

An End-to-End UAV–Cloud–AI Pipeline for Infrared-Based Real-Time Person Detection in SAR Missions

Seunghyun Kim¹, Jeonghyo Oh², Younggon Oh³, Impyeong Lee⁴

¹Dept. of Spatial Information Engineering, University of Seoul, Seoul 02504, Republic of Korea – glowring@uos.ac.kr

²Dept. of Spatial Information Engineering, University of Seoul, Seoul 02504, Republic of Korea – ohrora98@uos.ac.kr

³Dept. of Spatial Information Engineering, University of Seoul, Seoul 02504, Republic of Korea – ohrgon3446@uos.ac.kr

⁴Dept. of Spatial Information Engineering, University of Seoul, Seoul 02504, Republic of Korea – iplee@uos.ac.kr

Keywords: UAV, Cloud Computing, Infrared Sensor, Automatic Detection, SAR System

Abstract

To effectively respond to the growing number of search and rescue (SAR) incidents in mountainous areas, a real-time and automated detection system is essential. Traditional SAR operations still rely heavily on manual visual detection, which becomes significantly limited in low-light or night-time environments. The proposed system is an integrated UAV–Cloud–AI automated architecture designed to enable real-time detection even in low-visibility environments, with the goal of improving rescue efficiency. The entire pipeline—ranging from Uncrewed Aerial Vehicles (UAV) operation and IR data acquisition to MinIO-based storage, event-driven file conversion, SFTP-based transmission, and AI-based inference—is fully automated without manual intervention. To evaluate system latency, the entire process was divided into four key stages: image acquisition, cloud upload and conversion, server transmission, and object detection. A total of 231 Infrared (IR) images were collected across five sorties, with an average processing time of 12.4 seconds per image. The upload and conversion stage showed the longest delay at 10.474 seconds, while file transfer and model inference recorded stable performances of 0.532 seconds and 1.402 seconds, respectively. In addition, detection experiments using YOLOv12 demonstrated that the model consistently identified human targets in thermal imagery, even under complex backgrounds and low thermal contrast. This study experimentally validated the feasibility of a UAV-based SAR system capable of real-time detection and response. Its scalability and field applicability are expected to be further enhanced through the future integration of lightweight detection models and collaborative multi-drone architectures.

1. Introduction

According to the annual reports of the major European mountain rescue organizations, approximately 70,000 mountain rescue missions are carried out each year in Europe alone (Rega, 2023; KfV, 2024; Montagna.tv, 2024; Snosm, 2024; Bergwacht Bayem, 2024). These operations occur repeatedly, both during the day and at night, and the increasing number of incidents continues to place a heavy burden on rescue personnel. Consequently, a systematic search framework is required to effectively respond to the recurring nature of mountain rescue demands.

As mountain rescue requests increase, local Fire and Rescue Authorities are incorporating a variety of aerial assets into their search systems. Among these, manned aircraft are effective for surveying wide areas. However, they are limited by significant personnel and operational costs. As a more efficient alternative, drone-based search operations are being increasingly adopted (Ha et al., 2021; Cho et al., 2020).

Drone-based search operations offer a key advantage in their ability to operate in areas that are difficult for humans to access or under poor lighting conditions. To enhance detection performance in such challenging environments, infrared (IR) sensors are increasingly employed. These sensors detect thermal signals based on the body heat of victims, making them effective even when RGB imagery alone is insufficient for identification. IR sensors maintain high recognition accuracy in situations where visual identification is limited such as in mountainous terrain, shaded areas, or when subjects wear clothing similar in color to the background. Moreover, they are robust to changes in weather and lighting conditions, making

them a reliable sensing technology across diverse operational environments (Yeom, 2024).

When imagery collected by drones is transmitted rapidly, rescuers can assess the victim's location and situation before arriving on-site, thereby improving the accuracy of equipment and personnel deployment. Such real-time information is not only valuable to field responders but also plays a critical role in decision-making at the command center. By enabling the assessment of a victim's condition and surrounding environment without solely relying on verbal reports from field personnel, real-time situational sharing systems serve as a core technological foundation for collaborative decision-making in rescue operations.

However, conventional drone usage still requires rescuers to manually inspect the imagery for victim detection, which prevents the full advantages of drone mobility and rapid deployment from being realized (Gotovac et al., 2020). To address these limitations and enhance the operational efficiency of drones, the adoption of automated detection technologies is essential (Abdelnabi and Rabadi, 2024; Kundid Vasić and Papić, 2020). A prominent example is object detection using Convolutional Neural Networks (CNNs), which requires a fast and stable data transmission system for effective implementation (Tian et al., 2025).

When image data is rapidly transmitted from drones to the cloud server, it becomes possible to automatically detect victims in real time. This reduces the reliance on subjective human judgment and enables rescue operations to be carried out based on consistent detection performance (Kundid Vasić and Papić, 2020). Given the limited flight time of drones, the entire process from data collection and transmission to analysis and decision-making must be completed within a single flight without delay.

In order to satisfy these operational constraints, the automated detection system should be closely integrated with the drone-based processing pipeline and architected to function reliably under varying environmental conditions and time limitations.

In search and rescue missions, the speed of data collection and transmission has a direct impact on survival rates, making automation and real-time processing in drone-based systems essential. A UAV–Cloud–AI-based object detection system minimizes the delays and human dependency inherent in traditional rescue approaches, enabling faster and more efficient responses. Such systems play an increasingly critical role under real-world operational constraints, particularly in low-visibility conditions such as nighttime, where the limitations of conventional methods become more pronounced and rapid decision-making is crucial for locating victims.

To overcome these limitations, an automated detection system is required—one that transmits IR image data collected by drones to a cloud server and applies a deep learning-based object detection model, as illustrated in Figure 1. The system should automate the entire process—from data acquisition and storage to conversion, transmission, and inference—to minimize manual intervention and ensure consistent performance and real-time capability across various environments. Accordingly, this study proposes a drone-based near-real-time search platform to address these operational needs.

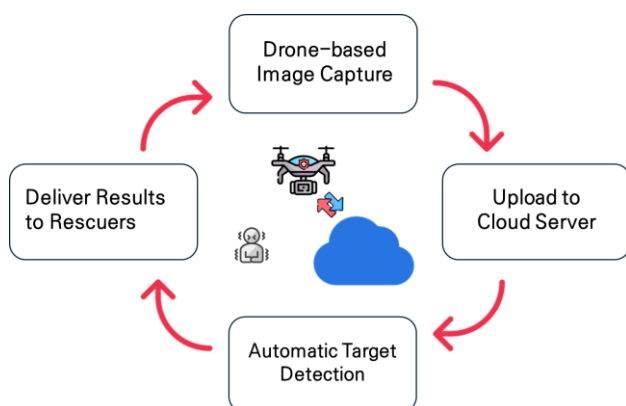


Figure 1. Conceptual diagram of drone-based real-time rescue system.

2. System Architecture and Methodology

2.1 Overview of the Automated SAR Pipeline

The UAV-based real-time automated detection pipeline proposed in this study consists of five stages, as illustrated in Figure 2. In Stage 1, thermal images are captured using an IR sensor mounted on the drone and transmitted to the remote controller. In Stage 2, the received images are uploaded to a cloud server, where they are converted to JPG format and processed for storage and preprocessing. Stage 3 involves transferring the converted images to a dedicated processing server for inference. In Stage 4, real-time inference is performed on the processing server—equipped with high-performance GPUs—using a deep learning-based object detection model to identify human targets. Finally, Stage 5 delivers the detection results to the responder to support decision-making and rescue operations. The proposed pipeline automates the entire process

from image acquisition to result delivery, thereby minimizing detection delays and enabling timely response in rescue missions.

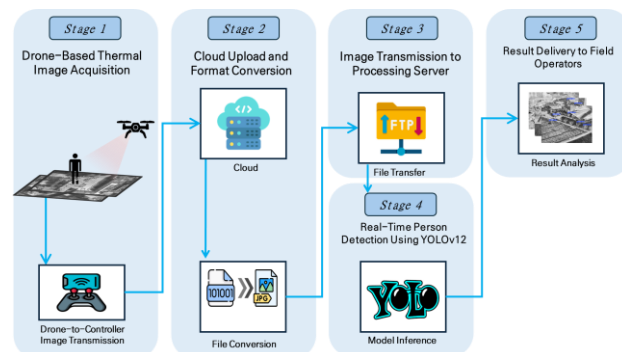


Figure 2. Workflow of the proposed UAV–Cloud–AI automated pipeline.

2.2 Drone-Based Thermal Image Acquisition

In this study, the DJI M30T drone, which supports real-time data acquisition and integration with external systems, was utilized for data collection. This drone employs DJI's OcuSync technology, enabling low-latency transmission of video and control signals between the drone and the remote controller, thereby allowing real-time video transfer. The collected data is then integrated with an Object Storage Service (OSS) and processed through a cloud pipeline based on the DJI Cloud API. To enhance the accuracy of human target detection, the drone is equipped with an IR sensor that effectively distinguishes the thermal signatures of human subjects from the background, enabling reliable operation in low-light or visually obstructed environments such as nighttime or mountainous terrain.

The overall structure for collecting, transmitting, and processing data in real time at the scene of an incident is as follows. Thermal data acquired through the IR sensor is transmitted to the remote controller via OcuSync wireless technology and then processed in real time on a cloud server integrated with the DJI Cloud API. This API supports various functions such as video data transmission, real-time monitoring, and remote control. In the actual system, it was further customized to establish a direct connection between the remote controller and the cloud server. In the implemented pipeline, the video data transmitted to the controller is automatically stored on the server, after which real-time inference is performed using a deep learning model. This system is designed to store field-acquired data in near real time and perform automated human detection based on the stored data, thereby significantly improving response speed and operational efficiency in rescue scenarios.

2.3 Cloud Upload and Format Conversion

Thermal images received from the drone are collected in real time and initially stored in binary file format before being converted to JPG format. This stage incorporates automatic detection and conversion triggers, which are essential for maintaining the system's real-time performance and automation. To efficiently manage data transmitted from the drone and remote controller, the cloud server was built on a locally hosted MinIO server. MinIO is a lightweight object storage system that enables easy deployment of a private cloud environment and supports flexible data handling in various experimental and operational scenarios. Furthermore, its compatibility with external processing systems and its integration with Watchdog-

based event detection and real-time processing functions make it well-suited for the proposed system.

In the actual operating environment, IR image data acquired by the drone and remote controller are uploaded to the MinIO server in near real time and stored in binary format. To automate the process of detecting uploaded data, converting binary files into JPG format, and transmitting them to the processing server, an event-driven detection system based on Watchdog was implemented. This system automatically handles both data conversion and transmission, with a focus on maximizing the efficiency of real-time data processing. Instead of continuously monitoring all changes within the folder, the detection mechanism is designed to respond only to specific events, such as the creation of new folders, thereby reducing system resource usage. As a result, once a binary file is detected upon upload, it is immediately converted into JPG format and passed to the next processing stage.

2.4 Image Transmission to Processing Server

The converted images are automatically transmitted to a separate processing server through a dedicated transfer stage, where inference is immediately performed using a deep learning model. This transmission process employs the Secure File Transfer Protocol (SFTP), which provides high security through encryption during file transfer and ensures stable performance even under unstable network conditions.

In particular, implementation in Python is straightforward using libraries such as *paramiko*, allowing flexible integration with various Python-based environments and system configurations. The entire system is developed in Python, ensuring broad library compatibility and enabling the construction of an end-to-end automated data processing pipeline—from data upload and format conversion to real-time inference.

2.5 Real-Time Person Detection Using YOLOv12

This stage serves as the core computational component of the automated pipeline, where deep learning-based object detection is performed using high-performance GPUs, taking JPG images transmitted from the cloud server as input. It is a critical element in implementing a real-time automated detection system aimed at overcoming the subjectivity and omission issues inherent in traditional search methods. In particular, the cloud-based computational environment compensates for the limited processing capabilities of UAV systems, enabling stable operation of complex models and full automation of the pipeline, thereby improving both the accuracy and speed of rescue operations.

In the actual system, a real-time operation structure capable of automatic inference was implemented within the data processing server. It was designed to detect images as soon as they are uploaded to the server and automatically perform inference through an object detection model. Through the event detection function linked to image upload, near-real-time detection is possible for each individual image.

In this study, considering the trade-off between real-time detection performance and accuracy, we adopted YOLOv12, the latest object detection model. The You Only Look Once (YOLO) series is a convolutional neural network (CNN)-based single-stage object detector designed to achieve a balance between high inference speed and accuracy. Owing to its user-friendly design and accessibility for non-experts in deep

learning, the YOLO series has been widely adopted across various application domains. Among these, YOLOv12 demonstrates excellent performance even with relatively small training datasets and offers enhanced accuracy and processing speed compared to previous versions in the series.

In the proposed detection stage, the model is configured to automatically detect human objects immediately upon the arrival of new data by linking the inference process to the image upload event. This configuration is expected to effectively fulfill the performance requirements of real-time automatic detection.

3. Experiments and Results

3.1 Experimental Setup

To evaluate the detection performance and processing time of the model, experiments were conducted at night (21:00 ~ 22:00) in a small mountainous area near an urban region. The test site featured low illumination and limited visibility, partially simulating external conditions encountered in real-world search scenarios. The total survey area was 20,967.8 m², and the Ground Sample Distance (GSD) was set to 3 cm.

A single flight from drone departure to return was defined as a sortie, and a total of five sorties were conducted using the IR camera mounted on the DJI M30T drone. For each sortie, 45 to 48 thermal images were captured at approximately 5-second intervals, resulting in a total of 231 IR images. The collected images have a resolution of 1280×1024 pixels, with an average file size of approximately 1.6 MB.

In this experiment, the processing time per image and the total time required for each sort were analyzed together to evaluate the time efficiency of the system. For the model used for inference, a public thermal image dataset centered on human objects captured using an IR sensor was used for learning (Speth et al., 2022; Suo et al., 2023). The dataset consists of a total of 1,172 images, of which 1,040 were used for learning and 134 for verification. The entire image contains a total of 3,501 human instances.

In this study, object detection learning was performed based on the YOLOv12 model, and the following hyperparameters were set for optimal learning performance. The initial learning rate was set to 0.01 to allow the model to converge effectively, and the final learning rate fraction was set to 0.01 so that the learning rate gradually decreased as learning progressed. AdamW was used as the optimization algorithm, weight decay 0.0005 was applied, and momentum was set to 0.937 to increase learning stability. Learning was performed for a total of 100 epochs, and the batch size was set to 16 to consider the balance between learning speed and memory efficiency. Model learning and inference were performed on a processing server equipped with NVIDIA RTX 3090 GPU, and the hardware environment enabled efficient processing of deep learning models requiring large-scale computation.

3.2 Time Points and Latency Structure

To evaluate the performance of the drone-based automatic detection pipeline, Stage 1 through 4 of the previously defined five-stage process were selected for analysis, and processing times were measured based on the completion time of each stage. Figure 3 illustrates the entire processing flow, from the moment an incident occurs to the point at which the detection

result is generated. The completion times of each stage are denoted as T1 through T4.

T1 corresponds to the time when the drone captures the video on-site. T2 is the point at which the image is stored on the cloud server and converted to JPG format. T3 marks the upload of the converted image to the object detection model, and T4 indicates the moment when inference is completed and the result is produced by the model.



Figure 3. Overview of the end-to-end processing pipeline segmented by key time points (T1–T4).

By measuring the time required for each section based on each point in time, it is possible to identify bottlenecks within the system and derive optimization directions for improving real-time performance. In this analysis, the entire processing process from image capture to object detection was targeted at stage 2 to 4, and the performance by section was analyzed as shown in Table 1

Stage	Calculation	Description
Stage 2	T2-T1	Capture → Upload/Conversion complete
Stage 3	T3-T2	Conversion complete → Transmission complete
Stage 4	T4-T3	Transmission complete → Inference complete

Table 1. Definitions of processing phases based on system time intervals

Stage 2 is the section where the drone takes the video, and the video is uploaded to the cloud storage and converted to JPG format. This is mainly affected by the Internet upload speed and conversion processing performance. Stage 3 is the section where the converted image is transmitted to the data processing server. This is mainly affected by the network transmission speed between servers. Stage 4 is the section where deep learning-based inference is performed on the image uploaded to the object detection model and the result is output. The processing time is determined by the model's inference speed.

3.3 Latency Analysis by Processing Stage

In this study, a drone-based automatic detection pipeline was constructed, and the performance of each processing section was summarized in Table 2. The table records the timestamps of each major processing step from the initial video capture T1 to the output of object detection results T4 in hours, presenting the sequential flow of data in chronological order. This allows for a direct assessment of whether each stage in the system operates in real time.

	T1	T2	T3	T4
1	21:41:55.000	21:42:00.966	21:42:01.495	21:42:02.784
2	21:41:59.000	21:42:05.152	21:42:05.678	21:42:06.952
3	21:42:04.000	21:42:10.417	21:42:10.952	21:42:12.228
4	21:42:09.000	21:42:17.590	21:42:18.158	21:42:19.414
5	21:42:13.000	21:42:18.981	21:42:19.491	21:42:20.779

Table 2. Time log of sequential processing steps

Based on the time point of each video frame, the data is divided into a sortie, which is a single flight unit, and the average processing time for each sortie is presented in Table 3. The analysis was performed by dividing the entire process from video shooting to object detection into stages to identify the processing time characteristics, bottlenecks, and improvement directions for each section. In addition, Figure 4 visualizes the time required for each processing stage across 46 frames within one sortie, complementing the data summarized in the table.

Sortie	No.Frames	Stage 2(s)	Stage 3(s)	Stage 4(s)	Stage 2~4(s)
Sortie 1	46	11.17	0.531	1.392	13.093
Sortie 2	48	10.67	0.543	1.411	12.624
Sortie 3	45	10.47	0.532	1.364	12.366
Sortie 4	47	9.57	0.526	1.37	11.466
Sortie 5	45	10.49	0.529	1.471	12.490
Average	-	10.474	0.532	1.402	12.408

Table 3. Stage-wise processing time for each sortie

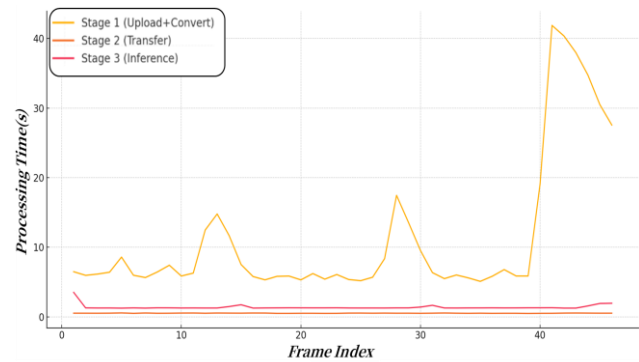


Figure 4. Stage-wise time analysis of 46 frames from a single UAV sortie.

The average processing time for Stage 2 was 10.474 seconds, making it the bottleneck with the highest delay in the entire pipeline. This stage was particularly sensitive to external environmental factors such as network conditions, storage delays, and file conversion performance. Therefore, Stage 2 requires the highest priority for optimization to enhance real-time performance. Technical solutions such as improving network speed, restructuring the cloud upload process, lightening the conversion algorithm, and implementing parallel processing are recommended. In contrast, the average processing time for Stage 3 was 0.532 seconds, indicating stable inter-server communication.

This section is considered to operate reliably without causing bottlenecks within the real-time system. Stage 4 recorded an average processing time of 1.402 seconds and demonstrated consistent inference performance across all sorties. This result confirms that the GPU-based computational environment and the optimized structure of the YOLOv12 model effectively support real-time inference performance.

Based on the analysis of each processing stage, it can be concluded that technical improvements in Stage 2 should be prioritized to ensure the real-time performance of the overall pipeline. The remaining stages are considered to have achieved performance levels that can be stably integrated into the current system architecture.

These findings provide experimental evidence that the proposed drone–cloud–AI-based detection system can achieve practical detection performance and time efficiency, even under nighttime or complex environmental conditions. Such results may serve as a foundation for future optimization of the system and its deployment in real-world scenarios.

3.4 Detection Accuracy Evaluation

When applied to thermal images captured at night using an IR sensor, the YOLOv12 model reliably detected numerous human subjects, even under challenging conditions such as complex backgrounds and weak thermal signatures.

As shown in Figure 5, the inference results indicate that human subjects were visually distinguishable in the IR images, with successful detection achieved in most cases, except for a few outliers. These findings confirm that the model maintains robust detection performance in low-light environments.

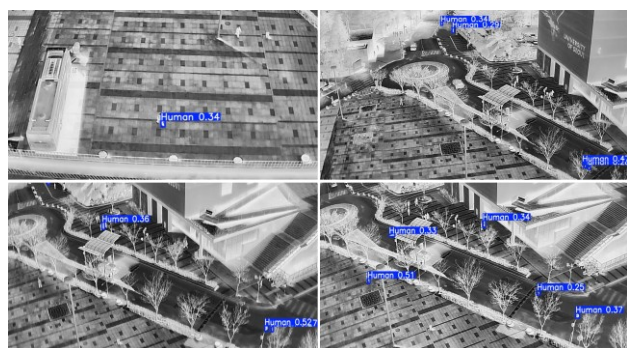


Figure 5. Visualization of YOLOv12 detection results on IR images.

Model training was conducted using 134 out of a total of 1,174 images as validation data. As a result, the model achieved a Precision of 0.896, a Recall of 0.847, and an AP50 of 0.915. These results indicate that the IR-based object detection system can perform effectively even in nighttime environments and suggest that IR sensors can serve as a viable detection tool in structural rescue scenarios where securing adequate lighting is difficult.

However, detection accuracy may be affected by various environmental factors, such as time-dependent ground temperature variations and object shape distortion caused by differences in camera angles between the drone and the training dataset. Therefore, building a dataset specialized for IR conditions and advancing detection models to be robust under diverse imaging conditions are critical to improving overall performance.

4. Conclusion and Future Work

To overcome the limitations of detection in mountainous terrain and nighttime environments, this study designed a drone–cloud–AI-based pipeline that automates the entire process from image acquisition to transmission and inference. The proposed system collects images using a DJI drone equipped with an IR sensor, stores them in the cloud via the DJI Cloud API and MinIO, and uses a Watchdog to detect the uploaded data, convert it into JPG format, and transmit it to the processing server via SFTP. Subsequently, another Watchdog instance on the server detects the incoming files and performs real-time

object detection using a YOLOv12-based model. This pipeline is designed as a real-time automatic detection system to enable near-real-time response in rescue scenarios.

The average processing time for the entire pipeline was measured at approximately 12.4 seconds per image, indicating the potential for near-real-time response in actual rescue scenarios. In particular, Stage 2 showed the highest latency, with an average processing time of 10.474 seconds, making it the primary bottleneck in the pipeline. This section was found to be highly sensitive to external factors such as network conditions, storage delays, and file conversion overhead. To optimize this stage, improvements in network speed, restructuring of the upload process, algorithm lightweighting, and the introduction of parallel processing are required.

In contrast, Stage 3 and Stage 4 demonstrated relatively stable processing times of 0.532 seconds and 1.402 seconds, respectively. These results indicate that the cloud-based computing environment and the optimized architecture of the YOLOv12 model contribute to the consistency and reliability of real-time detection. Moreover, in nighttime experiments using IR sensors, human subjects were clearly identifiable even under low illumination and complex background conditions, with stable detection achieved in most images.

This experimentally validates the effectiveness of the IR-based detection system for search-and-rescue operations conducted in low-light and challenging environments.

The training data used in this study consisted of infrared images collected under a single condition at a specific time point, which has limitations in terms of generalization to various periods and environmental settings. In particular, infrared-based detection is very sensitive to external factors such as temporal temperature changes, seasonal changes, and weather conditions, which can affect detection performance. Therefore, future research should focus on building an infrared-specific dataset that includes various temporal and seasonal conditions and improve the detection model to improve the robustness to environmental changes.

In future research, we aim to enhance the performance of detection models that are robust to external environmental variations by constructing IR datasets that reflect temperature conditions across different seasons and time periods. Additionally, we plan to explore methods for optimizing the data transmission structure and improving the efficiency of the conversion process to address the bottleneck observed in Stage 2, which accounted for the greatest delay in the entire pipeline.

Acknowledgements

This research was supported by Korea Institute of Marine Science & Technology Promotion(KIMST) funded by the Ministry of Oceans and Fisheries, Korea(RS-2022-KS221629).

References

- Abdelnabi, A.A.B., Rabadi, G., 2024: Human Detection from Unmanned Aerial Vehicles' Images for Search and Rescue Missions: A State-of-the-Art Review. *IEEE Access*. doi: 10.1109/ACCESS.2024.3479988.
- Bergwacht Bayem, 2024: Über uns – Organisation. Bavarian Mountain Rescue Service, Germany. <https://bergwacht-bayern.de/ueber-uns/organisation> (3 July 2025).

Cho, J.Y., Song, J.I., Jang, C.R., Jang, M.Y., 2020: A study on the utilization plan of drone videos for disaster management. *Journal of the Korea Academia-Industrial cooperation Society*, 21(10), 372–378. <https://doi.org/10.5762/KAIS.2020.21.10.372>.

Gotovac, S., Zelenika, D., Marušić, Ž., Božić-Štulić, D., 2020: Visual-based person detection for search-and-rescue with UAS: Humans vs. machine learning algorithm. *Remote Sensing*, 12(20), 3295. <https://doi.org/10.3390/rs12203295>.

Ha, K.H., Kim, J.H., Choi, J.W., 2021: A study on the application of drone in firefight field. *Journal of the Korea Academia-Industrial Cooperation Society*, 2021, 22.4: 321-330. <https://doi.org/10.5762/KAIS.2021.22.4.321>.

KFV, 2024: PA Bergrettungssymposium 2024. Kuratorium für Verkehrssicherheit (KFV), Austria. https://www.kfv.at/wp-content/uploads/2024/06/PA_Bergrettungssymposium-2024.pdf

Kundid Vasić, M., Papić, V., 2020: Multimodel deep learning for person detection in aerial images. *Electronics*, 9(9), 1459. <https://doi.org/10.3390/electronics9091459>.

Montagna.tv, 2024: Il Soccorso Alpino e Speleologico dà i numeri: ben 12.349 interventi effettuati nel 2023. <https://www.montagna.tv/236256/il-soccorso-alpino-e-speleologico-da-i-numeri-ben-12-349-interventi-effettuati-nel-2023> (3 July 2025).

Rega, 2023: Rega helped more patients than ever before. Rega – Swiss Air-Rescue. <https://www.rega.ch/en/news/news-from-the-world-of-rega/detail/rega-helped-more-patients-than-ever-before> (3 July 2025).

Snosm, 2024: Bilan d'activité 2023. Service National d'Observation des Secours en Montagne (SNOSM), France. https://www.snosm.fr/sites/default/files/2024-06/Bilan_snosm_2023_version%20site%20internet.pdf (3 July 2025).

Speth, S., Gonçalves, A., Rigault, B., Suzuki, S., Bouazizi, M., Matsuo, Y., and Prendinger, H., 2022: Deep Learning with RGB and Thermal Images onboard a Drone for Monitoring Operations. *Journal of Field Robotics*, 1–29. <https://doi.org/10.1002/rob.22082>.

Suo, J., Wang, T., Zhang, X., Chen, H., Zhou, W., and Shi, W., 2023: HIT-UAV: A high-altitude infrared thermal dataset for Unmanned Aerial Vehicle-based object detection. *Scientific Data*, Nature Portfolio. <https://doi.org/10.1038/s41597-023-02066-6>.

Tian, Y., Ye, Q., Doermann, D., 2025: YOLOv12: Attention-Centric Real-Time Object Detectors. *arXiv preprint arXiv:2502.12524*. <https://doi.org/10.48550/arXiv.2502.12524>.

Yeom, S., 2024: Thermal image tracking for search and rescue missions with a drone. *Drones*, 8(2), 53. <https://doi.org/10.3390/drones8020053>.