Object detection and augmented reality annotations for increased situational awareness in light smoke conditions

Ioanna Gkika*
Visual Computing Lab (VCL),
Information Technologies Institute (ITI),
Centre for Research and Technology - Hellas (CERTH),
Thessaloniki, Greece †
ioanna.gkika@iti.gr

Dimitrios Pattas
Visual Computing Lab (VCL),
Information Technologies Institute (ITI),
Centre for Research and Technology - Hellas (CERTH),
Thessaloniki, Greece †
dim.pattas@iti.gr

Konstantinos Konstantoudakis
Visual Computing Lab (VCL),
Information Technologies Institute (ITI),
Centre for Research and Technology - Hellas (CERTH),
Thessaloniki, Greece †
k.konstantoudakis@iti.gr

Dimitrios Zarpalas
Visual Computing Lab (VCL),
Information Technologies Institute (ITI),
Centre for Research and Technology - Hellas (CERTH),
Thessaloniki, Greece †
zarpalas@iti.gr

ABSTRACT

Innovative technologies powered by Computer Vision algorithms can aid first responders, increasing their situational awareness. However, adverse conditions, such as smoke, can reduce the efficacy of such algorithms by degrading the input images. This paper presents a pipeline of image de-smoking, object detection, and augmented reality display that aims to enhance situational awareness in smoky conditions. A novel smoke-reducing deep learning algorithm is applied as a preprocessing step, before state-of-the-art object detection. The detected objects and persons are highlighted in the user's augmented reality display. The proposed method is shown to increase detection accuracy and confidence. Testing in realistic environments provides an initial evaluation of the method, both in terms of image processing and of usefulness to first responders.

Keywords

Image Processing, Smoke, Augmented Reality, Deep Learning, Situational Awareness

INTRODUCTION

First responders (FRs) are often called to operate in dangerous conditions, with risks both to themselves and the civilian population. Situational awareness, a person's knowledge and perception of a changing environment, is a core concept in response operations, safeguarding the FRs' own safety and increasing their efficiency (Endsley 1995).

*corresponding author
†`http://vcl.iti.gr/`

In recent years, advanced technologies have been utilized to increase FRs' situational awareness, including drones (Sainidis et al. 2021), social media analytics (Snyder et al. 2019), and augmented reality (AR) (Christaki et al. 2022; Bhattarai et al. 2020). Vision naturally has a central role in situational awareness, and many elaborate solutions rely on Computer Vision (CV) algorithms, often powered by Artificial Intelligence (AI) (Lygouras et al. 2019; Kyrkou and Theocharides 2019), to detect objects of interest, including victims.

However, clarity of vision - be it human or computer - is seldom guaranteed in disaster response situations. FRs on the field may operate in vision-reducing conditions, including adverse weather, low light, or smoke. Such conditions inevitably reduce the efficacy of CV algorithms, just as they reduce human perception. To mitigate this effect, an additional pre-processing step can be applied to the input images or video degraded by adverse conditions, restoring some clarity before the main processing. Hence, there will be no need for a FR to walk to a location and lose precious time to recognise objects.

In this work in progress, we focus on improving CV in smoke and we present a pipeline of image de-smoking, object detection, and visualization of the output in AR, annotating important objects or persons in an FR's vicinity. The proposed methodology highlights the increased efficiency of object detection algorithms after our deep-learning (DL)-based de-smoking algorithm and presents a vision of future AR-enabled situational awareness.

The rest of this paper is organized as follows: Related Work discusses related previous work on image smoke removal, visual object detection, and increasing situational awareness through AR. System Architecture presents the architecture of the proposed system, describing the hardware and communication tools that make up the pipeline from image acquisition to AR display, while the software components, which are the main focus of this work, are described in Software components. Demonstration and evaluation presents the results of early testing in realistic conditions, while Conclusion and Future Work summarizes this paper and discusses future work on this subject.

RELATED WORK

Smoke removal

Images captured in inclement weather or environmental conditions like fog, haze or smoke usually have low contrast, low colour fidelity, the edge regions are degraded and generally they are not visually pleasing. This degradation can further reduce the efficiency of FRs during their mission and could expose them to danger. For this reason, there exists the need to create an image denoising algorithm to assist FRs on their mission.

Fog can be considered as the phenomenon where water particles or droplets are suspended in the air, while haze and smoke are created by small but dry particles. The degradation caused by these conditions, can also affect the performance of CV algorithms used for tasks like object detection, since they are mainly trained on images with clear weather conditions where the distribution of colours differs. Thus, fog, haze or smoke removal is a challenging but useful task that has gained great interest in the research community.

The methods used for fog or haze removal can be divided in two main categories: the hand-crafted priors-based and the data-driven. Before the wide application of the DL in image dehazing, most of the methods were based on different prior assumptions like dark channel (He et al. 2010), contrast maximization (Tan 2008), colour attenuation (Zhu et al. 2014) and non-local prior (Berman, Avidan, et al. 2016) and the Atmospheric Scattering Model (ASM) (McCartney 1976). Unlike these methods, the most recent works are data-driven and exploit large-scale datasets in order to train a DL network to remove the noise from an image (Cai et al. 2016; Ren, Liu, et al. 2016; Ren, Ma, et al. 2018; Qin et al. 2020; Fu et al. 2021; Engin et al. 2018; Dudhane and Murala 2019; Hong et al. 2020).

Early data-driven methods (Cai et al. 2016; Ren, Liu, et al. 2016) mainly use supervised learning and are dependent on the ASM. They exploit trainable CNNs to estimate the transmission map of a hazy image and then recover the corresponding clear one. At the next step, researchers developed DL-based networks that ignore completely the ASM and recover the clear image in an end-to-end manner (Ren, Ma, et al. 2018; Qin et al. 2020; Fu et al. 2021; Engin et al. 2018; Dudhane and Murala 2019; Hong et al. 2020; Li, Gou, et al. 2020).

In dehazing task, a major challenge is that it is impossible to acquire the same picture with and without noise simultaneously. Thus, the hazy images used for training the networks are created by three different ways: 1) adding synthetic haze in clear images (the clear images can be either real or synthetic), 2) using a professional haze machine, 3) capturing pairs of real hazy and clear images asynchronously. Some recent research works tried to overcome this issue using either unpaired training or Unsupervised Learning. These networks are able to restore the clear image without being trained on pairs of hazy and the corresponding ground-truth images. They are based on Generative Models (Engin et al. 2018; Dudhane and Murala 2019) Knowledge Distillation (Hong et al. 2020) or Zero-Shot Learning (Li, Gou, et al. 2020).

Object detection

Object detection algorithms could be extremely useful for improving the situational awareness of the FR as they can identify objects of interest, victims or obstacles, that they may not have noticed during their missions. This information can be visualised using an AR interface in order the FRs to have a live update of what is in front of them and make their mission shorter and less dangerous.

The state-of-the-art object detection algorithms can be divided into two main types: One-stage and two-stage object detectors. In the two-stage object detectors the models first propose a set of regions of interest and the classifier and bounding box regressor process only these regions of interest. In the one-stage object detectors the region proposal stage is skipped and the models predict the bounding boxes and the classes over the image. The most popular algorithms used for object detection include CNNs (Region-Based CNNs (R-CNN) (Girshick et al. 2014)) (two-stage), Fast R-CNN (Girshick 2015) (two-stage), and You Only Look Once (YOLO) (Redmon et al. 2016) (one-stage). For our experiments, we exploited the YOLO algorithm and specifically the YOLOv5s[1]. It is trained using the Microsoft Common Object in Context (COCO) dataset and is able to detect 80 different classes. It outputs the bounding boxes' pixel coordinates along with the objects label and its confidence.

Enhanced Situational Awareness via AR

Although AR is still an emerging technology, its possible usefulness for increasing situational awareness during missions has sparked a number of innovative applications and case studies. Bhattarai et al. 2020 present a firefighter aid system which receives input from depth and infrared cameras, detects persons and other objects of interest, and displays the results in a Microsoft HoloLens. Similarly, Hu et al. 2022 use ground-penetrating radar to detect voids in disaster rubble and present them to the FRs in 3D on AR. Zeman et al. 2022 harness AR to aid in indoor localization and navigation using points of interest (POIs) and known blueprints. Along similar lines, Christaki et al. 2022 describe a cross-tool AR POI creation, management and visualization system, to increase FRs' situational awareness and keep them updated of findings and changes in the course of the mission.

SYSTEM ARCHITECTURE

The proposed system's architecture is based on several hardware and software modules that implement image acquisition, de-smoking, object detection, and visualization annotations in HoloLens, as well as the communication between the above. This section describes the communication and hardware aspects used, while the software components, which are the main focus of this work, are presented in the next section, Software components.

Communication and data flow

The communication between the three hardware devices is the following. The camera streams images to the laptop using the Open GoPro Python SDK[2]. The communication is seamless using Bluetooth and Wi-Fi. The images are inserted to the desmoking algorithm and then to the object detection model. The final output of this process is a JSON file containing the coordinates of the bounding boxes, the labels of the objects and the confidence level of each prediction. The JSON messages are then sent to the Message Broker, using an MQTT protocol. For the visualization, the HoloLens communicates with the laptop seamlessly using Wi-Fi and subscribes to the Message Broker to receive the messages. It reads the JSON files and visualises the messages. The whole pipeline can be seen in Fig. 1. Although the communication is seamless using Wi-Fi and Bluetooth, the Wi-Fi is used only for internal communication, meaning that the whole process can perform without mobile signal.

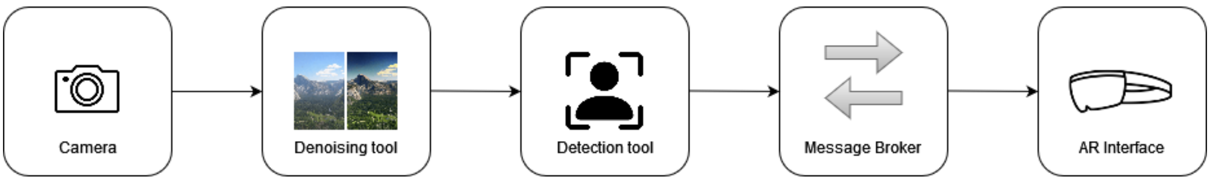

Figure 1. Architecture Diagram

[1]https://github.com/ultralytics/yolov5
[2]https://gopro.github.io/OpenGoPro/demos/python/sdk_wireless_camera_control

End-User Requirements

Since the proposing tool is targeted to rescue operations, apart from the technical view, it should also response to the actual needs that FRs have. In Table 1 we list the main requirements that the end-users requested and that the tool tried to be adapted to. In Section Demonstration and evaluation we further analyse how the tool has managed to address some of them and the reasons why other cannot be addressed so far.

Table 1. End-user requirements

Number	*Requirement*	*Description*
1	Range	Object detection effective distance $100m$
2	Preparation time	The tool should start within 1 to 2 minutes
3	Load	The tool should avoid overload of information
4	Thermal Resistance	The tool should work at a $-25°C, +150°C$ range
5	Battery autonomy	$8-12$ hours maximum battery autonomy

Hardware

Figure 2. The main hardware components used in the proposed system: GoPro camera for image acquisition and wireless connectivity, a powerful GPU-equipped laptop for processing, and the HoloLens 2 for visualization.

In terms of hardware, our tool requires a Helmet Mounted Camera (HMC) that will provide the noisy images to the tool, a laptop with a Graphics Processing Unit (GPU) that will be able to process the images in a short time and an AR interface that will visualise the detection to the FR. While performing our tests (intensive use), all of the hardware components could last around 2 hours when initially fully charged.

The camera that has been chosen is the GoPro Hero 10 Black. It is Waterproof to 33ft, has good low-light capabilities, has both wireless connection via Bluetooth and Wi-Fi and wired via Type-C, it provides GPS tagged photos and livestream video.

The laptop that was selected is the Razer Blade 14 - QHD 165Hz - GeForce RTX 3080 Ti - Black. Its small size and weight (14", and 1.78 kg) make it possible to be carried by users in a backpack during exercises, and its strong processing capabilities (GPU: Nvidia 3080 Ti 16 GB GDDR6) can facilitate fast inference of visual processing AI.

The AR interface that has been chosen is the HoloLens2.

SOFTWARE COMPONENTS

Our tool consists of 2 separate applications a) HoloLens 2 (Unity/C#), b) Laptop (Python). The HoloLens 2 application is responsible for the visualisation and everything else is being processed by the laptop's application.

Laptop app

The Laptop's application runs without internet connection on Python 3.10 , and requires CUDA GPU acceleration. Furthermore, it requires 1 Bluetooth and 2 Wi-Fi connections (1 for connectivity with the camera and 1 so that HoloLens and laptop are on the same LAN).

The laptop's application is responsible for 3 tasks:

- Establishing communication with the GoPro camera and the broker.
- Receiving frames from the camera stream, performing the pipeline (denoising, object detection, depth estimation) and exporting a json containing the necessary information to the broker.
- Initialising and running a server for receiving requests from the HoloLens, such as informing the HoloLens about the json and allowing the HoloLens to change the functionality of the laptop's app in real-time via voice commands.

Smoke removal

Since smoke and haze can cause a similar effect in an image, a dehazing model can also be applied for de-smoking. We exploited the DW-GAN (Fu et al. 2021), a state-of-the-art dehazing network that leverages the 2D discrete wavelet transform and the architecture of the Generative Adversarial Networks (GAN) to remove the haze efficiently. Additionally, we have developed a new dehazing model, which is also GAN-based.

For the training of our model we used two different datasets: The REalistic Single Image DEhazing (RESIDE) (Li, Ren, et al. 2018) and Foggy Cityscapes (Sakaridis et al. 2018). RESIDE is a large-scale dataset widely used for benchmarking of dehazing models and is publicly available. Depending on the purpose of use, it is divided into 5 subsets namely, Indoor Training Set (ITS), Outdoor Training Set (OTS), Synthetic Objective Testing Set (SOTS), Real-world Task-Driven Testing Set (RTTS) and Hybrid Subjective Testing Set (HSTS). Our training set consists of triplets of synthetic hazy images, their corresponding transmission maps and clear images and are part of the OTS (4,120 triplets) and ITS (2,798 triplets) subsets. Foggy Cityscapes is a dataset that inherits the data split and the annotations from Cityscapes (Cordts et al. 2016), and simulates fog on real images. From Foggy Cityscapes we use 20,000 triplets for training, keeping the same structure as with RESIDE.

Figure 3. First row: real hazy images. Second row: results of DW-GAN model. Third row: results of our model.

In order to train the dehazing model we employ a GAN-based scheme. The generator is composed of 4 downsampling convolutional layers, a bottleneck that consists of Fast Fourier Convolution residual blocks (Chi et al. 2020) and 4 upsampling layers. For discriminator, we utilized PatchGAN (Isola et al. 2017) to learn high frequency information and produce crisp images. Additionally, we employ a Depth Prediction Branch (DAB) which is formulated as a linear transform in the feature space (Xu and Zheng 2021). Since some parts of the image are partly occluded due to the interactions between haze and objects, learning explicitly the depth information of the scene is important for the model.

We combined multiple loss functions to supervise our model, L1 per pixel loss, a total variation smoothness prior, a high-level synthesis loss that minimizes the difference between generated and ground truth image in the feature space of a pre-trained network (in our case, VGG-19 (Simonyan and Zisserman 2014)), Hinge loss for the discriminator, and a berhu and gradient loss for supervising the depth map estimation branch.

As we can see in Fig. 3 our model is less prone to generating artifacts while the generated image is more natural and colour consistent.

Object detection

While employing a DL model for a rescue mission, we should keep in mind that FRs often work in environments with power constraints and the time is very critical for them. Hence, the models that we should choose for this application should not only be accurate enough, but also be lightweight enough to perform in near-real-time. Under these constraints the object detection model that we chose to be a part of our tool is YOLOv5s.

YOLOv5 is a family of 4 compound-scaled object detection models, namely YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. Each one of them has different detection accuracy, inference time and complexity. We chose YOLOv5s model for our pipeline, since it is the smallest and therefore has the smallest inference time and also shows a good performance.

As we have already mentioned, YOLOv5s can detect 80 different classes. However, not all of them are useful for a rescue mission. In order to avoid overload of information (Requirment 3), we have only kept the 22 most useful ones.

AR visualization

Holo app

The HoloLens 2 app is reading messages from the broker in the json format and visualises bounding boxes along with the label of the object. When a voice command is given, the HoloLens 2 app informs the laptop's app to change its functionality depending on the command.

The HoloLens app is written in Unity v2020.3.36f1. The app has 2 functionalities, requesting json containing the bounding boxes data and sending voice commands. These are implemented by communicating with the python server on the laptop's app.

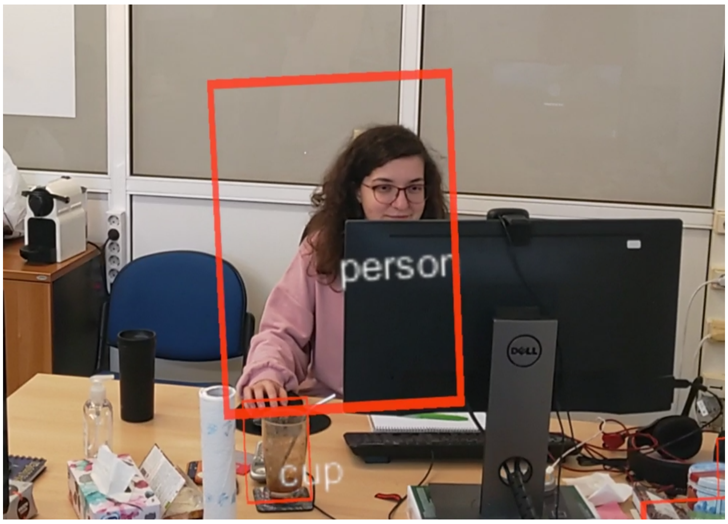

Figure 4. HoloLens 2 Visualisations

Firstly, the HoloLens app is requesting json from the laptop app multiple times per seconds, depending on the frame rate of the camera. The json data is being parsed, and used for the accurate placement of 2D boxes at 4 meters distance. In the same position a 3DText is placed with the proper label. The correct x and y axis placement of the bounding box is calculated by the intrinsics of the GoPro camera and extrinsics of the relative position in relation to the HoloLens.

Furthermore, there are voice commands implemented, which when given will change the funcitonality of the laptop's app in real-time. Voice commands such as, "Denoise On/Off", "Detection On/Off", to allow the user to disable the detections if need be and/or see the difference that the denoising algorithm makes.

DEMONSTRATION AND EVALUATION

The presented system was tested and evaluated by FRs in three events featuring close-to-realistic conditions and organized by associated FR organizations. These early tests have served two main purposes: to test the efficiency and limits of the proposed solution in variable conditions, and to gather feedback from the intended future users of this technology. The tool is able to start operating within $1-2$ minutes as requested by the end-users (Requirement 2).

Each trial proceeded in four broad steps: 1) participating FRs were presented with the system and given some time to become familiar with its functions, voice commands, and AR visualizations; 2) the testing area was prepared, including the placement of several objects and persons, and the presence of non-toxic smoke, if applicable; 3) participating FRs used the system and noted its efficacy in detecting objects, people and parts of people in different distances and smoke densities; and 4) the participating FRs gave their feedback to the researchers, discussing the system's performance, clarity of visualizations, usability, bugs, and wishes for future features.

The first field trial was conducted in Greece, prior to the system's integration with the HoloLens. It held in an abandoned quarry regularly used by local FRs for drills. The de-smoking and object detection algorithms were tested during one of such drills, with over 20 FRs participating, including both men and women. After objects and people were placed in a room, it was filled with smoke several pictures of it were acquired. The noisy images were processed by the proposed smoke-removal algorithm, and both the original and the enhanced images were used as input to the YOLOv5s object detection algorithm.

(a) **(b)**

Figure 5. Object detection during the field trial in Greece: object detections (a) without desmoking and (b) after desmoking

As shown in the example in Fig. 5 the object detection exhibited a significant improvement in performance, as the de-noised image accurately detects the person, increases the confidence of the detection of the bottle and successfully detects the cup.

The second field trial took place in Germany, in the context of a larger exercise involving more than 50 FRs. In this test, the system was complete and integrated with the HoloLens 2. 11 men and 9 women FRs, including both local volunteers and visiting personnel from other EU countries, tested the system over two days, in smoke-filled rooms, with varying intensities of daylight. Activities included both individual tests and the use of the system in short search-and-rescue scenarios. Fig. 6 shows person detection before and after de-smoking during testing.

Finally, the system was tested in an outdoor scenario in Spain. These tests did not include the presence of smoke, focusing rather on the accuracy of person and object detection in various distances, orientations, lighting conditions, and occlusion. 13 men and 6 women FRs took part in these tests.

Based on our tests we found that there is a distance threshold that the tool can detect people reliably, around $20-30m$, failing to complete the Requirement 1. That is mainly because of the wide field-of-view of the GoPro Camera. Furthermore, we performed tests based on visible parts of people, such a limb, and the algorithm did manage to detect it in most cases.

(a

(b

Figure 6. Object detection (a without de-smoking and (b with de-smoking

Feedback and Difficulties

Feedback from all three trials was generally positive. Suggested improvements concerned both visual and functional aspects. Desired improvements on visual elements included clearer bounding boxes for object detection, fewer overlaps, and customized colors, among others. Regarding functionality, FRs wished for detections of some additional object classes, relevant to the use cases of their particular units. Such included discarded articles of clothing (which may help track a missing person), obstacles, openings, and footprints or ski-prints on the snow. In addition, object and person detection at larger distances was another common request.

The main difficulty we encountered in real case scenarios was the lighting conditions. In conditions of low light the images captured from the camera are too dark and we cannot get any information out of it. In terms of hardware, although our initial plans were to use only one device (the HoloLens) for capture images from its own built-in camera and perform any denoising and detection algorithm on it, this proved to be impossible as it is limited in many aspects, mainly computationally. Thus, we split the load to a GoPro camera and a laptop, performing only the visualisation on the HoloLens and failed to limit the weight and amount of hardware devices a FR will use in a mission. Furthermore, according to Requirements 4 and 5 all of the hardware devices are prone to overheating and are limited by their short battery autonomy especially in visual computing tasks that our tool requires.

CONCLUSION AND FUTURE WORK

Conclusion

The proposed system describes a standalone AR tool for increased situational awareness during response missions, especially in the presence of smoke. All system components – camera, processing laptop, and AR device – can be carried or worn by the user and connected wirelessly to each other, making the system suitable for independent use, regardless of connectivity or any other external parameters. The de-smoking algorithm has been shown to improve object detection in the presence of semi-transparent smoke both qualitatively (more accurate detections and fewer misclassifications) and quantitatively (increased confidence). The complete system has been tested and evaluated in near-realistic conditions by its intended future users, FRs involved in search-and-rescue missions. Both technical and usability feedback was collected from those tests, and will be used to drive future developments.

Future work

As this is a work in progress, research and development is continuing along several paths, attempting to address both technical performance and requirements acquired from user feedback. The following is a list of major points we will work on in the near future:

- **Depth Estimation**: FRs often have the need to know the exact distance of each detected object. Thus, as an additional step in our whole pipeline we could add the estimation of the absolute distance of each object of interest. This information can be further visualised on the AR interface with many different ways e.g. showing the distance as a number next to the bounding box, adding a circle underneath each object etc. A different approach to this problem could have also been the alteration of the existing 2D object detection to a 3D one, so as to provide cuboids instead of rectangles.
- **Specific Classes**: Since every object detection algorithm covers a more generic variety of objects, there is the need to create an object detector that focuses on the classes that are useful for a rescue mission (e.g. pieces of clothes dropped by the victim in the snow, ski traces, holes on the ground etc.).
- **Long Distance Detections**: FRs in rescuer operations in open spaces, such as mountains, have the need to be able to detect people and objects of interest in longer distances than the human eye can see. This requirement can be implemented with object detection algorithms and is one of the approaches viable for future work.
- **Better Visualisations**: We can add better visualisations and more options to provide FRs with visualisation of choice, e.g. adding different color bounding boxes for different objects or distances, or changing the shape of bounding box to match that of the physical object.
- **Additional Denoising Algorithms**: We have included as a preprocessing step of denoising in our proposed tool only a dehazing / desmoking algorithm. As a future work we are planning to include image denoising algorithms for additional weather and environmental conditions such as models that remove noise caused by rain, snow, low-light and high contrast conditions.
- **Optimisations**: The whole application runs in relatively low frame per second and not in real-time. With better optimisations the whole pipeline could reach real-time 30 fps or near real-time.

ACKNOWLEDGEMENTS

Acknowledgements temporarily hidden to preserve anonymity

This research has been supported by the European Commission funded program RESCUER, under H2020 Grant Agreement 101021836. The authors would like to thank the Hellenic Rescue Team (HRT - Greece), International Search And Rescuer (ISAR - Germany), the Medical Emergency Service of the Community of Madrid (SERMAS - Spain) and the Special High-Altitude Rescue Group (GERA - Spain) for hosting the field trials and providing valuable feedback to the proposed system.

REFERENCES

Berman, D., Avidan, S., et al. (2016). “Non-local image dehazing”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1674–1682.

Bhattarai, M., Jensen-Curtis, A. R., and Martínez-Ramón, M. (2020). “An embedded deep learning system for augmented reality in firefighting applications”. In: *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, pp. 1224–1230.

Cai, B., Xu, X., Jia, K., Qing, C., and Tao, D. (2016). “Dehazenet: An end-to-end system for single image haze removal”. In: *IEEE Transactions on Image Processing* 25.11, pp. 5187–5198.

Chi, L., Jiang, B., and Mu, Y. (2020). “Fast fourier convolution”. In: *Advances in Neural Information Processing Systems* 33, pp. 4479–4488.

Christaki, K., Tsiakmakis, D., Babic, I., Guillaume, I., Konstantoudakis, K., Giunta, G., Dimou, A., Balet, O., and Daras, P. (2022). “Augmented Reality Points of Interest for Improved First Responder Situational Awareness”. In: *Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM), Tarbes, France*, pp. 755–770.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., and Schiele, B. (2016). “The cityscapes dataset for semantic urban scene understanding”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3213–3223.

Dudhane, A. and Murala, S. (2019). “Cdnet: Single image de-hazing using unpaired adversarial training”. In: *2019 IEEE Winter conference on applications of computer vision (WACV)*. IEEE, pp. 1147–1155.

Endsley, M. R. (1995). “Toward a Theory of Situation Awareness in Dynamic Systems”. In: *Human Factors* 37.1, pp. 32–64.

Engin, D., Genç, A., and Kemal Ekenel, H. (2018). "Cycle-dehaze: Enhanced cyclegan for single image dehazing". In: *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 825–833.

Fu, M., Liu, H., Yu, Y., Chen, J., and Wang, K. (2021). "Dw-gan: A discrete wavelet transform gan for nonhomogeneous dehazing". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 203–212.

Girshick, R. (2015). "Fast r-cnn". In: *Proceedings of the IEEE international conference on computer vision*, pp. 1440–1448.

Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). "Rich feature hierarchies for accurate object detection and semantic segmentation". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 580–587.

He, K., Sun, J., and Tang, X. (2010). "Single image haze removal using dark channel prior". In: *IEEE transactions on pattern analysis and machine intelligence* 33.12, pp. 2341–2353.

Hong, M., Xie, Y., Li, C., and Qu, Y. (2020). "Distilling image dehazing with heterogeneous task imitation". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3462–3471.

Hu, D., Chen, L., Du, J., Cai, J., and Li, S. (2022). "Seeing through Disaster Rubble in 3D with Ground-Penetrating Radar and Interactive Augmented Reality for Urban Search and Rescue". In: *Journal of Computing in Civil Engineering* 36.5, p. 04022021.

Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2017). "Image-to-image translation with conditional adversarial networks". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134.

Kyrkou, C. and Theocharides, T. (2019). "Deep-Learning-Based Aerial Image Classification for Emergency Response Applications Using Unmanned Aerial Vehicles." In: *CVPR workshops*, pp. 517–525.

Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., and Wang, Z. (2018). "Benchmarking single-image dehazing and beyond". In: *IEEE Transactions on Image Processing* 28.1, pp. 492–505.

Li, B., Gou, Y., Liu, J. Z., Zhu, H., Zhou, J. T., and Peng, X. (2020). "Zero-shot image dehazing". In: *IEEE Transactions on Image Processing* 29, pp. 8457–8466.

Lygouras, E., Santavas, N., Taitzoglou, A., Tarchanidis, K., Mitropoulos, A., and Gasteratos, A. (2019). "Unsupervised human detection with an embedded vision system on a fully autonomous UAV for search and rescue operations". In: *Sensors* 19.16, p. 3542.

McCartney, E. J. (1976). "Optics of the atmosphere: scattering by molecules and particles". In: *New York*.

Qin, X., Wang, Z., Bai, Y., Xie, X., and Jia, H. (2020). "FFA-Net: Feature fusion attention network for single image dehazing". In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 34. 07, pp. 11908–11915.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You only look once: Unified, real-time object detection". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788.

Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., and Yang, M.-H. (2016). "Single image dehazing via multi-scale convolutional neural networks". In: *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*. Springer, pp. 154–169.

Ren, W., Ma, L., Zhang, J., Pan, J., Cao, X., Liu, W., and Yang, M.-H. (2018). "Gated fusion network for single image dehazing". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3253–3261.

Sainidis, D., Tsiakmakis, D., Konstantoudakis, K., Albanis, G., Dimou, A., and Daras, P. (2021). "Single-handed gesture UAV control and video feed AR visualization for first responders". In: *Proceedings of the International Conference on Information Systems for Crisis Response and Management (ISCRAM), Blacksburg, VA, USA*, pp. 23–26.

Sakaridis, C., Dai, D., and Van Gool, L. (2018). "Semantic foggy scene understanding with synthetic data". In: *International Journal of Computer Vision* 126, pp. 973–992.

Simonyan, K. and Zisserman, A. (2014). "Very deep convolutional networks for large-scale image recognition". In: *arXiv preprint arXiv:1409.1556*.

Snyder, L. S., Karimzadeh, M., Stober, C., and Ebert, D. S. (2019). "Situational awareness enhanced through social media analytics: A survey of first responders". In: *2019 IEEE International Symposium on Technologies for Homeland Security (HST)*. IEEE, pp. 1–8.

Tan, R. T. (2008). "Visibility in bad weather from a single image". In: *2008 IEEE conference on computer vision and pattern recognition*. IEEE, pp. 1–8.

Xu, J. and Zheng, C. (2021). "Linear semantics in generative adversarial networks". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9351–9360.

Zeman, K., Stusek, M., Masek, P., Mlynek, P., and Hosek, J. (2022). "Augmented Reality-Aided Indoor Localization and Navigation for Next-Gen Rescue Services". In: *2022 14th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. IEEE, pp. 202–206.

Zhu, Q., Mai, J., and Shao, L. (2014). "Single image dehazing using color attenuation prior." In: *BMVC*.