FramCo: Frame corrupted detection for the Open RAN intelligent controller to assist UAV-based mission-critical operations

Ciro J. A. Macedo [**Federal Institute of Goiás** | *ciro.macedo@ifg.edu.br*]
Elton V. Dias [**Federal University of Goiás** | *eltondias@inf.ufg.br*]
Cristiano Bonato Both [**University of Vale do Rio dos Sinos** | *cbboth@unisinos.br*]
Kleber Vieira Cardoso [**Federal University of Goiás** | *kleber@ufg.br*]

✉ *Postgraduate Program in Computer Science, Universidade Federal de Goiás*
Alameda Palmeiras, Quadra D, Câmpus Samambaia, Goiânia, GO, 74.690-900, Brazil.

Abstract Unmanned Aerial Vehicles (UAVs) and communication systems are fundamental elements in Mission Critical services, such as Search and Rescue (SAR) operations. UAVs can fly over an area, collect high-resolution video information, and transmit it back to a ground base station to identify victims through a Deep Neural Network object detection model. However, instabilities in the communication infrastructure can compromise SAR operations. For example, if one or more transmitted data packets fail to arrive at their destination, the high-resolution video frames can be distorted, degrading the application performance. In this article, we explore the relevance of computer vision application information, complementing the functionalities of Radio Access Network Intelligent Controllers for managing and orchestrating network components, through FramCo - a frame corrupted detection based on EfficientNet. Another contribution from this article is an architectural element that explores the components of the Open Radio Access Network (O-RAN) standard specification, with an assessment of a complex use case that explores new market trends, such as SAR operations assisted by UAV-based computer vision. The experimental results indicate that the proposed architectural element can act as an external trigger, integrated into the O-RAN cognitive control loop, significantly improving the performance of applications with sensitive Key Performance Indicators (KPIs).

Keywords: Communication networks, Artificial Intelligence, UAVs, Mission-critical operations.

1 Introduction

The financial impact generated by natural disasters on the global economy is considerable. In addition to economic losses, disasters can impact wildlife and reap thousands of human lives [Our World in Data, 2024]. Immediately after a disaster occurs, Search and Rescue (SAR) operations begin. These operations are critical because time is vital and any delay can result in severe consequences, e.g., loss of human lives [Roldan *et al.*, 2019]. Unmanned Aerial Vehicles (UAVs) can perform SAR missions autonomously, thus, reducing human demand. UAVs are agile, fast, have low operating costs, and can be used to fly over an area where the victims can be located. UAVs can collect high-resolution video information and transmit it to the vehicular ground Base Station (BS). Some UAV models [Da Jiang Innovations, 2024] perform real-time video capture and transmission with a maximum rate of 80 Mbps and a frame rate of 30 frames per second (fps) under favorable communication distances and conditions. In this context, the video content collected by UAVs can be processed by software solutions based on Artificial Intelligence (AI) and Machine Learning (ML) techniques, which run as edge computing services, as illustrated in Figure 1.
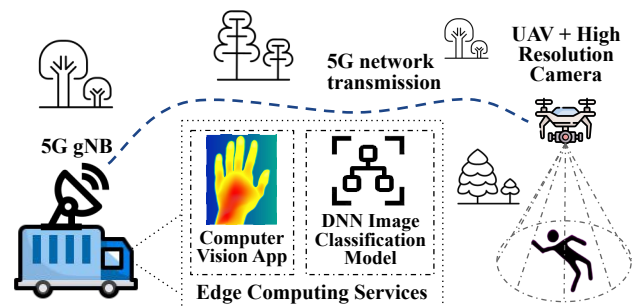
Communication systems are also essential in SAR opera-

Figure 1. SARs assisted by UAVs.

tions, allowing the integration of UAVs and improving these operations [Macedo *et al.*, 2022a]. However, some technological challenges need to be overcome. Current communication networks do not support the demands of highly dynamic scenarios such as SAR operations integrated with UAVs. This situation worsens in disaster scenarios, where the network infrastructure may be affected and often does not allow continuous communication between several teams, including UAVs. The adaptability of the network has been investigated, considering 5th Generation (5G) and Beyond 5G (B5G) scenarios through AI as a managing and orchestrating tool that makes up the communication infrastructure [Both *et al.*, 2022]. Moreover, standardization efforts toward AI-

powered communication networks are being disseminated by different organizations such as 3rd Generation Partnership Project (3GPP), European Telecommunications Standards Institute (ETSI), and Open Radio Access Network (O-RAN) Alliance [3GPP-TR22.125, 1 09; Wang *et al.*, 2020; O-RAN Working Group 2, 1 10].

This work explores the features of the Artificial Intelligence/Machine Learning (AI/ML) components that integrate the RAN Intelligent Controllers (RICs) [O-RAN Working Group 1, 2 03] specified in the O-RAN reference architecture [Polese *et al.*, 2022]. We consider a scenario involving User Equipment (UE) as a UAV connected to an O-RAN ground vehicle BS, collecting high-resolution video in real-time and transmitting it to a Deep Neural Network (DNN) object detection model [O-RAN Working Group 1, 2 03]. In this scenario, external application information is essential to support the O-RAN AI/ML components. For example, the quality of the video frames that arrive at the DNN object detection model is fundamental to supporting critical mission operations assisted by UAVs. To the best of our knowledge, only our previous paper Macedo *et al.* [2022a] presented an initial approach to the problem. However, it lacks a component that can act as an external trigger, integrated into the O-RAN cognitive control loop, with the aim of improving the performance of applications with sensitive Key Performance Indicators (KPIs), such as SARs operations.

In this article, we propose FramCo, a frame corrupted detection for the intelligent controller O-RAN that assists UAV-based mission-critical operations. The frame corrupted detection is based on DNN image classification model with a Convolutional Neural Networks (CNNs) architecture called EfficientNet [Tan and Le, 2020], and is used as an external trigger integrated into the O-RAN cognitive control loop, with the aim of improving the performance of UAV-based computer vision applications in SARs operations [Macedo *et al.*, 2022a]. By acting as a complementary element to UAV-based computer vision applications, FramCo assumes the same requirements of computer vision applications. In this sense, the UAVs involved must be able to capture and transmit 4K images resolution at 30 fps. Factors such as flight time, speed, and autonomy are relevant for SARs operations but do not impact FramCo's functionality, since it is not running on the UAV. FramCo operates on a nearby infrastructure located at the edge, as illustrated in Figure 1.

The main contributions of this work are: (i) show the feasibility of the O-RAN AI pipeline to support mission-critical operations assisted by UAV; (ii) analyze the information quality provided by external applications in the internal operation of the O-RAN AI pipeline; (iii) explore the usage of DNN to enhance O-RAN components providing support to an AI application in SAR operation. FramCo exhibits at least 90% accuracy in classifying corrupted frames, even under high packet loss.

The remainder of this article is organized as follows. Section 2 presents the related work concerning 5G/B5G, considering the main initiatives towards AI solutions, UAV, and SARs operations. Section 3 introduces O-RAN reference architecture explaining how our AI pipeline can enhance O-RAN components. Section 4 shows the experiments with AI for UAV-based SAR operations. Finally, conclusions and future work are discussed in Section 6. Moreover, Table 3 summarizes the commonly-used acronyms.

2 Related work

In recent years, telecommunications standardization bodies, e.g., O-RAN, ETSI, and 3GPP, have focused on specifications to make the 5G ecosystem more efficient and optimized. Much of the effort is towards AI and ML to deal with the complexity of new applications and use cases presented by existing market trends. Moreover, critical mission applications using UAVs have been widely investigated in the literature [Shule *et al.*, 2020] [Kulkarni *et al.*, 2020] [Hellaoui *et al.*, 2023], which must be supported by telecommunications infrastructure [3GPP-TR22.125, 1 09; Roldan *et al.*, 2019]. However, many applications have requirements beyond the capabilities of the elements that make up these infrastructures [Wu *et al.*, 2021].

In Zhou *et al.* [2019], the authors present concepts about expanding the use of AI at the Edges of the Network, giving rise to the concept of Edge Intelligence (EI). In the article, the authors survey recent advances in EI, listing architectures, frameworks, and emerging key technologies for a deep learning model toward training/inference at the network edge. For the authors, one of the main challenges is the limited computing and storage resources at the network edge, which can affect the performance and accuracy of edge intelligence models. Still on the use of AI at the edge of the network, Xu *et al.* [2020] present a comprehensive survey that covers the architectures, challenges, and applications of Edge Intelligence. The work is structured based on four primary mainstays: edge caching, edge training, edge inference, and edge offloading. For each of them, the authors discuss the main problems from a practical perspective, present the techniques that can be adopted for each of these problems, and discuss the objectives of applying each of the techniques.

In Shule *et al.* [2020], the authors provide an overview of the challenges in aerial navigation and navigation involving multiple UAVs. The work highlights requirements associated with constant monitoring of the location of UAVs, where, depending on the scenario, there is a need for a high-precision solution, capable of providing real-time information in the order of tens of centimeters, or in some cases more extreme, in the order of centimeters. These demands are directly associated with the capacity of the communication infrastructure to provide robust and real-time connectivity, when we consider the use of UAVs in some more extreme scenarios, such as SAR operations. In this sense, Kulkarni *et al.* [2020] explore a scenario of a SAR operation, assisted by a single UAV, in an indoor environment, where the Global Position System (GPS) signals tend to not present satisfactory levels of reliability. The authors consider that the victims to be rescued have smart devices that emit Radio Frequency (RF) signals. Through the RF signals, they try to locate victims as quickly as possible, using reinforcement learning techniques.

In Hellaoui *et al.* [2023], the authors explore the challenges of providing multiple services in unmanned UAVs communication networks. The authors propose a solution

Figure 2. High-resolution frames: without distortion and boundary-box people and with packet loss conditions.

for resource allocation and UAV deployment to support Ultra-Reliable Low Latency Communications (uRLLC) and Enhanced Mobile Broadband (eMBB) for the Internet of Things (IoT) in a UAV-enabled aerial access network. The proposed solution is based on two interactive algorithms, aimed to optimize uRLLC and eMBB services, balancing of the effective rate and the transmission delay.

In Saif *et al.* [2023], the authors discusses aspects related to the role of UAVs in mission critical operations. The authors demonstrate that UAVs are relevant in disaster response, providing expanded coverage, fast deployment, improved connectivity, energy efficiency, and low cost solutions. The authors consider some aspects of 5G technology to be relevant, in particular its ability to provide improved performance, ultra-low latency, expanded coverage, resilient connectivity and energy efficiency. According to the authors, all these aspects are essential to improve communication capabilities and support an effective approach to the use of UAVs.

In Wu *et al.* [2021], the authors discuss concepts of deep learning-based UAV object detection and tracking methods. Due to efficient and adaptive data gathering UAVs capabilities, they arise as a point of interest in Computer Vision (CV) applications. According to the authors, deep learning-based object detection from UAV videos faces many challenges, such as image degradation, uneven object intensity, small object size, and real-time problems such as perspective specificity, background complexity, and scale. However, recent advances in wireless communication and information compression technology, enable a higher data rate and longer transmission distance, making almost no-delay data link transmission possible and so making the processing of images obtained from UAVs viable.

In Alawada *et al.* [2023], the authors propose a disaster and crisis management system using UAVs for smart cities to optimize UAV routes to reduce energy consumption and improve effectiveness in disaster response. The authors explore disaster-affected areas and direct UAVs to identify groups of victims using a swarm-based optimization algorithm and analyze performance metrics such as delay, throughput, and performance. While the work provides a solid foundation for using UAVs in smart city disaster and crisis management through the implementation of a swarm optimization algorithm to manage UAV paths efficiently, FramCo goes a step further by employing AI to detect frame corruption in video data, a common issue in UAV-transmitted video streaming

under poor connectivity conditions. This approach represents a significant step forward in the field of UAV communications, offering enhanced capabilities for disaster response and other critical applications compared to existing systems that primarily focus on optimizing UAV navigation and data collection without addressing the quality of the transmitted data.

Recent research proposes adopting DNN models, generating, as a consequence, a high throughput between edge and cloud networks in the context of UAVs-based computer vision applications. For example, Li *et al.* [2022b] present a set of use cases, providing an overview of deployable 5G network concepts, including architecture options, system performance analysis, and coexistence. The same authors Li *et al.* [2022a] provide an overview of service requirements for public safety mission-critical communications, identifying key technical challenges and explaining how 5G New Radio (NR) features evolved to meet the emerging safety critical needs. Using AI/ML at the edge of the network has become more robust with the O-RAN Alliance, which arises with the general objective of standardizing an architecture and a set of interfaces to Radio Access Network (RAN) [Polese *et al.*, 2022]. In this sense, FramCo makes use of the concepts presented in these works, demonstrating the feasibility of an AI-driven approach, which not only improves the robustness of data transmission from UAVs in challenging communication environments but also demonstrates a specialized solution that directly contributes to the effectiveness of SAR operations, highlighting the potential of AI to address challenges in the context of public safety communications supported by 5G.

In Bertizzolo *et al.* [2021], the authors address the challenge of enabling high data-rate uplink cellular connectivity for UAVs within the context of 5G O-RAN architectures, highlighting the specific issues that arise when UAVs, which are more likely to have line-of-sight (LoS) connections to base stations, engage in high data-rate transmissions (e.g., video streaming). According to the authors, this situation potentially leads to uplink inter-cell interference and performance degradation for neighboring terrestrial users. The authors propose a low-complexity, closed-loop control system designed for O-RAN architectures that jointly optimizes the UAVs spatial location and transmission directionality. Although the system proposed by the authors demonstrates the feasibility of supporting efficient video streaming, reducing the impact of uplink interference on the network, external

factors related to different use cases can influence the results presented in real scenarios. In this sense, an approach that directly integrates AI into the application and communication layers of the infrastructure, similar to that presented in FramCo, becomes extremely relevant.

Our earlier work Macedo *et al*. [2022a] explores AI and ML components of O-RAN specification, aiming to improve support for UAV-based computer vision applications in SARs operations via RICs. We emulate a scenario in which a certain number of data packets may fail to arrive at the destination, producing different levels of degradation in the high-resolution video frames. The left side of Figure 2, illustrates a high-resolution frame in its original state, without any distortion, the right side illustrates a corrupted frame, reconstructed with corrupted pixels after being transmitted through a communication infrastructure with instabilities. The results showed the correlation between instabilities in the communication infrastructure and the degradation of the DNN object detection model. Despite the promising results, there is a relevant gap assuming that the DNN object detection model has a specific objective, e.g., maximizing the number of detected victims. Moreover, a second issue of service metrics needs to be improved to correctly indicate the performance degradation in the UAV-based computer vision application to the AI and ML components from RICs.

While packet loss and throughput reduction naturally impact video quality negatively, it is not trivial to map this degradation in the service metric, i.e., the number of detected victims. Furthermore, our previous work did not explore how external information is incorporated into AI and ML components from RICs. O-RAN already describes this functionality but focuses on the network infrastructure, not specific services or applications. This article approaches these two issues.

Table 1 summarizes how the state-of-the-art directly relate to this work and compares the works according to the main elements for supporting UAV-based critical missions, i.e., UAVs, O-RAN, AI, computer vision application, cloud, and critical mission operations. Section 3 presents the O-RAN architecture and the incorporation of an external source of information to AI and ML components from RICs, focused on the context of improved support for UAV-based computer vision applications in SAR operations.

3   Frame corrupted detection for RIC to assist UAV-based mission-critical operations

The 5G specification introduces considerable improvements over previous generations. Features such as increased bandwidth, improved data transmission speed, and low latency allow for *near-real-time* communication of massive volumes of information [3GPP-TR22.125, 1 09]. These features are essential in UAV-based mission-critical operations, where data collection, analysis, and sharing in *near-real-time* are relevant to saving human lives. Moreover, the low latency of 5G significantly contributes to better control of UAVs, making them more precise when remotely operated by res-

cue teams.

We adopted the following assumptions for building the corrupt frame detection considering UAV-based mission-critical operations: (i) SAR operation as a critical mission; (ii) application requirements associated with KPIs of a DNN object detection model based on high-resolution real-time video; (iii) UAV-based high-resolution video stream; (iv) DNN image classification model for application monitoring through frame corrupted detection, and (v) 5G O-RAN networks specifications, due to its flexibility and components disaggregation.

Considering the five assumptions, the O-RAN reference architecture was structured based on four guidelines:

- **Virtualization.** Introduction of new components to manage and optimize the network infrastructure and its operations, covering edge-positioned systems as well as virtualization platforms.
- **Disaggregation.** Division of the BS into Central Unit (CU), Distributed Unit (DU), and Radio Unit (RU), following the 3GPP recommendations for the RAN segmentation.
- **Open interfaces.** The inclusion of open interfaces connecting different O-RAN architecture components enables competition between suppliers and allows interoperability between the CU, DU, and RU.
- **Intelligent, data-driven control through RICs.** Components programmability, allowing the execution of optimization routines that orchestrate the RAN.

Figure 3 illustrates our view of the interaction between the frame-corrupted detector and the components of the O-RAN reference architecture. UE Layer represents a UAV carrying a high-resolution camera, flying over an area, collecting high-resolution video information, and transmitting it back to vehicular ground BS. *O-RAN Layer* represents the reference architecture of O-RAN [Polese *et al*., 2022], acting in a mission-critical scenario. Edge Computing Service Layer contains a DNN object detection model for processing high-resolution video information and a DNN image classification model that integrates with the O-RAN architecture for providing control information about the object detection application.

The interactions between the reference architecture elements are also illustrated in Figure 3 through several control loops. The *control loop* labeled with purple color is responsible for updating UAV control with insights produced by third-party applications, i.e., *xApps* and *rApps* designed to provide value-added services to support the RAN optimization process. The *control loop* labeled with red color acts in the near-real-time RAN Intelligent Controller (near-RT RIC) with insights produced by the non-real-time RAN Intelligent Controller (non-RT RIC). The *control loop* labeled with green color performs Next Generation Node Base (gNB) optimization with insights produced by near-RT RIC, while the *control loop* blue does the same for eNB. Finally, the *control loop* labeled with yellow color reports information about application performance, which is performed through the X1 interface by the frame-corrupted detector to the *rApps*.

In the following, we present additional information about RICs, *rApps*, *xApps*, Service Management and Orchestra-

Article	UAV	O-RAN basic	AI	Computer Vision	Cloud	Critical Mission	O-RAN with Enrichment Information
[Zhou *et al.*, 2019]			✓		✓		
[Shule *et al.*, 2020]	✓		✓				
[Kulkarni *et al.*, 2020]	✓		✓	✓		✓	
[Hellaoui *et al.*, 2023]	✓				✓	✓	
[Wu *et al.*, 2021]	✓		✓	✓		✓	
[Li *et al.*, 2022b]	✓		✓			✓	
[Alawada *et al.*, 2023]	✓		✓			✓	
[Li *et al.*, 2022a]	✓			✓	✓		
[Polese *et al.*, 2022]		✓	✓		✓		
[Saif *et al.*, 2023]	✓				✓	✓	
[Bertizzolo *et al.*, 2021]	✓	✓	✓	✓			
[Macedo *et al.*, 2022a]	✓	✓	✓	✓	✓	✓	
This work	✓	✓	✓	✓	✓	✓	✓

Table 1. UAVs assisted by AI/ML and focused on critical mission operations over 5G infrastructures exploring computer vision applications.
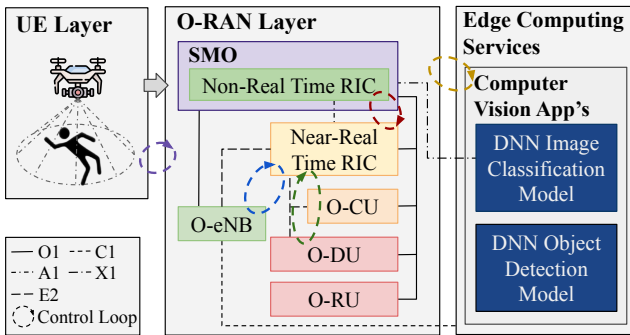
Figure 3. O-RAN reference architecture with mission-critical operations on computer vision applications.

tion (SMO), O-RAN open interfaces, the DNN object detection model, the DNNs image classification model, and how integrating these AI elements with intelligent controller O-RAN can improve enhanced support for mission-critical operations assisted by UAVs.

3.1 Service Management and Orchestration Framework

SMO handles all management, orchestration, and automation procedures to control RAN components. The non-RT RIC is part of SMO, as illustrated in Figure 3. Through the A1 and O1 interfaces, the SMO components interact with the other components, enabling the data collection to serve as input to AI/ML models. The outputs of the AI/ML models are intended to facilitate network monitoring and control [O-RAN Working Group 2, 1 10].

3.2 E2, O1, and A1 Interfaces

One of the O-RAN structuring guidelines was the inclusion of the E2, O1, and A1 open interfaces, connecting different architecture components and enabling interoperability among CU, DU, and RU. The E2 interface interconnects the CU and DU elements to the near-RT RIC. Through this interface, the near-RT RIC can collect RAN metrics and act in control procedures of the CU and DU elements [O-RAN Working Group 3, 2 03]. The O1 interface enables SMO to manage the lifecycle of O-RAN components. It is an interface focused on operation and maintenance activ-

ities, allowing initialization/configuration activities of components and performance assurance control [O-RAN Working Group 1, 1 02]. The A1 interface connects non-RT RIC and near-RT RIC. Through this interface, non-RT RIC can forward high-level optimization goals and manage ML models (e.g., deploy, update, or uninstall ML trained models used in *xApps*) [O-RAN Working Group 2, 1 07, 1 10].

3.3 Non-real-time RAN Intelligent Controller

The non-RT RIC is contained inside the SMO and is one of the core components of the O-RAN reference architecture. It was designed to handle the near-RT RIC, to support the execution of third-party applications, acting on control actions over the RAN, with timescales larger than 1 s. Furthermore, because it is located inside the SMO, non-RT RIC can influence the SMO operations and indirectly control all RAN components connected to SMO through the A1 and O1 open interfaces. The non-RT RIC is composed of three main elements, as illustrated in Figure 4: (i) Data management and exposure component are responsible for managing data and exposing services in the context of SMO, e.g., incorporating the internal information produced by O-CU/O-RU and the external information produced by FramCo to the AI/ML workflow; (ii) *rApps* are designed to provide value-added services to support the RAN optimization process; and (iii) AI/ML workflow is responsible for the first steps that make up the cognitive control loop. Determining how to position each AI/ML workflow element depends directly on the characteristics of the use case [O-RAN Working Group 2, 1 10]. Considering improved support for mission-critical operations assisted by UAV, the elements that make up the AI/ML workflow are distributed between non-RT RIC and near-RT RIC.

The non-RT RIC box at the top of Figure 4 presents the *rApps*, the non-real-time ML workflow, and the data management and exposure component. The near-RT RIC box at the bottom of the figure shows the *xApps* and the inference host responsible for real-time insights through ML models. O-CU/O-DU provides data for non-RT RIC ML workflow offline training, represented by the number 1 arrow. Moreover, FramCo provides external data for non-RT RIC ML workflow training, represented by number 2. O-CU/O-DU provide data for near-RT RIC ML online learning, represented by number 3. The near-RT RIC ML online learning
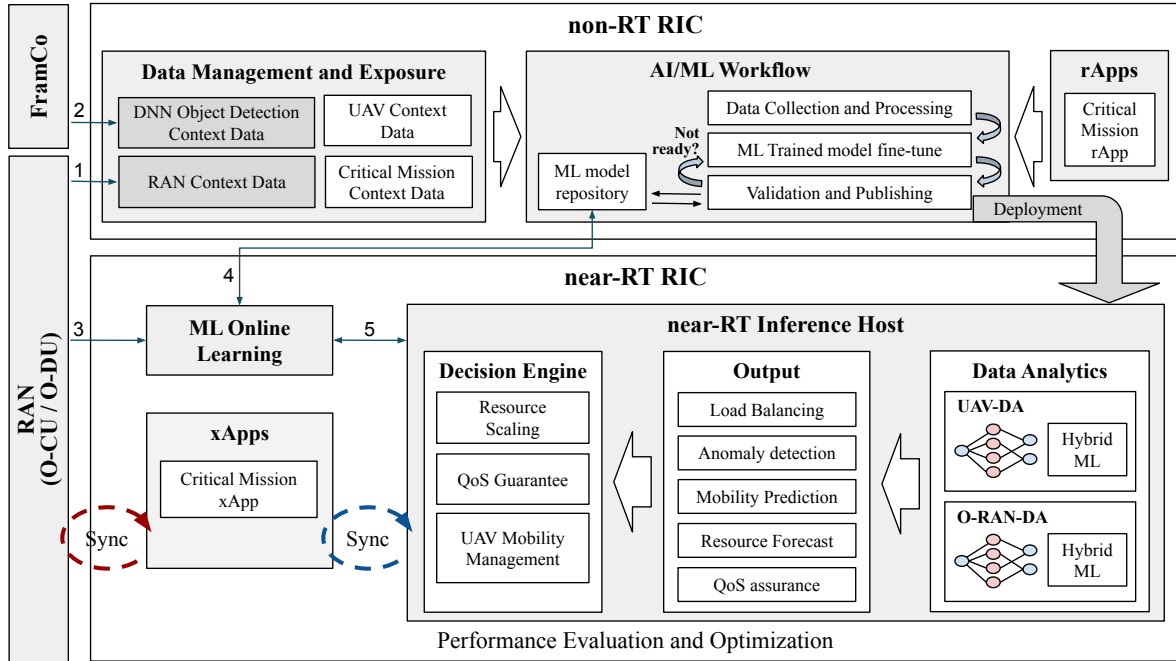
Figure 4. Interaction between the internal components of non-RT RIC and near-RT RIC improved support for mission-critical operations assisted by UAVs.

can interchange ML training models, stored at non-RT RIC ML model repository, represented by the number 4 arrow. The near-RT RIC ML online learning can deploy ML models and receive performance feedback information for online learning from the inference host, shown by number 5. The sync labeled with blue color provides ML decision insights to *xApps*. The sync labeled with red color is responsible for control action and guidance over RAN components.

The AI/ML non-RT RIC workflow is distributed into following four main steps:

- **Data collection/processing** - AI/ML models can be used to compose the cognitive control loop focus on improved support for mission-critical operations assisted by UAV. An appropriate AI/ML model to work with resource forecast may not be suitable for anomaly detection or Quality of Service (QoS) assurance. Each AI/ML model can use a specific data type with distinct granularity and collect over different time scales. The data collection/processing step obtains the data over the O1, A1, E2, and X1 interfaces, formats, and models these data according to the AI/ML model specifications.

- **AI/ML trained model fine-tune** - An AI/ML model needs to be trained and tested before being deployed in the network infrastructure. A completely untrained AI/ML model will not be deployed in the network [O-RAN Working Group 2, 1 10]. An AI/ML trained model fine-tune runs online fine-tune and updates AI/ML models previously trained offline.

- **Validation/publishing** - An AI/ML trained model must be validated to ensure reliability. If the validation step is successful, the AI/ML model becomes available in the ML model repository. However, if any anomaly is detected, the AI/ML model returns to the previous step.

- **Deployment** - A trained and validated AI/ML model can be deployed in containerized image format or

through an AI/ML model description file [O-RAN Working Group 2, 1 10]. Once the AI/ML model is deployed and activated, online data produced by the communication infrastructure elements will be used for inference and actuation in control procedures, which include: (i) 3GPP and Non-3GPP events across all different managed elements over O1 and E2 interfaces, and (ii) Enrichment information from non-RT RIC to near-RT RIC over A1 interface.

The network conditions could be highly dynamic, considering that this work involves detecting possible victims through high-resolution video analysis with DNN models and that this video is streamed in real-time by a UAV. The requirements for the AI application and the mission objectives may change over time. In this sense, it is the responsibility of mission-critical *rApp* to enrich the Performance evaluation and Optimization loop placed in near-RT RIC. This enables mission-critical *xApp* to act in near-real-time changes to RAN components in scenarios that can produce broad improvement in the SAR operation.

3.4 Near-real-time RAN Intelligent Controller

The near-RT RIC connects O1, A1, and E2 interfaces, in addition to host the *xApps* and the components required to operate and manage the *xApps*. We placed the ML inference host as part of near-RT RIC, considering improved support for mission-critical operations assisted by UAV and control of KPIs with low latency demand and uplink/downlink service asymmetry. After training and validation steps in non-RT RIC, ML models are deployed int the near-RT RIC inference host.

We consider the existence of a Data Analytics component inside near-RT RIC, which can contain several ML data an-

alytics models. Each component is trained and actuates in a specific context, such as working in a UAV context decision, e.g., navigation commands or flight status data reporting, or decision making associated with RAN control action and guidance. The outputs produced by ML Data Analytics components may correlate with the following actions: (i) Load Balancing, (ii) Anomaly Detection, (iii) Mobility Prediction, (iv) Resource Forecast, and (v) QoS Assurance. Considering improved support for SAR operations, these correlations are explored by the decision engine to generate actions aimed at RAN resource scaling, QoS targeted at SAR operation, and UAV mobility management. Finally, through a regular synchronization process, the insights generated by the near-RT RIC inference host stimulate *xApps* to actuate on the RAN elements, generating a cognitive control loop, able to actuate on the elements of the communication infrastructure to improve support for mission-critical operations assisted by UAV.

Data produced by O-RAN internal components are collected through the O1 interface and forwarded to non-RT RIC, as shown in Figure 5. Moreover, external data is collected through the X1 interface and used to train the AI/ML elements. This point appears out as a relevant contribution of this work. Considering a scenario, where the entire communication infrastructure must support an AI/ML application, collecting data about the health status of this application and using this data to fine-tunning AI/ML models, aimed at operating the communication infrastructure becomes quite relevant. This is provided by FramCo. Through the X1 interface, metrics about the quality of frames delivered to You Only Look Once (YOLO) are collected and used to train internal AI/ML internal elements. AI/ML workflow involves non-RT RIC and near-RT RIC. The AI/ML elements are trained, validated, and can be internally deployed and stored in non-RT RIC. These same trained elements can be forwarded and deployed in the inference host through the A1 or O1 interfaces to actuate in the performance evaluation and optimization loop. Inside the performance evaluation and optimization loop, data produced by O-RAN internal components are used for online learning. The training host continuously manages the AI/ML models. When it detects severe performance degradation, it can request the stored (previously well-performing) AI/ML model from the ML model repository in the non-RT RIC.

In the following, we discuss details of the DNN object detection model, the DNN image classification model, and how integrating these AI elements with intelligent controller O-RAN can improve enhanced support for mission-critical operations assisted by UAVs.

3.5 DNN object detection model

We use DNN YOLO v3 [Redmon and Farhadi, 2018] for object detection. The object detection task determines the location and the class of certain objects present in a particular video frame, i.e., in an image. In SAR operations, the "objects" of interest are potential victims to be rescued by the search teams. YOLO suits this task because it goes straight from the image pixels to the bounding box coordinates and probabilities of each object class involved.

YOLO divides each video frame into an $S \times S$ grid of cells. Another cell grid is responsible for predicting each object in the image. Each cell grid can also predict bounding boxes that delimit the detected objects, as illustrated in Figure 6. The bounding box prediction has the following components: (i) x and y coordinates representing the center of the box, (ii) w and h values corresponding to the bounding box dimensions that delimit the detected object, and (iii) the confidence value reflecting the presence or absence of an object of any of the class considered. YOLO can identify objects in an image and classify them according to predefined classes using the Intersection over Union (IoU) metric [Rezatofighi *et al.*, 2019].

The accuracy of tools such as YOLO depends on the quality of the images used as input. In our context, the Distortion Level (DL) in frames or even compression in the images provided to the tool can compromise the object detection process and negatively affect the SAR operations. These distortions can be caused by abnormal behavior in the communication infrastructure, which needs to be tracked. Nevertheless, it cannot count on YOLO to obtain information related to this issue.

3.6 DNN image classification model

We employed CNN EfficientNet [Tan and Le, 2020] as the main component to detect the corrupted frames. EfficientNet uses compound coefficients to scale up models effectively. Instead of randomly scaling up over the three dimensions, compound scaling scales all three dimensions while maintaining a balance between these dimensions of the network. Formally, a CNN can be defined as follows:

$$\mathcal{N} = \bigodot_{i=1...s} \mathcal{F}_i^{L_i} \left(X_{\langle H_i, W_i, C_i \rangle} \right), \qquad (1)$$

where i represents the stage number, $\mathcal{F}_i$ defines the convolution operation for the ith stage, and L_i represents the number of times $\mathcal{F}_i$ is repeated in stage i. H_i, W_i, and C_i denote the input tensor shape for stage i [Tan and Le, 2020]. Finding a set of proper coefficients to scale deep (L_i), width (C_i), and input resolution (H_i and W_i) is a complex task since the search space is huge. EfficientNet is formulated from two ground rules to restrict the search space: (i) all layers in the scaled models use the same convolution operations as the baseline network, and (ii) these layers must be scaled uniformly with a constant ratio. Considering these rules, Equation (1) can be adjusted as follows:

$$\mathcal{N}(d, w, r) = \bigodot_{i=1...s} \hat{\mathcal{F}}_i^{d \cdot \hat{L}_i} \left(X_{\langle r \cdot \hat{H}_i, r \cdot \hat{W}_i, w \cdot \hat{C}_i \rangle} \right), \qquad (2)$$

where w, d, and r are scaling width, depth, and resolution coefficients. $\hat{\mathcal{F}}_i$, $\hat{L}_i$, $\hat{H}_i$, $\hat{W}_i$, and $\hat{C}_i$ are predefined parameters in the baseline network and substitute the previous coefficients. EfficientNet uses a compound coefficient to uniformly scale network width, depth, and resolution, as fol-
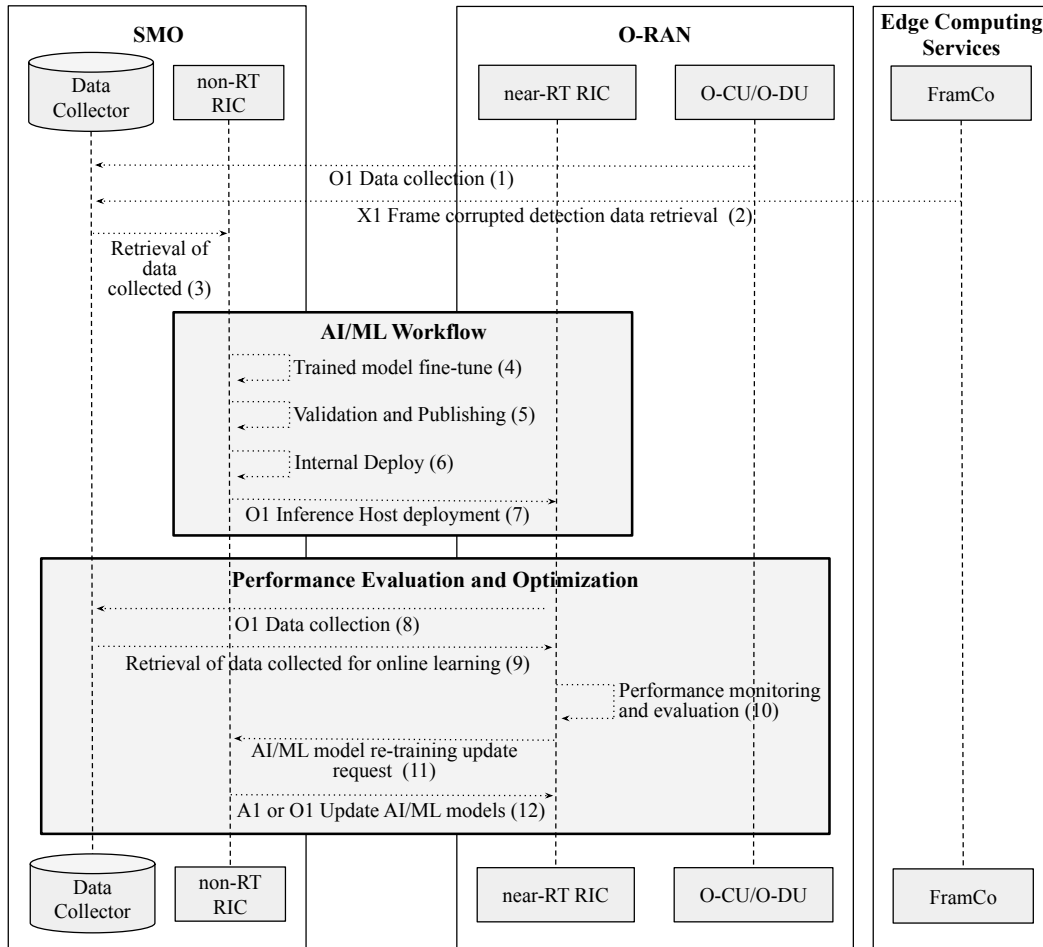
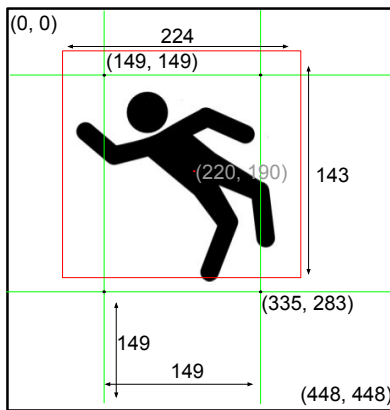**Figure 5.** AI/ML models training and distribution flow diagram.

Figure 6. YOLO bounding box for object detection.

lows:

$$depth: d = \alpha^\phi$$

$$width: w = \beta^\phi$$

$$resolution: r = \gamma^\phi \qquad (3)$$

$$subject\ to: \alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$$

$$\alpha \geq 1, \beta \geq 1, \gamma \geq 1,$$

where ϕ is an integer value representing a global scaling factor controlling the available resources. α, β, and γ determine how to assign these resources to depth, width, and resolution. Based on the definition of the frame corrupted detection for

RIC to assist UAV-based mission-critical operations, we describe the sets of experiments and their results acting on the communication infrastructure elements in the following.

4 Methodology and Results

We consider a SAR operation to demonstrate the usefulness and flexibility of the O-RAN elements, in which a UAV flies over an area where the victim is located, collects high-resolution video information, and transmits it back to a vehicular ground BS. Edge computing services are deployed at the ground BS, processing local video, and providing control information. DNN YOLOv3 [Redmon and Farhadi, 2018] is employed to act as an object detection application. Finally, we use DNN EfficientNet [Tan and Le, 2020] to analyze the frames' quality, identify possible distortions, and notify O-RAN AI/ML components. The methodology used to carry out the experiments and their results are discussed below.

4.1 Methodology

In our experiments, we employed the network configuration illustrated in Figure 7. This configuration comprises a Cloud node, O-RAN node, vehicular ground BS, and UE representing UAV emulated in a containerized Web server that streams

a high-resolution video through the network. Our vehicular ground BS node provides the communication, representing the RAN. The BS node is emulated by a virtual switch created by Open vSwitch (OVS) and connects UE to the O-RAN infrastructure. The cloud node represents where all the services and applications external to O-RAN are running. In our case, the cloud node hosts YOLOv3 and EfficientNet models. We employ a node from the Open-Access Research Testbed for Next-Generation Wireless Networks (ORBIT)[1] [Testbeds, 2024] equipped with a 10-core Xeon(R) E5-2640 CPU @ 2.40GHz and a Tesla P100 12GB GPU to ensure the requisite hardware capabilities. We also employed the Linux Traffic Control (tc) [Linux Foundation, 2024] tool to emulate different levels of packet loss in the experiments.
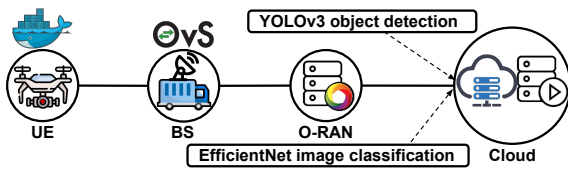

Figure 7. Network configuration used in the experiments.

In this work, the cloud node is composed of two elements: (i) a Darknet algorithm with YOLOv3 [Redmon and Farhadi, 2018] pre-trained model with Pascal VOC dataset [Everingham *et al.*, 2015] and (ii) an EfficientNet image classification model, trained using a large set of images (or video frames) obtained exclusively from UAVs [Macedo *et al.*, 2022b]. Each video frame belongs to two possible classes: i) images corrupted by packet loss in the communication infrastructure; and ii) original images without distortion. To generate corrupted images, we employed a Real Time Streaming Protocol (RTSP) server to stream each original frame in 4K resolution and artificially injected packet loss, ranging from 9% to 12%, using tc. This range of values was selected to cover the specific region of interest in our study, based on the visual results caused by packet loss. As we administered lower packet loss injections, distortions remained visually minor, gradually intensifying as we increased the percentages. Eventually, a significant portion of image content was lost, resulting in increasingly pronounced and disruptive distortions. Our approach made two versions of each video frame available: a corrupted one and the original one.

To improve the evaluation of FramCo's accuracy, we categorized the corrupted images into distinct levels. Thus, we employed the ImageHash Python library [Buchner, 2022] to compare each video frame transmitted by the communication infrastructure with the original equivalent high-resolution video frame. The results were categorized into five DLs. The first DL is the most sensitive, where minimal differences between the two frames are considered. The sensitivity is reduced as the DL increases, so fewer differences are considered.

In our experiment, we stream a 4K video recorded by a UAV, showing a group of people on the highway. The high-resolution images are more suitable for smaller object detection [Zhang *et al.*, 2021], and nowadays, there are several

models of UAVs properly designed to capture 4K images. We selected a 3-second video clip with a frame rate of 30 fps for assessing the performance of our DNN object detection model. To simulate a high-resolution video stream from a UAV, we utilized a containerized FFmpeg instance deployed on the UE and cloud nodes equipped with an RTSP server. The FFmpeg instance on the UE was responsible for streaming the video to the cloud nodes via the RTSP protocol. On the cloud node's end, the streaming data was received and frames were extracted for subsequent use in both YOLOv3 object detection and the FramCo classification process.

As previously described, application information may be useful to enrich the O-RAN cognitive control loop [Macedo *et al.*, 2022a]. In this work, this information is associated with the quality of the frames delivered to the DNN object detection model. This information is especially critical because corrupted frames can prevent the accurate detection of objects positioned within these frames during the image reconstruction process. Bandwidth, high latency, and packet loss are the most common factors contributing to increased corrupted frames [Nishio *et al.*, 2021]. Figure 8 shows the behavior of the YOLOv3 object detection in a scenario of packet loss.
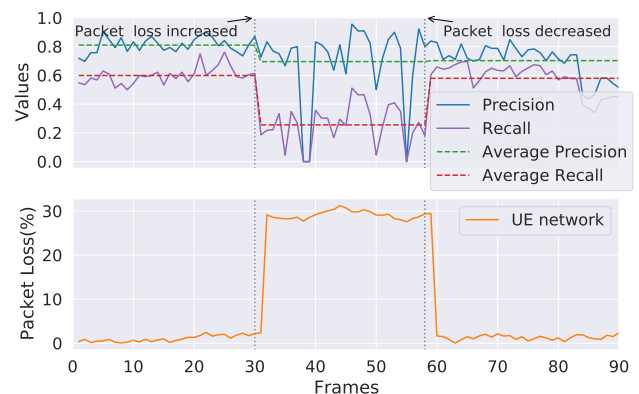

Figure 8. YOLO behavior in a scenario of packet loss.

The x-axis represents a video stream timeline (fps), and at a given moment, we note a disturbance in the average of the recall caused by an increased packet loss, and the values returned to the previous averages. Frames transmitted under normal communication infrastructure conditions (e.g., intervals 1 to 30 and 60 to 90 in Figure 8) are entirely reconstructed and delivered to YOLOv3 object detection in a format equivalent to that shown on the left side of Figure 2. However, frames transmitted under abnormal conditions of communication infrastructure (e.g., instabilities associated with packet loss, illustrated by frames in the range 31 to 59 in Figure 8) are delivered to YOLOv3 object detection in a format equivalent to that shown on the right side of Figure 2, considerably impacting the functional object detection area in the frame.
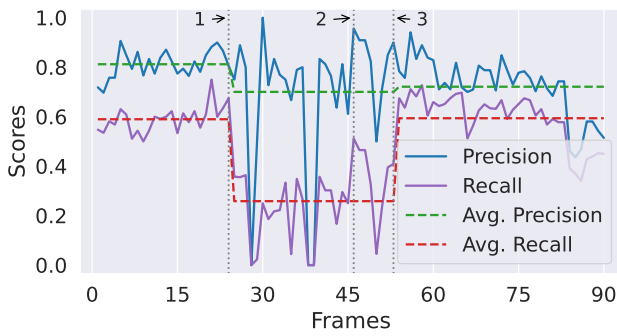
Figure 9. Complete emulated scenario with the high-resolution video stream and FramCo acting to correct instabilities in the communication infrastructure.

Figure 9 illustrates a complete emulated scenario where FramCo employs the EfficientNet image classification model to assess the quality of frames delivered to the YOLOv3-based system, communicating any findings to the O-RAN AI/ML components that work to reestablish the communication infrastructure. Label 1 illustrates the moment when the communication infrastructure has an increase in packet loss. This abnormal behavior causes the values associated with YOLOv3 recall measures to be reduced considerably. The same frames analyzed by YOLOv3 are also classified by FramCo, which identifies the anomaly at the moment indicated by Label 2, communicating any findings to the O-RAN AI/ML components that work to reestablish the communication infrastructure, indicated by Label 3. In the experiments, the average time between the beginning of the abnormal behavior of the communication infrastructure and the detection of corrupted frames, with consequent activation of the O-RAN AI/ML components by FramCo, is approximately 800 ms, which is an acceptable value considering a SAR operation scenario and also considering performance of DNN inference workloads [Liang *et al.*, 2023]. Table 2 summarizes the main parameters used in the experiments and their respective values. The GitHub repository [2] provides additional details, including all configuration parameters used in YOLOv3 and EfficientNet (e.g., initial weights, initial and final learning rates, moving average decay, and classes).

Parameter description	Value
Object Detection Algorithm	YOLOv3 [Redmon and Farhadi, 2018]
Object Detection Dataset	Pascal VOC [Everingham *et al.*, 2015]
Image Classification Algorithm	EfficientNet [Tan and Le, 2020]
Image Classification Dataset	UAV Images Packet Loss Distortions [Macedo *et al.*, 2022b]
Video Stream Duration	3 seconds
Video Stream Frame Rate	30 fps
Video Stream Resolution	4K (3840 x 2160 pixels)

**Table 2.** Parameters values used in the experiments.

4.2 Results

In Figure 10, we present the scores of recall and precision reported by YOLOv3 along an experiment under normal conditions, i.e., with the communication infrastructure presenting

[2]https://github.com/LABORA-INF-UFG/paper-CECK-2024

no packet loss. The X axis represents the evaluated frame number and the Y axis represents the score value obtained by YOLOv3 evaluation, concerning recall and precision. The recall represents the true positive identification reported for a given frame under normal conditions, and precision quantifies the number of positive class predictions that belong to the positive class. The average recall value near 0.6 is acceptable because the frames were obtained from a UAV at a high altitude. This characteristic represents a difficult scenario for the detection of small objects [Zhang *et al.*, 2021].
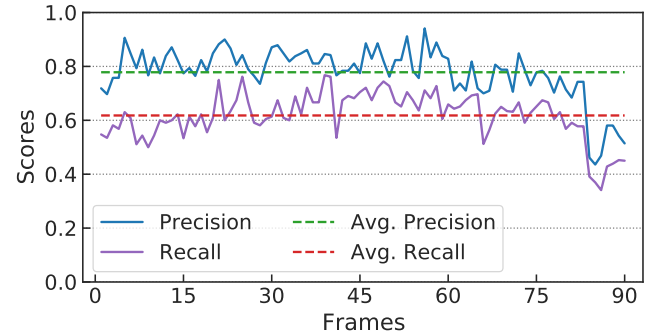
Figure 10. YOLO behavior in no packet loss conditions - Original frames.
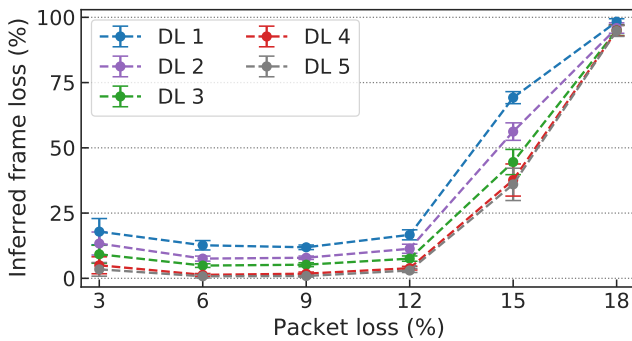
We assume that the performance of YOLOv3 presented in Figure 10 is the best available, since there is no packet loss. The impact of the packet loss is shown in Figure 11, which presents the average recall and precision under degraded network conditions. In Figure 11, the X axis represents the percentage of packet loss in the communication infrastructure and the Y axis represents the score value obtained by YOLOv3 evaluation, concerning recall and precision. The performance is quite stable until a certain level, achieving 12% in our experiments, but there is a fast degradation beyond this point. This behavior illustrates the performance degradation of an AI object detection application presented in Figure 8, which may be correlated with the quality of the frames transmitted by the communication infrastructure. In our experiment, the majority of the frames transmitted under packet loss above 12% were corrupted, looking similar to right side of the Figure 2. Such degradation caused a lot of frame information to be lost and made it impossible for YOLO to identify people in the image. Moreover, these results show that FramCo can be used as an external trigger integrated to the O-RAN cognitive control loop.
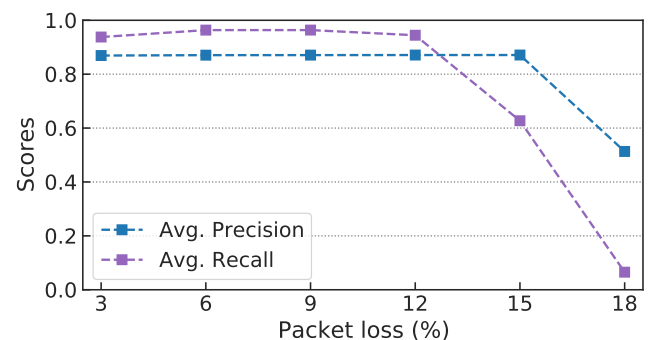
Figure 11. YOLO behavior in no packet loss conditions - Distorted frames.
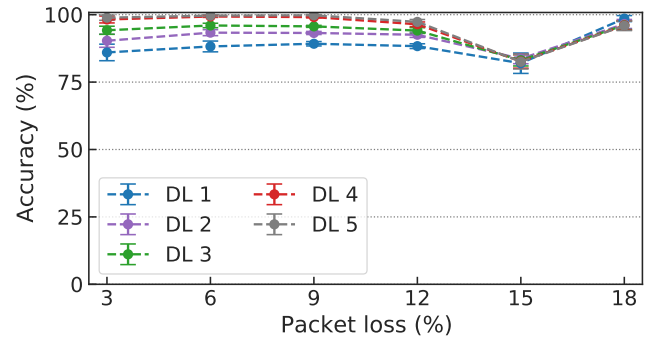
In Figure 12, we present the inferred frame loss according to the ImageHash as a function of the packet loss. The Y axis represents the value obtained by ImageHash classification function and the X axis represents the percentage of packet loss in the communication infrastructure. The inferred frame loss is affected by the assigned DL value. Lower DLs (specifically, DL1 and DL2) may classify frames as lost even if they exhibit minimal corruption, while higher DLs (specifically, DL4 and DL5) classify frames as lost only when they are nearly completely corrupted. The DL3 exhibited the highest accuracy on classifying visually corrupted images.

Figure 13. FramCo accuracy in an incremental packet loss scenario.

5 Discusions

The combination of AI, UAVs, and 5G networks, aiming to improve support in SAR operations, represents an advance in terms of efficiency and effectiveness. Our work highlights an application of AI, based on YOLOv3, which analyzes frames from a high-resolution video transmitted through the 5G communication infrastructure, aiming to identify evidence of the presence of victims and report the information collected to a remote ground station. Moreover, AI is integrated into the communication infrastructure and optimizes the infrastructure's elements through EfficientNet [Tan and Le, 2020]. Considering the use of 5G communication infrastructure acting in a SAR operation context as the main scope of our work, we simplify the AI application scenario that makes use of the 5G infrastructure through the use of a Darknet algorithm with YOLOv3 [Redmon and Farhadi, 2018] pre-trained model with Pascal VOC dataset [Everingham *et al.*, 2015] and detecting potential victims from images obtained by UAVs.

Figure 12. ImageHash categorization of frames transmitted by the communication infrastructure and delivered to the YOLOv3.

Figure 13 presents the FramCo accuracy as a function of the packet loss. Each frame is delivered to YOLOv3 for object detection and to FramCo for classification into two possible classes: (i) corrupted or (ii) not-corrupted. The accuracy was determined by comparing the computed results to the ground truth, which was established using the inferred frame loss reported by the ImageHash tool. Specifically, a result was considered correct when it was the same as the outcome provided by ImageHash. The accuracy is at least 90% for most DLs until a packet loss of 12% in our experiments. The worst accuracy of FramCo is 80% in the presence of a packet loss of 15%, which is also the point of highest divergence among DLs, as shown in Figure 12. This can be partially related to the training dataset, which is mainly composed of images corrupted with packet loss levels ranging from 9% to 12%. Therefore, FramCo exhibits a satisfactory performance even in the most difficult scenario. At higher packet loss rates, such as 18%, we observe DLs converging because FramCo's classification task becomes simpler. This occurs because the frames are nearly 100% corrupted, as indicated by all DLs in Figure 12.

FramCo can deliver relevant information to the SMO through the X1 interface in the scenario presented in the experiments. AI and ML components from RIC can act to improve the communication infrastructure by correlating the application metrics with the other traditional network metrics, such as packet loss and delay. Moreover, the results show the impact of the communication infrastructure in critical mission scenarios in which time is relevant, and performance penalties can result in severe consequences. In this context, adopting AI/ML is critical, but it is far from trivial, as illustrated.

Although the approach of simplifying the AI application that uses 5G infrastructure shows promise from the point of view of evaluating the behavior of the communication infrastructure, the results presented in Figure 10 exhibit limitations in detecting potential victims, mainly in terms of the low recall observed in the experiments. This outcome can be attributed to a classic computer vision problem associated with the AI model's difficulty in small object detection [Zhang *et al.*, 2021] in images captured from high altitudes, a common scenario in SAR operations. In this context, the discussion about the need for YOLO fine-tuning, specifically trained with images obtained by UAV, is quite relevant from the perspective of improving computer vision models. Given that the main objective of our work is to demonstrate the interaction between 5G communication infrastructure and AI, an AI model more aligned with real-world conditions and SAR operations should significantly improve victim detection performance. Even though helpful, using YOLOv3 [Redmon and Farhadi, 2018] pre-trained model with the Pascal VOC dataset should not be the ideal solution for the specificity and complexity of images captured by UAVs at different altitudes and conditions.

Additionally, the diversity of UAV models, each with unique characteristics, e.g., from fixed wings to rotary wings and the ability to fly at low or high altitudes [Zeng *et al.*, 2016], can influence the quality of captured and transmitted

frames. In this context, building datasets that reflect this diversity and the variety of SAR scenarios, e.g., SAR operations in mountains or offshore, can significantly improve the performance of computer vision models, such as those used in this work. These datasets allow training AI models capable of considering specific visual details of each situation, reducing the incidence of false positive detections and improving detection reliability. Exploring the construction of specific datasets for training AI models to act in SAR operations represents a promising work front. These datasets can significantly enhance support for UAV-based computer vision applications in SAR operations by considering the diversity of UAV platforms and SAR scenarios. Moreover, these datasets should improve detection results in SAR operations and serve as an integral part of the evolution of future communications infrastructure, where AI plays a central role.

Finally, given the potential for improvement in detection outcomes and the importance of integrating AI with the infrastructure of future communication networks, deepening studies to adapt FramCo to other scenarios through the refinement of AI for specific applications in UAVs emerges as a promising direction. Such an approach can expand FramCo's scope of action, expanding the potential for applying these technologies in several other scenarios while respecting the necessary adaptations. The combination of UAVs, computer vision, and the robustness of 5G communications infrastructure opens up many possibilities beyond SAR missions (e.g., environmental and agricultural monitoring, urban security, and inspection of critical infrastructures). UAVs equipped with computer vision systems, supported by the high-speed and low-latency data transmission capabilities of 5G networks, can provide real-time information for quick decision-making. Therefore, adapted versions of FramCo, especially associated with developing specific datasets, improve the effectiveness of SAR operations and establish a solid foundation for future innovations in other applications.

6   Conclusion

In this work, we explored the functionalities of RAN intelligent controllers, as described in the O-RAN standard specification. We propose to extend the O-RAN architecture in the context of critical mission operations assisted by UAVs. This extension includes a new component for providing external information to enhance AI/ML capabilities, which is also AI/ML-based but focused on application health. The experiments illustrate the benefits of our proposal, mainly the high accuracy in detecting the effective impact of the communication degradation in the application performance. Moreover, we showed how to integrate this information into the O-RAN cognitive control loop and its potential benefits. The preliminary results demonstrate the difficulties found during SAR operations, particularly when UAVs are use for assistance. The results also highlight the potential benefits of employing intelligent edge strategies in alignment with the O-RAN architecture proposal.

As future work, we are interested in evaluating more challenging scenarios with multiple UAVs, including an online adaptation of the communication infrastructure. Further-

more, we are interested in investigating the benefits of the interaction among intelligent components operating at various points within the emerging communication infrastructures.

Declarations

Acknowledgment

Author's Contributions

Conceptualization, Macedo, C.J.A; Both, C.B. and Cardoso, K.V.; methodology Macedo, C.J.A; Dias, E.V; Both, C.B. and Cardoso, K.V; software Macedo, C.J.A. and Dias, E.V; validation Macedo, C.J.A; Dias, E.V; Both, C.B. and Cardoso, K.V.; writing–original draft preparation Macedo, C.J.A. and Dias, E.V; writing–reviewand editing, Macedo, C.J.A; Dias, E.V; Both, C.B. and Cardoso, K.V.

Competing Interests

The authors declare that they have no competing interests.

References

3GPP-TR22.125 (2021-09). Unmanned Aerial System (UAS) support in 3GPP. Technical report, 3rd Generation Partnership Project (3GPP). Available at: `https://www.3gpp.org/` Version 17.0.4.

Alawada *et al.* (2023). An unmanned aerial vehicle (uav) system for disaster and crisis management in smart cities. *Electronics*, 12(4):1051. DOI: 10.3390/electronics12041051.

Bertizzolo *et al.* (2021). Streaming from the air: Enabling drone-sourced video streaming applications on 5g open-ran architectures. *IEEE Transactions on Mobile Computing*, 22(5):3004–3016. DOI: 10.1109/TMC.2021.3129094.

Both *et al.* (2022). System Intelligence for UAV-Based Mission Critical with Challenging 5G/B5G Connectivity. *ITU Journal on Future and Evolving Technologies*, 2. DOI: 10.48550/arXiv.2102.02318.

Buchner, J. (2022). Image Hashing library. Available at `https://github.com/JohannesBuchner/imagehash`. Accessed: Oct-19-2022.

Da Jiang Innovations (2024). DJI Mavic 3. Available at: `https://www.dji.com/br/support/product/mavic-3`. Accessed: Jul-13-2024.

Table 3. Summary of acronyms.

Acronym	Definition	Acronym	Definition
3GPP	3rd Generation Partnership Project	5G	5th Generation
5GC	5G Core	AI	Artificial Intelligence
AI/ML	Artificial Intelligence / Machine Learning	B5G	Beyond 5G
BS	Base Station	CV	Computer Vision
CNN	Convolutional Neural Network	CUDA	Compute Unified Device Architecture
DNN	Deep Neural Network	DL	Distortion Level
DU	Distributed Unit	EI	Edge Intelligence
eMBB	Enhanced Mofbile Broadband	ENI	Experiential Networked Intelligence
ETSI	European Telecommunications Standards Institute	FANET	Flying ad-hoc network
fps	Frames per Second	GPU	Graphics Processing Unit
GPS	Global Position System	gNB	Next Generation Base Station
HTTP	Hypertext Transfer Protocol	IoT	Internet of Things
IoU	Intersection over Union	IQA	Image Quality Assessment
ITU-T	International Telecommunication Union	JSON	JavaScript Object Notation
KPI	Key Performance Indicator	LoS	line-of-sight
ML	Machine Learning	mMTC	massive Machine-type communication
mULLC	Massive Ultra Reliable Low Latency Communication	near-RT RIC	near-real-time RAN Intelligent Controller
non-RT RIC	non-real-time RAN Intelligent Controller	NR	New Radio
NWDAF	Network Data Analytics Function	OVS	Open vSwitch
ORBIT	Open-Access Research Testbed for Next-Generation Wireless Networks	O-RAN	Open Radio Access Network
QoE	Quality of Experience	QoS	Quality of Service
RAN	Radio Access Network	R-CNN	Region-Based Convolutional Neural Network
RIC	RAN Intelligent Controller	RF	Radio Frequency
RU	Radio Unit	RTSP	Real Time Streaming Protocol
SAR	Search and Rescue	SMO	Service Management and Orchestration
uRLLC	Ultra-Reliable Low Latency Communications	UAV	Unmanned Aerial Vehicle
UE	User Equipment	VOC	Visual Object Classes
YOLO	You Only Look Once	WMV	Wireless Mesh Network

Everingham *et al*. (2015). The Pascal Visual Object Classes Challenge: A Retrospective. *International Journal of Computer Vision*, 111(1):98–136. DOI: 10.1007/s11263-014-0733-5.

Hellaoui *et al*. (2023). On supporting multi-services in uav-enabled aerial communication for the internet of things. *IEEE Internet of Things Journal*, pages 1–1. DOI: 10.1109/JIOT.2023.3262920.

Kulkarni, S., Chaphekar, V., Uddin Chowdhury, M. M., Erden, F., and Guvenc, I. (2020). Uav aided search and rescue operation using reinforcement learning. In *2020 SoutheastCon*, volume 2, pages 1–8. DOI: 10.1109/SoutheastCon44009.2020.9368285.

Li, J. *et al*. (2022a). 5G New Radio for Public Safety Mission Critical Communications. *IEEE Communications Standards Magazine*, 6(4):48–55. DOI: 10.1109/MCOMSTD.0002.2100036.

Li, J. *et al*. (2022b). Toward Providing Connectivity When and Where It Counts: An Overview of Deployable 5G Networks. *IEEE Communications Standards Magazine*, 6(4):56–64. DOI: 10.1109/MCOMSTD.0003.2100094.

Liang *et al*. (2023). Model-driven cluster resource management for ai workloads in edge clouds. *ACM Trans. Auton. Adapt. Syst.*, 18(1). DOI: 10.1145/3582080.

Linux Foundation (2024). Traffic Control in the Linux kernel. Available at: `https://man7.org/linux/man-pages/man8/tc.8.html`. Accessed: Jul-13-2024.

Macedo *et al*. (2022a). Improved support for UAV-based computer vision applications in Search and Rescue operations via RAN Intelligent Controllers. *XL Brazilian Symposium on Telecommunications and Signal Processing*, XL. DOI: 10.14209/sbrt.2022.1570817420.

Macedo *et al*. (2022b). UAV images packet loss distortions. Available at: `https://www.kaggle.com/dsv/4136188`. Accessed: Jul-12-2024.

Nishio *et al.* (2021). When wireless communications meet computer vision in beyond 5g. *IEEE Communications Standards Magazine*, 5(2):76–83. DOI: 10.1109/MCOM-STD.001.2000047.

O-RAN Working Group 1 (2021-02). O-RAN Operations and Maintenance Interface 4.0. Technical specification, O-RAN.WG1.O1-Interface.0-v04.00. Available at: `https://specifications.o-ran.org/download?id=32`.

O-RAN Working Group 1 (2022-03). O-RAN Architecture-Description 6.0. Technical specification, O-RAN.WG1.O-RAN-Architecture-Description-v06.00. Available at: `https://specifications.o-ran.org/download?id=641`.

O-RAN Working Group 2 (2021-07). A1 interface: General aspects and principles 2.03. Technical specification, ORAN-WG2.A1.GAP-v02.03. Available at: `https://specifications.o-ran.org/download?id=642`.

O-RAN Working Group 2 (2021-10). O-RAN AI/ML workflow description and requirements 1.03. Technical specification, O-RAN.WG2.AIML-v01.03. Available at: `https://specifications.o-ran.org/download?id=158`.

O-RAN Working Group 3 (2022-03). O-RAN Near-Real-time RAN Intelligent Controller Architecture & E2 General Aspects and Principles 2.01. Technical specification, O-RAN.WG3.E2GAP-v02.01. Available at: `https://specifications.o-ran.org/download?id=123`.

Our World in Data (2024). Economic damage by natural disaster type, 1900 to 2023. Available at: `https://ourworldindata.org/grapher/economic-damage-from-natural-disasters`. Accessed: Jul-13-2024.

Polese, M. *et al.* (2022). Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges. *CoRR*, abs/2202.01032. DOI: 10.48550/arXiv.2202.01032.

Redmon, J. and Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *CoRR*, abs/1804.02767. DOI: 10.48550/arXiv.1804.02767.

Rezatofighi *et al.* (2019). Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 658–666. DOI: 10.1109/CVPR.2019.00075.

Roldan *et al.* (2019). ITU Guidelines for national emergency telecommunication plans; ITU-D Emergency Telecommunications Disaster Response. Technical report, International Telecommunication Union (ITU). Available at: `https://www.itu.int/en/ITU-D/Emergency-Telecommunications/Documents/2020/NETP-guidelines.pdf`.

Saif *et al.* (2023). Skyward bound: Empowering disaster resilience with multi-UAV-assisted B5G networks for enhanced connectivity and energy efficiency. *Internet of Things*, 23:100885. DOI: 10.1016/j.iot.2023.100885.

Shule, W. *et al.* (2020). UWB-Based Localization for Multi-UAV Systems and Collaborative Heterogeneous Multi-Robot Systems. *Procedia Computer Science*, 175:357–364. DOI: 10.1016/j.procs.2020.07.051.

Tan, M. and Le, Q. V. (2020). EfficientNet: Rethinking model scaling for convolutional neural networks. *CoRR*, abs/1905.11946. DOI: 10.48550/arXiv.1905.11946.

Testbeds, N. R. (2024). Open-Access Research Testbed for Next-Generation Wireless Networks. Available at `https://www.orbit-lab.org/`. Accessed: Jul-13-2024.

Wang, Y. *et al.* (2020). From Design to Practice: ETSI ENI Reference Architecture and Instantiation for Network Management and Orchestration Using Artificial Intelligence. *IEEE Communications Standards Magazine*, 4(3):38–45. DOI: 10.1109/MCOMSTD.001.1900039.

Wu, X. *et al.* (2021). Deep Learning for Unmanned Aerial Vehicle-Based Object Detection and Tracking: A Survey. *IEEE Geoscience and Remote Sensing Magazine*, pages 2–35. DOI: 10.1109/MGRS.2021.3115137.

Xu, D. *et al.* (2020). Edge Intelligence: Architectures, Challenges, and Applications. *CoRR*, abs/2003.12172. DOI: 10.48550/arXiv.2003.12172.

Zeng, Y. *et al.* (2016). Wireless communications with unmanned aerial vehicles: opportunities and challenges. *IEEE Communications Magazine*, 54(5):36–42. DOI: 10.1109/MCOM.2016.7470933.

Zhang, H. *et al.* (2021). An empirical study of multi-scale object detection in high resolution UAV images. *Neurocomputing*, 421:173–182. DOI: 10.1016/j.neucom.2020.08.074.

Zhou *et al.* (2019). Edge Intelligence: Paving the Last Mile of Artificial Intelligence With Edge Computing. *Proceedings of the IEEE*, 107(8):1738–1762. DOI: 10.1109/JPROC.2019.2918951.