A Vision-aided Open Radio Access Network for Obstacle-aware Wireless Connectivity

Carolina Simões, André Coelho, Manuel Ricardo
INESC TEC, Faculdade de Engenharia, Universidade do Porto, Portugal

up201900208@edu.fe.up.pt, andre.f.coelho@inesctec.pt, manuel.ricardo@inesctec.pt

***Abstract*—High-frequency radio networks, including those operating in the millimeter-wave bands, are sensible to Line-of-Sight (LoS) obstructions. Computer Vision (CV) algorithms can be leveraged to improve network performance by processing and interpreting visual data, enabling obstacle avoidance and ensuring LoS signal propagation. We propose a vision-aided Radio Access Network (RAN) based on the O-RAN architecture and capable of perceiving the surrounding environment. The vision-aided RAN consists of a gNodeB (gNB) equipped with a video camera that employs CV techniques to extract critical environmental information. An xApp is used to collect and process metrics from the RAN and receive data from a Vision Module (VM). This enhances the RAN's ability to perceive its surroundings, leading to better connectivity in challenging environments.**

***Index Terms*—5G, obstacle-aware network, O-RAN, Vision-aided Radio Access Network.**

I. Introduction

Line-of-Sight (LoS) obstruction, caused by obstacles blocking the visual path between transmitter and receiver, has been a challenge for wireless communications and mobile networks. LoS obstruction results in signal attenuation causing degradation of communications quality, or even loss of connectivity. The millimeter-wave (mmWave) and sub-Terahertz (THz) frequencies, adopted respectively in 5G and 6G networks, are particularly sensitive to LoS obstruction.

Addressing the challenge of LoS obstruction and developing networking solutions capable of adapting to the evolving demands of wireless communications infrastructures is crucial. Furthermore, there is growing recognition of the importance of open, interoperable radio architectures for accelerating the technological progress.

Within 6G networks, mobile Base Stations (BSs) or mobile gNodeBs (gNBs) will help achieving ubiquitous network connectivity, with application in non-terrestrial, vehicular, and ad-hoc networks. However, unlocking the full potential of mobile BSs requires developing new software applications capable of controlling the Radio Access Network (RAN) in near real time. The xApps and rApps, defined by the O-RAN architecture [1], are a promising approach for controlling the mobile BSs by supporting the integration of radio, sensing, and vision-based information.

This work was supported by the CONVERGE project which has received funding under the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101094831.

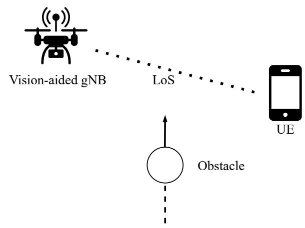

Fig. 1. Mobile gNB with a video camera to manage Line-of-Sight (LoS) obstructions to User Equipment (UE) caused by dynamic obstacles.

This approach enables the perception-aided mobile RAN, leveraging real-time environmental awareness to improve signal propagation and mobile devices localization. Computer Vision (CV) technology employing state-of-the-art algorithms, sensors, and video cameras, may improve network capabilities beyond traditional telecommunications by optimizing network performance through the proactive identification of obstacles to prevent signal attenuation and blockage, as illustrated in Figure 1. The integration of CV with communications technologies offers significant advantages, improving network responsiveness, adaptability, and overall performance.

The main contribution of this paper is a vision-aided RAN that integrates video-based information into the O-RAN architecture. The proposed solution extracts relevant information from video feeds and makes it available to O-RAN xApps through the exchange of messages containing video-derived information. These messages are designed to enhance network performance and avoid LoS obstructions. We also propose a novel xApp that processes video-derived information and RAN metrics, and uses it for the suitable placement of the mobile gNB and configuration of the RAN. By proposing a vision-aided real-time controller we enhance the gNB's capabilities and allow it to make informed decisions based on real-time environmental perception.

The remainder of this paper is structured as follows. Section II presents the related work. Section III explains the system design. Section IV details the system implementation. Section V discusses the validation carried out, including the main results obtained. Finally, Section VI refers to the main conclusions and directions for future work.

ISBN: 978-3-903176-71-3

II. Related Work

Recent works explore the use of CV to improve mobile networks. In [2], a Machine Learning (ML) solution that uses video data from base stations to improve wireless communications is presented. This includes anticipating blockages and facilitating user hand-offs in advance through a two-component deep learning architecture employing You Only Look Once (YOLO) v3 [3]. It incorporates visual and radio information to proactively predict blockages and manage seamless user hand-offs. Another work [4] proposes a CV-based approach to predict mmWave beams and blockages using RGB images captured by cameras and sub-6 GHz channel data. This approach avoids the need for explicit channel knowledge for beam training, demonstrating the potential of CV and deep learning to improve mmWave system capabilities and address wireless communications challenges. In [5], a real-world evaluation that leverages visual data and ML techniques to proactively predict dynamic mmWave link blockage is described. The study employs a CV-based solution to process visual data captured by infrastructure-mounted video cameras. Using the DeepSense 6G dataset [6], which includes multi-modal sensing and communications data, the study highlights the potential of integrating CV techniques in communications networks to mitigate link blockages and improve network performance.

In the context of O-RAN deployments, [7] introduces OpenRAN Gym, a framework designed for data-driven experimentation within the Open-RAN ecosystem. This framework supports the development, training, and testing of xApps – data-driven applications designed for Near-Real-Time RAN Intelligent Controllers (Near-RT RICs), defined in the O-RAN architecture. The xApps integrate service models for communications with RAN nodes over the E2 interface, incorporating data-driven logic units that host Artificial Intelligence (AI)/ML models for RAN inference, enabling intelligent closed-loop RAN control.

Regarding mobile gNBs, private standalone on-demand 5G network is presented in [8], where a mobile robotic platform carrying a gNB monitors the radio conditions of served User Equipments (UEs) and is remotely controlled in real-time using its video cameras. Similarly, [9] details the development of a private standalone 5G network with a mobile RAN employing the O-RAN architecture. This mobile RAN, also carried by a mobile robotic platform, employs an xApp to autonomously collect RAN metrics, analyze them, and optimize the placement of the mobile RAN, improving the wireless connectivity between UEs and the gNB. This work demonstrates the potential of the O-RAN architecture to facilitate the deployment of vendor-agnostic components and improve the flexibility of 5G and 6G networks. However, these works do not integrate CV.

As such, there remains a gap in the literature regarding solutions that integrate vision-based information with the O-RAN architecture, which represents an opportunity for relevant contributions to the O-RAN framework.

TABLE I
Description of each message type.

Message type	Description
Blockage	Sent when an obstacle is currently blocking the UE.
Prior-blockage	Sent when an obstacle is predicted to block the UE based on its current trajectory.
Post-blockage	Sent when an obstacle that was previously blocking the UE is no longer doing so.

III. Vision-aided Open RAN

In order to achieve a vision-aided, open RAN, we developed a Vision Module (VM) that employs CV techniques to detect and track obstacles that block LoS while sending relevant messages to connected services. The VM categorizes messages into three types: *blockage*, *prior-blockage*, and *post-blockage*. Table I describes each message type.

The payload of the messages varies according to their type. The fields include an identifier for the detected obstacle and UE, the type of detected object (e.g., a person), and the obstacle location within the video frame. *Prior-blockage* message fields include the obstacle velocity and the predicted time until obstruction. As for *Blockage* messages, the duration of obstruction is included.

A. Vision Module (VM)

The VM includes functionalities for information exchange, detection and tracking, image processing, and auxiliary tasks, such as coordinate transformations and message structuring.

The VM captures frames from a camera and processes them one at a time to track the location of the UE, which is marked with an Augmented Reality University of Córdoba (ArUco) marker [10] for easy identification. ArUco markers, commonly used in CV for camera pose estimation and object tracking, help distinguish the UE from obstacles. Open Source Computer Vision Library (OpenCV) [11] detects the ArUco marker and determines the Region of Interest (RoI) and its identification, while the VM periodically reports the UE's location within the frame. The *YOLOv8n* model, combined with Bag of Tricks Simple Online and Realtime Tracking (BoT-SORT) [12], detects and tracks obstacles, generating a tracking history with unique identifiers and positional data. If an obstacle blocks the ArUco marker, a blockage message is generated, and a post-blockage message is sent once the LoS is reestablished between the gNB and UE. The VM can also predict potential blockages based on object velocity and tracking history, allowing for proactive repositioning of the gNB to maintain LoS and ensure good channel quality, assessed by the Signal-to-Noise Ratio (SNR).

B. O-RAN architecture

In order to integrate CV into the O-RAN architecture, we leveraged the Near-RT RIC to deploy an xApp that manages both Radio Frequency (RF) metrics and VM messages. Communications between the VM and xApp are achieved through a new E2' interface, which is based on the O-RAN

ISBN: 978-3-903176-71-3

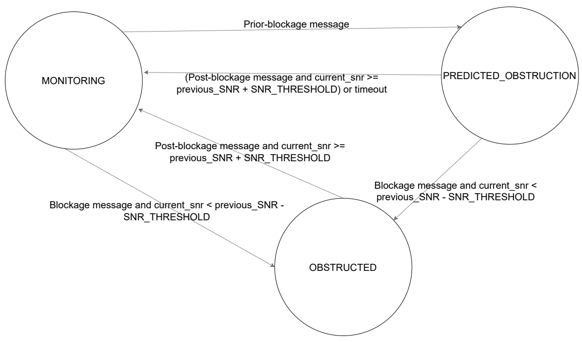

Fig. 2. State machine of the developed Vision and Radio xApp (VIRA).

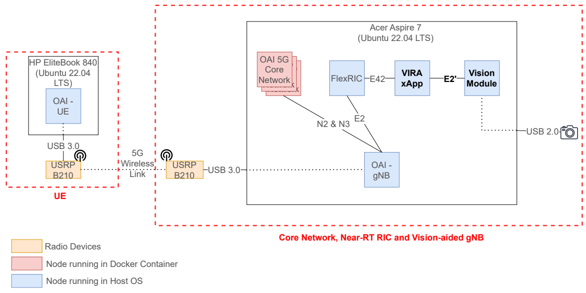

Fig. 3. Architecture of the proposed solution [17].

E2 interface and E2 Application Protocol (E2AP), as depicted in Figure 3. The E2' interface ensures reliable data exchange using a Stream Control Transmission Protocol (SCTP) socket connection and an Abstract Syntax Notation One (ASN.1) schema for message structuring. This design enables the VM to communicate with the xApp, allowing for the integration of video-derived data into mobile RAN management.

C. Vision and Radio xApp (VIRA)

The novel Vision and Radio xApp (VIRA), highlighted in Figure 3, integrates CV data with SNR values from the network to monitor network conditions and detect obstacles using a state machine with three states— `MONITORING`, `OBSTRUCTED`, and `PREDICTED_OBSTRUCTION`—based on SNR levels and obstacle detection, as depicted in Figure 2. In `MONITORING`, the system monitors obstacles and SNR levels. If a *Prior-blockage* message is received, it transitions to `PREDICTED_OBSTRUCTION`. If a *Blockage* message is received and the current SNR decreases, the system transitions to the `OBSTRUCTED` state. In the `OBSTRUCTED` state, the system awaits a *Post-blockage* message and an increase on the SNR. If this occurs, the state returns to `MONITORING`. In the `PREDICTED_OBSTRUCTION` and the `OBSTRUCTED` states, the VIRA xApp suggests the repositioning of the gNB.

IV. System Implementation

The system consists of two main logical units, represented by two dashed rectangles in Figure 3. The right-hand side rectangle includes the 5G Core Network, the Near-RT RIC, and the vision-aided gNB; the left-hand side rectangle represents the UE.

The 5G network components were implemented using the open-source software package OpenAirInterface (OAI) [13], which provides the components required to deploy the RAN and Core Network. Mosaic5G's FlexRIC [14] was selected to implement the Near-RT RIC due to its lightweight nature and ease of deployment, as it is executable from a single file.

The VM was developed using a combination of OpenCV, Ultralytics YOLO, BoT-SORT, and ASN1Tools. OpenCV was chosen for its well-documented Application Programming Interface (API) and its native support for ArUco markers. Ultralytics YOLO and BoT-SORT were selected for their simplicity and good performance in object detection and tracking, making them suitable for the intended application.

The OAI Core Network, FlexRIC, Vision Module (VM), and gNB were deployed on an Acer Aspire A715-74G laptop. The OAI Core Network was deployed in containers using Docker for ease of deployment and management. The VIRA xApp was deployed alongside FlexRIC to minimize latency between the two components. The gNB was implemented using OAI software. The laptop was connected to a Universal Software Radio Peripheral (USRP) B210 Software-Defined Radio (SDR) device [15] via a Universal Serial Bus (USB) 3.0 interface. The USRP B210 was equipped with two W5084K [16] dipole antennas tailored for the 3.6 GHz frequency band.

For the VM, an LL-4196 video camera was used, offering Full High-Definition resolution at a frame rate of 30 frames/s. The video camera was connected to the laptop using a USB 2.0 interface.

The UE was deployed on an HP EliteBook 840 laptop, equipped with a USRP B210 SDR and two W5084K dipole antennas to establish 5G radio communications with the gNB.

V. System Validation

In order to validate the implemented solution, we defined a use case based testing scenario within an indoor environment, as depicted in Figure 4. The scenario consisted of an obstacle moving from left to right, from the gNB's video camera perspective (right-hand side in Figure 4), at a constant velocity.

The objective was to assess the impact of blockages on the LoS between the gNB and the UE. By keeping both the gNB and the UE in fixed positions, we introduced an obstacle to observe its effect on signal quality. This scenario also aimed to validate the accuracy of the messages sent by the VM concerning the presence of blockages.

Based on data collected from the VM and the SNR values obtained from the RAN via the Near-RT RIC, VIRA instructed repositioning the gNB to restore the LoS with the UE.

Figure 5 presents a graph of the results collected during the experiment, illustrating the scenario in which an obstacle passes between the gNB and the UE. The y-axis represents the SNR variation over time, correlated with each message received from the VM. Each message is represented by a different color, as the legend indicates. Note that situations

ISBN: 978-3-903176-71-3

Fig. 4. Moving obstacle use case.

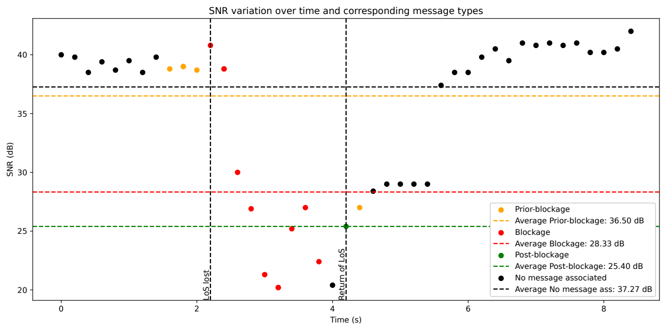

Fig. 5. SNR variation over time and the respective VM messages.

may arise in which SNR values are presented without an associated message, such as when detected obstacles are not expected to block the LoS.

The graph is divided into three distinct periods, separated by vertical transition lines. In the first period, as the obstacle moves towards the UE, *Prior-blockage* messages are sent. The SNR remains high because the LoS between the gNB and UE is not obstructed, indicating a strong received radio signal.

The first transition, labeled ′LoS lost′, identifies the beginning of the blockage period. It is important to refer that, because the ArUco marker is positioned to the left of the UE's SDR (from the gNB's video camera perspective), the blockage is detected from the CV perspective before the LoS is lost from the radio perspective. This explains the decrease in the SNR only after receiving some *Blockage* messages.

In the second period, as the obstacle is directly in front of the UE, the first *Blockage* message is received, indicating the start of the obstruction. As expected, the average SNR decreases during this period. The second transition, labeled ′Return of LoS′, identifies the end of the blockage, which allows for higher SNR values.

In the third period, as the obstacle moves away from the UE, and no longer blocks the LoS, the average SNR increases, approaching the values measured in the first period. This increase in SNR denotes the restoration of a higher signal and confirms the passage of the obstacle.

VI. Conclusions

We implemented a vision-aided RAN that integrates CV techniques into 5G/6G networks using the O-RAN architecture. The proposed solution demonstrates the potential of CV to improve the performance and autonomy of wireless networks by enabling environmental perception, especially in dynamic scenarios where LoS blockages occur. The solution was validated in a limited but representative scenario, confirming the initial intuition we had about the value of using CV to improve the radio access network's performance.

For future work, we aim at improving the integration of sensing into communications by considering additional data, such as 3D mapping for more precise obstacle detection and UE tracking. The proposed solution may also benefit from integrating VIRA with mobile robotic platforms for autonomous gNB repositioning to proactively prevent LoS blockages. Extending use case scenarios and conducting experimental tests on mmWave frequencies will further validate this approach, especially for high-frequency communications.

References

[1] O-RAN Alliance, "O-RAN Architecture Description 12.0," O-RAN Working Group 1, Technical Specification O-RAN.WG1.OAD-R003-v12.00, June 2024, r003. [Online]. Available: https://www.o-ran.org/specifications

[2] G. Charan, M. Alrabeiah, and A. Alkhateeb, "Vision-Aided 6G Wireless Communications: Blockage Prediction and Proactive Handoff," *IEEE Transactions on Vehicular Technology*, vol. 70, pp. 10 193–10 208, 10 2021.

[3] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv preprint arXiv:1804.02767*, 2018. [Online]. Available: https://arxiv.org/pdf/1804.02767.pdf

[4] M. Alrabeiah, A. Hredzak, and A. Alkhateeb, "Millimeter Wave Base Stations with Cameras: Vision-Aided Beam and Blockage Prediction," in *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, May 2020, pp. 1–5.

[5] G. Charan and A. Alkhateeb, "Computer Vision Aided Blockage Prediction in Real-World Millimeter Wave Deployments," in *2022 IEEE Globecom Workshops (GC Wkshps)*, 2022, pp. 1711–1716.

[6] A. Alkhateeb, G. Charan, T. Osman, A. Hredzak, J. Morais, U. Demirhan, and N. Srinivas, "DeepSense 6G: A Large-Scale Real-World Multi-Modal Sensing and Communication Dataset," *IEEE Communications Magazine*, vol. 61, no. 9, pp. 122–128, 2023.

[7] L. Bonati, M. Polese, S. D'Oro, S. Basagni, and T. Melodia, "Intelligent Closed-loop RAN Control with xApps in OpenRAN Gym," in *European Wireless 2022; 27th European Wireless Conference*, 2022, pp. 1–6.

[8] D. Maia, A. Coelho, and M. Ricardo, "Obstacle-aware On-demand 5G Network using a Mobile Robotic Platform," in *2022 18th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 2022, pp. 470–473.

[9] G. Queirós, P. Correia, A. Coelho, and M. Ricardo, "Autonomous control and positioning of a mobile radio access node employing the o- ran architecture," in *2024 19th Wireless On-Demand Network Systems and Services Conference (WONS)*, 2024, pp. 25–28.

[10] S. Garrido-Jurado, R. Munoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marin-Jimenez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, June 2014.

[11] I. Culjak, D. Abram, T. Pribanic, H. Dzapo, and M. Cifrek, "A brief introduction to OpenCV," in *2012 Proceedings of the 35th International Convention MIPRO*, 2012, pp. 1725–1730.

[12] N. Aharon, R. Orfaig, and B.-Z. Bobrovsky, "BoT-SORT: Robust Associations Multi-Pedestrian Tracking," *arXiv preprint arXiv:2206.14651*, 2022.

[13] OpenAirInterface, "Open Air Interface Repository," https://gitlab.eurecom.fr/oai/openairinterface5g, accessed: 2024-08-13.

[14] M. Project, "FlexRIC: Flexible RAN Intelligent Controller," https://gitlab.eurecom.fr/mosaic5g/flexric, 2024.

[15] E. Research, "USRP B210 Product Page," https://www.ettus.com/all-products/ub210-kit/, accessed: 2024-08-13.

[16] P. Antennas, "W5084k Dipole Antenna," https://www.farnell.com/datasheets/3812050.pdf, accessed: 2024-08-13.

[17] C. Simões, "Vision-aided Open Radio Access Network," MSc. Thesis, Faculdade de Engenharia, Universidade do Porto, Portugal, 2024.

ISBN: 978-3-903176-71-3