv12 model for detection. The inference results are validated and evaluated using predefined performance metrics, ensuring that the deployed system maintains both accuracy and speed in real rescue missions.

4. RESULTS AND DISCUSSION

Simulation of the Payload-Carrying Mechanism

The proposed rescue payload delivery mechanism was simulated to evaluate its capability in carrying and releasing various types of emergency supplies. As illustrated in Figure 8, the system is designed to transport multiple items, including an instant noodle carton, a medical supply box, and life buoys of different sizes. The instant noodle carton measures 40.7 × 31 × 24 cm, while the medical box has dimensions of 38 × 27 × 22 cm. The life buoys come in multiple sizes, with the largest one having an outer diameter of 760 mm and an inner diameter of 460 mm, weighing approximately 4.3 kg. The mechanism's arms can adjust their grip to secure these payloads during flight and release them accurately at the target location. This simulation phase focused on verifying that the mechanical design can accommodate payloads of varying shapes and sizes while maintaining stability for drone operations in real-world rescue scenarios.

Figure 8. Various Payloads for Drone-Based Rescue Missions: Food Supply Box, Medical Kit, and Different Sizes of Life Buoys

In the case of carrying instant noodle boxes or medi–cal kit boxes, the four arms operate in a coordinated manner to securely hold the payload. The terminal link (Joint 7) is positioned parallel to the base of the carried object, providing vertical support. One or two of the subsequent links (e.g., Joint 5 and Joint 6) are pressed against the object's sides to ensure a firm grip. The remaining joints are adjusted to keep the payload at the geometric center beneath the drone, thereby main–taining balance and preventing asymmetric load distri–bution.

Figure 9 illustrates the simulation of carrying a single box. In this configuration, all four arms coor–dinate to clamp and support the box at its bottom and sides, ensuring stability during aerial transport. The corresponding joint configurations for this case are presented in Table 1. Figure 10 shows the simulation scenario for carrying two boxes stacked horizontally. The arms are adjusted to simultaneously grip and stabilize both boxes, ensuring that the load is securely held without slipping or tilting. The joint angles for this configuration are provided in Table 2.

Table 1. Joint angles of the four arms for the single-box carrying case

No.	Joint 1	Joint 2	Joint 3	Joint 4	Joint 5	Joint 6
Arm 1	34.87^0	0^0	0^0	0^0	-34.87^0	90^0
Arm 2	9.58^0	39.69^0	0^0	-49.27^0	0^0	90^0
Arm 3	34.87^0	0^0	0^0	0^0	-34.87^0	90^0
Arm 4	9.58^0	39.69^0	0^0	-49.27^0	0^0	90^0

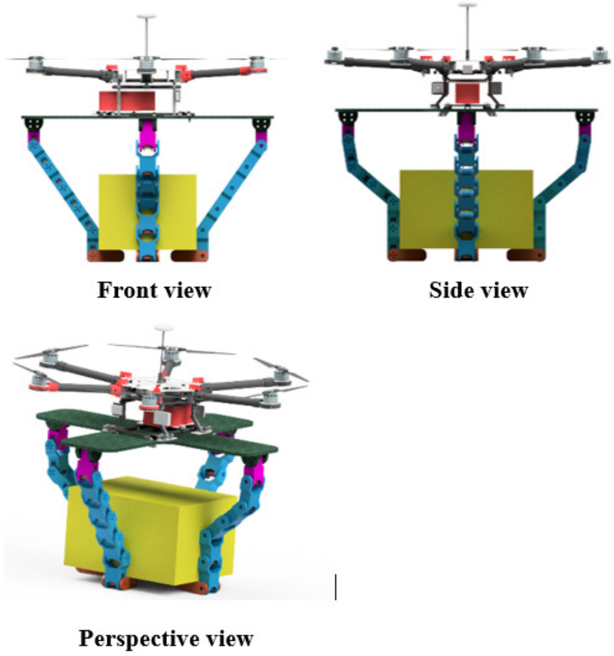

Figure 9. Simulation of the drone arm mechanism carrying one box.

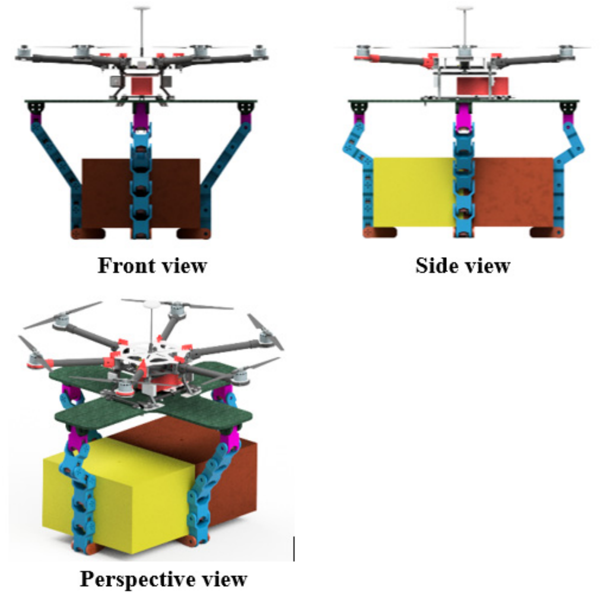

Figure 10. Simulation of the drone arm mechanism carrying two boxes.

Table 2. Joint angles of the four arms for the dual-box carrying case

No.	Joint 1	Joint 2	Joint 3	Joint 4	Joint 5	Joint 6
Arm 1	-27.06^0	82.35^0	55.29^0	0^0	0^0	90^0
Arm 2	9.58^0	39.69^0	0^0	-49.27^0	0^0	90^0
Arm 3	-27.06^0	82.35^0	55.29^0	0^0	0^0	90^0
Arm 4	9.58^0	39.69^0	0^0	-49.27^0	0^0	90^0

For carrying lifebuoys, the arms operate indepen–dently rather than in coordination. Each arm is bent into a hook-shaped configuration that wraps tightly around the lifebuoy, preventing it from swaying or shifting during flight (Figure 11). This secure grip ensures stability even in windy or turbulent conditions. The specific joint angles for forming the hook shape are provided in Table 3.

An important advantage of the proposed mechanism is that once the drone approaches the target location, the arms can actively release the rescue items directly into the hands of the victim. This capability is particularly valuable when the victim is physically and mentally exhausted and may not be able to detach the payload if conventional passive delivery methods were used.

Figure 11. Simulation of the drone arm mechanism carrying a lifebuoy in a hook-shaped configuration.

Table 3. Joint angles of the four arms for the hook-shaped configuration

Lifebuoys	Joint 1	Joint 2	Joint 3	Joint 4	Joint 5	Joint 6
Big	0^0	0^0	56^0	37^0	56^0	37^0
Small	0^0	0^0	66.3^0	44.68^0	66.3^0	44.68^0

The results of the YOLOv12 detection model

The training and validation curves are illustrated in Figure 12. The box regression loss, classification loss, and distribution focal loss (DFL) all steadily decreased and converged after approximately 100 epochs, indi–cating stable learning. The validation losses followed a similar trend, demonstrating that the model was not overfitting.

In terms of detection metrics, both precision and recall reached near-optimal levels, with values of 0.984 and 1.000, respectively. The mean Average Precision at IoU threshold 0.5 (mAP@50) achieved 0.995, reflecting the model's excellent capability to localize and classify human instances in challenging flood scenarios. Meanwhile, the stricter metric mAP@50–95 attained a value of 0.652, which indicates reasonable gene–ralization across multiple IoU thresholds but also suggests potential improvement when detecting under high-precision bounding box alignment.

The quantitative results on the validation set are summarized in Table 4. The YOLOv12 model achieved near-perfect performance on the test dataset, surpassing conventional baselines in both accuracy and reliability.

Table 4. Performance of yolov12 on human detection dataset

Precision (P)	Recall (R)	mAP@50	mAP@50–95
0.984	1.00	0.995	0.652

Additionally, the average inference speed was ~5.7 ms per image, including preprocessing, model infe–rence, and post-processing. This indicates that the trained YOLOv12 detector can be deployed on-board a UAV in real-time conditions, which is crucial for time-sensitive flood rescue operations.

The proposed human detection model was evaluated under various real-world flood scenarios, including close-range, far-range, and crowded conditions. The results demonstrate that the model consistently achieved accurate detections across different contexts.

For close-range detection, as illustrated in Fig. 13, individuals located on rooftops or partially submerged in floodwater were successfully detected with high confidence. The bounding boxes tightly aligned with the human contours, proving the model's robustness against variations in posture and partial occlusion caused by water or surrounding objects.

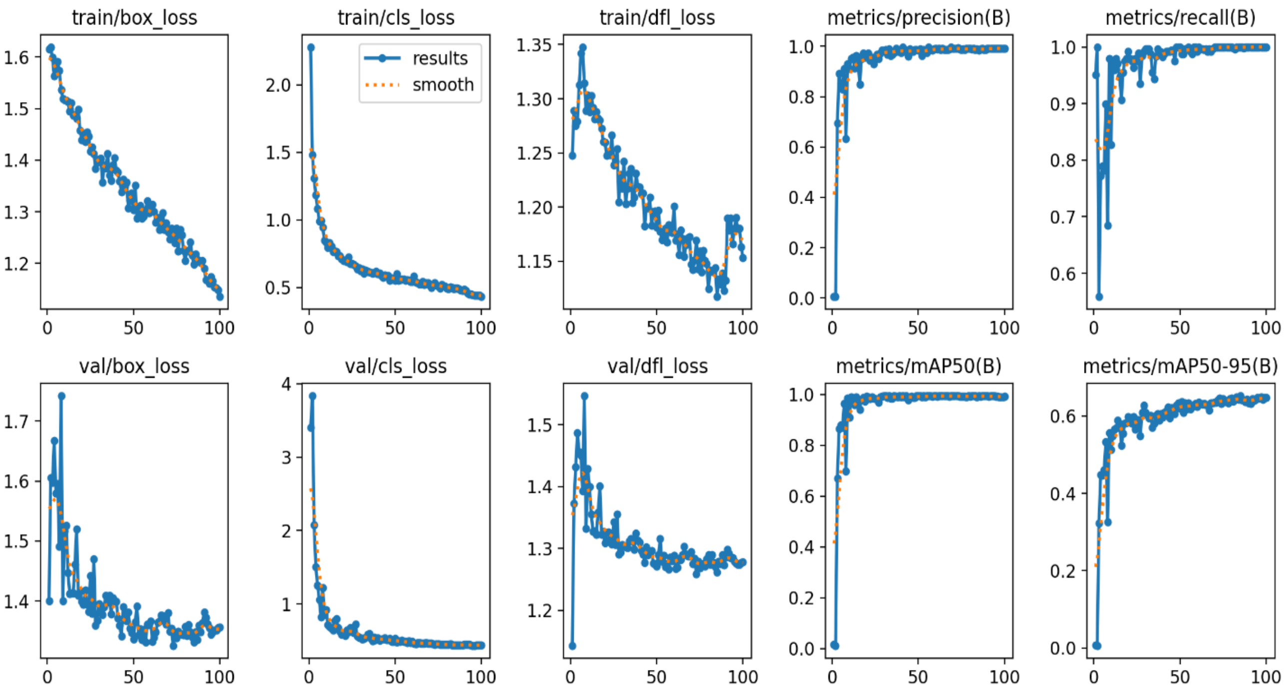

Figure 12. Training and validation curves results.

Figure 13. Close-range detection: accurate identification of individuals in proximity under flood conditions.

In far-range scenarios, shown in Fig. 14, the model was able to detect both single individuals and groups of people at a distance. Even in challenging conditions where people appeared relatively small in the image, the detection performance remained reliable. However, in cases where individuals were standing close to each other, the model occasionally grouped them into a single bounding box, which suggests a limitation when distinguishing between closely spaced targets in low-resolution regions.

a. Far-range detection with multi-person.

b. Far-range detection with one-person.

Figure 14. Far-range detection results in flood scenarios.

Figure 15. Crowded Scenario: Multi-Person Detection in a Flooded Environment.

For crowded environments, as seen in Fig. 15, the model maintained its ability to localize multiple persons simultaneously, even in complex backgrounds with rescue boats and trees. This indicates that the trained model generalizes well to high-density situations and is capable of supporting rescue operations where rapid detection of multiple people is critical.

Overall, these results confirm that the developed model is effective in handling diverse real-world scenarios. While minor challenges remain in distinguishing individuals in very close proximity, the detection accuracy is sufficiently high to ensure reliable operation in emergency flood-rescue missions.

The integration of the YOLOv12-based detection model with the flexible multi-payload delivery mechanism creates a synergistic rescue platform: YOLOv12 enables rapid and accurate victim detection, ensuring the drone approaches the target quickly, while the delivery mechanism allows the system to transport essential supplies and release them precisely at the victim's location. Although this study has not yet been validated through real-world experiments, the integration of YOLOv12-based human detection and the proposed multi-payload delivery mechanism demonstrates strong feasibility. The experimental results of YOLOv12 on diverse flood scenarios confirm its robustness in detecting victims under different conditions, while the structural design of the flexible carrying arms shows the potential to securely transport and release various rescue supplies. Therefore, this initial research provides a solid foundation for future implementation and real-world testing, where the system can be further optimized to meet practical rescue requirements.

5. CONCLUSION

This study presented the development of a novel drone-based rescue platform that integrates an intelligent human detection model (YOLOv12) with a flexible multi-payload delivery mechanism. The proposed system addresses two key challenges in flood rescue scenarios: (1) the rapid and reliable detection of victims in complex environments, and (2) the secure transportation and release of essential supplies, including life buoys, food packages, and medical kits.

Through simulation, the flexible arm mechanism demonstrated the ability to adapt to different payload shapes and sizes, ensuring stable transportation. The arms could operate cooperatively to grasp rigid boxes or independently to hook and secure circular life buoys, preventing oscillations during flight. Additionally, YOLOv12 achieved high detection accuracy with a mean Average Precision (mAP50) of 99.5%, proving effective across near-range, far-range, and crowded scenarios. This ensures that victims can be quickly identified and rescue operations can be executed with minimal delays.

Although this work remains at a preliminary stage without real-world testing, the combination of robust computer vision with an adaptive mechanical design highlights the strong feasibility of the proposed approach. The research establishes a foundation for future implementation, where hardware integration and field experiments will be conducted to validate system performance under practical conditions. Furthermore, future work will focus on optimizing the arm's control algorithms, improving payload stability under turbulent environments, and enhancing the robustness of the detection model in adverse weather conditions.

ACKNOWLEDGMENT

We acknowledge Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for supporting this study.

REFERENCES

[1] S. Yeom, "Thermal Image Tracking for Search and Rescue Missions with a Drone," Drones, Vol. 8, No. 2, 2024. https://doi.org/10.3390/drones8020053

[2] C.K. Long U and N. K. Toan, "Emergency UAV Delivery Framework: A Hybrid Approach to GPS Navigation and Visual Landmark Detection," International Journal of Mechanical Engineering and Robotics Research, Vol. 14, No. 4, pp. 374-383, 2025. doi: 10.18178/ijmerr.14.4.374-383

[3] V. Garg et al., "Drones in last-mile delivery: A systematic review on efficiency, accessibility, and sustainability," Transportation Research Part D: Transport and Environment, vol. 123, 103831, 2023. doi: 10.1016/j.trd.2023.103831

[4] K. M, O. M, P. B, R. K V and S. K. S, "Unmanned Aerial Systems in Search and Rescue: A Comprehensive Review and Future Directions," 2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI), Lalitpur, Nepal, 2024, pp. 15-18, doi: 10.1109/ICMCSI61536.2024.00008.

[5] M.D. Milenković-Babić, V.D. Stefanović-Gobeljić, B.G. Ostojić, B.Z. Dovatov, B.B. Ostić, M.B. Gligorijević, "The Rescue System Implemetation in the Medium Range UAV," FME Transactions, Vol. 53, pp. 482-489, 2025.

[6] P.R. Harris, S.E. Miller & J.D. Telford, "Drones in disaster management: A review of applications and recommendations for future research," International Journal of Disaster Risk Reduction, Vol. 50, 101682, 2020.

[7] S.A. Ghauri, M. Sarfraz, R.A. Qamar, M.F. Sohail, S.A. Khan, "A Review of Multi-UAV Task Allocation Algorithms for a Search and Rescue Scenario," Journal of Sensor and Actuator Networks, Vol. 13, No. 5, 47. https://doi.org/10.3390/jsan13050047

[8] J. Xu et al., "YoloOW: A Spatial Scale Adaptive Real-Time Object Detection Neural Network for Open Water Search and Rescue From UAV Aerial Imagery," in IEEE Transactions on Geoscience and Remote Sensing, vol. 62, pp. 1-15, 2024, Art no. 5623115, doi: 10.1109/TGRS.2024.3395483.

[9] S. Qi, B. Lin, Y. Deng, X. Chen, Y. Fang, "Mini–mizing Maximum Latency of Task Offloading for Multi-UAV-Assisted Maritime Search and Rescue," in IEEE Transactions on Vehicular Technology, vol. 73, no. 9, pp. 13625-13638, Sept. 2024, doi: 10.1109/TVT.2024.3384570.

[10] M. Marin, P. Mirosavljević, "Analysis of the per–formance and kinematics of the movement of UAV," FME Transaction, Vol. 51, No 4, pp. 627-636, 2023

[11] A. Claesson, L. Svensson, P. Nordberg, M. Ringh, M. Rosenqvist, T. Djarv, J. Samuelsson, O. Hernborg, P. Dahlbom, A. Jansson, J. Hollenberg, "Drones may be used to save lives in out of hospital cardiac arrest due to drowning," Resuscitation, Vol. 114, pp. 152–156, 2017. https://doi.org/10.1016/j.resuscitation.2017.01.003

[12] Y. Karaca, M. Cicek, O. Tatli, A. Sahin, S. Pasli, M.F. Beser & S. Turedi, "The potential use of unmanned aircraft systems (drones) in mountain search and rescue operations," The American Jour–nal of Emergency Medicine, Vol. 36, No. 4, pp. 583–588, 2018. https://doi.org/10.1016/j.ajem. 2017.09.025

[13] D.C. Schedl, I. Kurmi, O. Bimber, "An autonomous drone for search and rescue in forests using airborne optical sectioning," Science Robotics, 2021. https://doi.org/10.1126/scirobotics.abg1188.

[14] A.M. Russell Bernal, J. Cleland-Huang & W. Scheirer, "Psych-Occlusion: Using Visual Psycho–physics for Aerial Detection of Occluded Persons During Search and Rescue," 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), pp. 3383–3395, 2025. https://doi.org/10.1109/WACV51458.2025.00344

[15] B. Mishra, D. Garg, P. Narang, V. Mishra, "Drone-surveillance for search and rescue in natural dis–aster," Computer Communications, Vol. 156, pp. 1–10, 2020. https://doi.org/10.1016/j.comcom.2020.03.012

[16] S. Meenakshi, R.K.P.K.P Pavaimalar, S. Ravi, N. Kumaran & S.B. Priya, "AI-Driven Autonomous Robots for Search and Rescue Operations in Disaster Zones," 2025 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), pp. 1–6, 2025. https://doi.org/10.1109/ICDSAAI65575.2025.11011570

[17] A. Albanese, V. Sciancalepore, X. Costa-Pérez, "SARDO: An Automated Search-and-Rescue Dro–ne-Based Solution for Victims Localization," IEEE Transactions on Mobile Computing, Vol. 21, No. 9, pp. 3312–3325, 2022. https://doi.org/10.1109/TMC.2021.3051273

[18] M.S. Hii, P. Courtney, P.G. Royall, "An Evaluation of the Delivery of Medicines Using Drones," Drones, Vol. 3, No. 3, 52, 2019. https://doi.org/10.3390/drones3030052

[19] A.S. Kristensen, D. Ahsan, S. Mehmood, S. Ahmed, "Rescue Emergency Drone for Fast Res–ponse to Medical Emergencies Due to Traffic Accidents," International Journal of Health and Medical Engineering, Vol. 11, No. 11, pp. 637–641, 2017.

[20] L. Liu & Z. You, "Drone transports medical supplies to Puerto Rico based on shortest path," IOP Conference Series: Earth and Environmental Science, Vol. 512, No. 1, 012140, 2020. https://doi.org/10.1088/1755-1315/512/1/012140

[21] B. Rabta, C. Wankmüller & G. Reiner, "A drone fleet model for last-mile distribution in disaster relief operations. International Journal of Disaster Risk Reduction, Vol. 28, pp. 107–112, 2018. https://doi.org/10.1016/j.ijdrr.2018.02.020

[22] R. R. Patel, P.J. Parmar, M.J. Sutariya, "MedDrone Rescue: A Survey on Drones for Medical Supplies in Rural, Terrained Areas and Search & Rescue Operations," 2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS), pp. 1682–1689, 2025.

[23] A. A. Bany Abdelnabi, G. Rabadi, "Human Detection from Unmanned Aerial Vehicles' Images for Search and Rescue Missions: A State-of-the-Art Review," in IEEE Access, vol. 12, pp. 152009-152035, 2024, doi: 10.1109/ACCESS.2024.3479988.

[24] S. Ma, Y. Zhang, L. Peng, C. Sun, B. Ding, Y. Zhu, "OWRT-DETR: A Novel Real-Time Transformer Network for Small-Object Detection in Open-Water Search and Rescue from UAV Aerial Imagery," in IEEE Transactions on Geoscience and Remote Sensing, vol. 63, pp. 1-13, 2025, Art no. 4205313, doi: 10.1109/TGRS.2025.3560928.

[25] H. Liu, Y. P. Tsang, C. K. M. Lee, Y. Wang, F. -Y. Wang, "Internet of UAVs to Automate Search and Rescue Missions in Post-Disaster for Smart Cities," 2024 IEEE Intelligent Vehicles Symposium (IV), Jeju Island, Korea, Republic of, 2024, pp. 614-619, doi: 10.1109/IV55156.2024.10588641.

[26] Z. Fu, Y. Xiao, F. Tao, P. Si & L. Zhu, "DLSW-YOLOv8n: A Novel Small Maritime Search and Rescue Object Detection Framework for UAV Images with Deformable Large Kernel Net," Drones, Vol. 8, No. 7, 310, 2024. https://doi.org/10.3390/drones8070310

[27] Bei Liu, Jiangliang Jin, Yihong Zhang, Chen Sun, "WRRT-DETR: Weather-Robust RT-DETR for Drone-View Object Detection in Adverse Weather", Drones, vol.9, no.5, pp.369, 2025.

[28] C. Velavan, A. Balamurugan, L. Murugasamy, "5G and Edge Computing Enabled Search and Rescue Drones," 2025 8th International Conference on Trends in Electronics and Informatics (ICOEI), pp. 579–586, 2025. https://doi.org/10.1109/ICOEI65986.2025.11013465

[29] A. Alnoman, A.S. Khwaja, A. Anpalagan, & I. Woungang, "Emerging AI and 6G-Based User Localization Technologies for Emergencies and Disasters," IEEE Access, Vol. 12, pp. 197877–197906, 2024. https://doi.org/10.1109/ACCESS.2024.3399446S.

РАЗВОЈ ПЛАТФОРМЕ ЗА СПАСАВАЊЕ ЗАСНОВАНЕ НА ДРОНУ СА ИНТЕ–ЛИГЕНТНИМ МЕХАНИЗМОМ ЗА ДЕТЕКЦИЈУ ЉУДИ И ИСПОРУКУ ВИШЕСТРУКОГ ТЕРЕТА

Т.К. Туан, В.Т.Н. Хан, Т. Кујен, Д.Ф. Тин, Ф.Т. Дат, Д.А. Дуј

Поплаве често стварају ванредне ситуације у којима је приступ жртвама озбиљно ограничен брзим покретом воде, дубоким тереном и опасним преп–рекама. Традиционалне операције спасавања у так–вим условима су обично споре, захтевају много ресурса и излажу спасиоце значајним ризицима, што ограничава ефикасност реаговања у ванредним ситуацијама. Овај рад представља концептуални дизајн платформе беспилотне летелице са веш–тачком интелигенцијом за брзо спасавање у условима поплава, посебно у опасним и непри–ступачним подручјима. Систем интегрише прила–гођени, реконфигурабилни механизам терета који се састоји од четири зглобна крака постављена испод дрона, свака са шест степени слободе покретана RC сервомоторима. Ова конфигурација пружа високу прилагодљивост, омогућавајући безбедан транспорт и активно ослобађање разноврсних залиха за спаса–вање (нпр. медицинске кутије, неопходне робе, бове за спасавање). Паралелно, модел за детекцију људи заснован на дубоком учењу, изграђен на YOLOv12, тренира се на снимцима из ваздуха снимљеним у сценаријима поплава како би се брзо откриле и локализовале жртве из перспективе беспилотне летелице. Излази детекције пружају тренутне сиг–нале за циљ, омогућавајући дрону да брзо одреди локацију жртве и испоручи терет директно у досег, чиме се смањује когнитивни и физички терет и за оператера и за жртву. Механичке симулације четво–рокраког механизма за више типова корисних терета, заједно са резултатима обуке YOLOv12 детектора, указују на изводљивост предложеног приступа и његову погодност за временски кри–тичне услове на терену. Генерално, комбинација локализације жртве уз помоћ вештачке интели–генције и флексибилног, активног механизма корис–них теретова смањује људску интервенцију, побољ–шава тачност испоруке и подржава паметно, ефи–касно и лако распоредиво решење за операције реаговања на катастрофе.