

## Article

# DDPG-Based UAV-RIS Framework for Optimizing Mobility in Future Wireless Communication Networks

Yasir Ullah <sup>1</sup>, Idris Olalekan Adeoye <sup>1</sup>, Mardeni Roslee <sup>1,\*</sup>, Mohd Azmi Ismail <sup>2</sup>, Farman Ali <sup>1</sup>,  
Shabeer Ahmad <sup>3</sup>, Anwar Faizd Osman <sup>4</sup> and Fatimah Zaharah Ali <sup>5</sup>

<sup>1</sup> Center for Wireless Technology, Faculty of Artificial Intelligence and Engineering, Multimedia University, Cyberjaya 63100, Malaysia; yasirullah415@gmail.com (Y.U.); idrisoadeoye@gmail.com (I.O.A.); drfarmanali.optics@gmail.com (F.A.)

<sup>2</sup> Centre of Excellence for Intelligent Network, Telekom Malaysia Research & Development, Cyberjaya 63000, Malaysia; mohdazmi@tmrnd.com.my

<sup>3</sup> School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China; shabeer@bupt.edu.cn

<sup>4</sup> Telekom Malaysia (TM) Technology Services Sdn Bhd, Kuala Lumpur 50672, Malaysia; anwar.faizd@tm.com.my

<sup>5</sup> College of Engineering, Universiti Teknologi MARA (UiTM), Shah Alam 40450, Malaysia; fatimah\_zaharah@uitm.edu.my

\* Correspondence: mardeni.roslee@mmu.edu.my

**Abstract:** The development of beyond 5G (B5G) future wireless communication networks (FWCN) needs novel solutions to support high-speed, reliable, and low-latency communication. Unmanned aerial vehicles (UAVs) and reconfigurable intelligent surfaces (RISs) are promising techniques that can enhance wireless connectivity in urban environments where tall buildings block line-of-sight (LoS) links. However, existing UAV-assisted communication strategies do not fully address key challenges like mobility management, handover failures (HOFs), and path disorders in dense urban environments. This paper introduces a deep deterministic policy gradient (DDPG)-based UAV-RIS framework to overcome these limitations. The proposed framework jointly optimizes UAV trajectories and RIS phase shifts to improve throughput, energy efficiency (EE), and LoS probability while reducing outage probability (OP) and HOF. A modified K-means clustering algorithm is used to efficiently partition the ground users (GUs) considering the newly added GUs as well. The DDPG algorithm, based on reinforcement learning (RL), adapts UAV positioning and RIS configurations in a continuous action space. Simulation results show that the proposed approach significantly reduces HOF and OP, increases EE, enhances network throughput, and improves LoS probability compared to UAV-only, RIS-only, and without UAV-RIS deployments. Additionally, by dynamically adjusting UAV locations and RIS phase shifts based on GU mobility patterns, the framework further enhances connectivity and reliability. The findings highlight its potential to transform urban wireless communication by mitigating LoS blockages and ensuring uninterrupted connectivity in dense environments.

**Keywords:** DDPG; GUs connectivity; RIS configuration; LoS probability; UAV trajectory; UAV-RIS; user grouping



Academic Editor: Emmanouel T. Michailidis

Received: 18 April 2025

Revised: 4 June 2025

Accepted: 13 June 2025

Published: 15 June 2025

**Citation:** Ullah, Y.; Adeoye, I.O.; Roslee, M.; Ismail, M.A.; Ali, F.; Ahmad, S.; Osman, A.F.; Ali, F.Z. DDPG-Based UAV-RIS Framework for Optimizing Mobility in Future Wireless Communication Networks. *Drones* **2025**, *9*, 437. <https://doi.org/10.3390/drones9060437>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The use of higher frequency bands such as millimeter wave (mmWave) and Terahertz (THz) in beyond 5G (B5G) networks is expected to greatly increase network capacity, support a large number of devices, and enable applications such as virtual reality (VR)

and augmented reality (AR) [1–5]. To meet these requirements, UAVs have emerged as a promising solution. Unlike traditional ground base stations (GBSs), UAVs can be deployed flexibly, adjust their positions as needed, and provide cost-effective wireless coverage [6–8]. With their ability to establish line-of-sight (LoS) links with user equipment (UEs), UAVs improve data rates, connectivity, and service reliability [9–11].

Despite these features, UAV-based communication faces several challenges, especially in dense urban environments. High-rise buildings and other obstacles like trees block LoS paths, limiting coverage and causing user disconnections. Additionally, UAVs are highly mobile, which can cause frequent handovers, and maintaining stable connections over long distances can be difficult. These issues make UAV-only communication systems less reliable in dense urban environments.

To address these issues, reconfigurable intelligent surfaces (RIS) technology has been proposed as a promising solution. RIS consists of an extensive array of passive reflecting elements that can intelligently adjust the phase of incoming signals to improve wireless transmission [12–15]. Using adjusting and redirecting signal phenomena around obstacles, RIS can enhance network coverage and data rates in dense areas [16]. Since RIS elements are passive and do not require active power sources, they offer an energy-efficient alternative to traditional relay systems. However, RIS alone is not enough to fully solve the challenges of UAV-based communication. While RIS can improve signal transmission, it cannot handle the frequent link disruptions caused by UAV mobility and user movement. Therefore, integrating UAVs with RIS can provide a more robust solution using both technologies. The UAV ensures flexible coverage, while the RIS enhances signal quality by reflecting signals toward intended users.

This work explores UAV-integrated RIS (UAV-RIS) to improve mobility management and HO performance in dense urban heterogeneous network (HetNet) environments. The UAV provides LoS communication, while the RIS dynamically adjusts signal reflections to mitigate blockages. To maximize throughput, improve energy efficiency (EE), and ensure seamless connectivity, we jointly optimize UAV trajectories and RIS phase shifts using a deep deterministic policy gradient (DDPG) approach. Simulation results demonstrate that the proposed UAV-RIS framework significantly reduces HO failures (HOF), enhances network reliability, and improves overall communication performance, highlighting its potential for B5G wireless networks.

### 1.1. Related Work

UAVs and RIS have emerged as essential solutions for addressing the challenges of establishing LoS and reliable connectivity to GUs. Most existing studies focus on either deploying UAVs (static and mobile) or deploying RIS on buildings to maximize the communication performance of GUs. However, these approaches significantly limit their application in dynamic, real-time GU communication within densely populated urban environments. The integration of RIS in UAV-aided communication systems has recently gained significant attention due to improved system performance [17]. For example, a UAV-based scheme for enhancing mobility management and reducing handover (HO) rate in FWCNs is proposed in [18], utilizing reinforcement learning (RL) to optimize UAV trajectories and enhance LoS links, ensuring seamless connectivity for GUs during transitions between cells. Similarly, a multi-UAV-mounted BS framework for serving multiple IoT devices is proposed in [19], which optimizes 3D layouts and resource allocation to minimize uplink transmission power. However, UAV flexibility is overlooked, which reduces coverage and the ability to adjust positions in response to real-time communication demands. Other studies, such as [20], focus on lowering the HO rate and interference using deep Q-network (DQN) algorithms for UAV communications. However, they do not fully address challenges such as limited coverage at

high altitudes or LoS blockages. Additionally, ref. [21] introduces a RIS-assisted HO scheme utilizing deep reinforcement learning (DRL) to mitigate mm-wave channel blockages, significantly reducing HO rate and improving spectral efficiency (SE) through joint adjustment of beamformers and RIS phase shifts. In contrast, RIS-assisted UAV systems have become pivotal in enhancing communication performance, mainly when direct communication links are unavailable. Several studies have explored UAV-RIS systems with various optimization techniques to improve network performance in FWCNs. For instance, ref. [22] investigated a UAV-RIS downlink transmission system and proposed a successive convex approximation (SCA)-based algorithm to optimize UAV trajectory and RIS beamforming. In [23], RIS in UAV-assisted communication systems was explored to achieve substantial performance improvements in UAV-aided cellular networks. Another study, [24], used DRL to optimize UAV trajectory, RIS phase shifts, and power allocation, aiming to maximize the sum rate for mobile UEs. Similarly, ref. [25] developed a RIS-assisted UAV communication system with alternating optimization algorithms to jointly optimize UAV mobility, resource allocation, and RIS scheduling, ensuring improved system rate and heterogeneous QoS for each mobile user. In [26], an RIS-assisted UAV framework employing an energy-efficient UAV deployment (EEUD) algorithm is proposed to maximize energy efficiency by jointly optimizing RIS phase shifts, UAV trajectory, and BS transmit power. The study in [27] investigated UAV-assisted RIS deployment to enhance connectivity and SNR when direct communication links are unavailable. Furthermore, ref. [28] utilized a DRL scheme to improve communication efficiency between the ground base station (GBS) and mobile vehicles. A similar approach was adopted in [29], which employed the deep deterministic policy gradient (DDPG) algorithm to optimize BS power, RIS reflection coefficients, and UAV positioning to maximize the communication rate. In [30], a total energy harvesting strategy for a UAV-RIS system is proposed to enhance energy efficiency (EE) and meet QoS requirements. Other studies, such as [31], have used the UAV-RIS technology to improve downlink secrecy rates by optimizing transmit power allocation, RIS beamforming, and UAV trajectory, demonstrating the potential for high performance in FWCNs. The comparison of the proposed work with existing models, in terms of system configuration, optimization techniques, mobility consideration, and limitations, is presented in Table 1.

**Table 1.** Comparison of the proposed work with existing related studies.

Ref.	System Configuration	Optimization Technique	Mobility Consideration	Remarks
[23]	RIS-assisted UAV	RIS location optimization for gain maximization	UAV users	Does not consider GU mobility; focuses only on RIS deployment in buildings.
[24]	RIS-assisted UAV	DRL-based rate enhancement	Mobile GUs	Limited joint design for RIS and UAV trajectories.
[29]	IRS-based UAV-NOMA	DRL-based beamforming and UAV optimization	Mobility clustering not considered	Lacks user mobility clustering despite IRS integration.
[30]	UAV-RIS-SWIPT	SWIPT-based energy optimization	Static users	Does not considered mobile users, focus on static users with energy-harvesting RIS.
[31]	RIS-UAV Mobile Vehicle	Successive convex approximation-based secrecy rate maximization	UAV is assumed to be static	No dynamic UAV mobility; RIS for secrecy enhancement.
<b>This Work</b>	UAV-RIS with user clustering	DDPG-based joint UAV trajectory and RIS phase optimization	Dynamic mobility management of mobile GUs	Fully dynamic, adaptive UAV-RIS integration with user clustering.

## 1.2. Motivations and Contributions

From the aforementioned studies, it is evident that relatively few investigations have focused on RIS-empowered UAV communication systems aimed at enhancing the performance of FWNs. Much of the existing research has concentrated on maximizing network throughput, secrecy, and sum rate while reducing energy consumption. However, none of these studies have addressed GU mobility management and HO performance in FWNs through the joint utilization of UAVs and RIS. Deploying low-cost RIS on UAVs provides multiple transmission paths for GUs, improving communication quality and overcoming obstacles and limitations in terrestrial wireless networks. Motivated by the research works in [18,21,22], which focus on UAV-only, RIS-only, and UAV-RIS deployments, respectively, this paper proposes an RIS-assisted UAV system that employs DDPG and DQN algorithms to manage GU mobility and improve HO/mobility performance in dense urban environments. Furthermore, the proposed framework enhances communication by adjusting RIS phase shifts and utilizing UAV mobility to optimize its trajectory based on the GU's location and requirements, ensuring stable link quality and improved connectivity. The main contributions of this paper are as follows:

- A novel UAV-RIS framework is proposed to enhance signal strength and improve LoS connectivity between UAVs and GUs in dense urban environments by addressing communication disruptions caused by ground obstacles.
- A modified K-means clustering algorithm is introduced for efficient user partitioning, alongside a DDPG algorithm to intelligently optimize UAV trajectories and RIS configurations simultaneously in a continuous action space for managing GUs mobility.
- The proposed framework, utilizing the DDPG algorithm, significantly improves key network performance metrics, including HOF, OP, EE, throughput, and LoS probability, compared to state-of-the-art schemes.

The rest of this paper is organized as follows: Section 2 presents the system model for the RIS-assisted UAV communication system. Section 3 describes the problem formulation and discusses the proposed iterative solution. Performance results are then presented and analyzed in Section 4. Finally, the conclusions and future work are outlined in Section 5.

## 2. System Model

This section introduces the system model for downlink data transmission to manage GU mobility in 5G/B5G networks by utilizing a mobile RIS-assisted UAV wireless communication network. The detailed architecture of the system model is depicted in Figure 1, which consists of a BS operating at the mmWave frequency band, equipped with multiple antennas [32]. We assume that the direct communication links between the GBS and the GUs are obstructed by tall buildings or other environmental impediments, so we employ UAV-RIS technology to improve the downlink data transmission service and GUs mobility performance. The BS is located at the center point of an urban area denoted by  $C$  and serve GUs are randomly distributed in that area denoted by  $U$ . These GUs are partitioned into different groups based on SINR to enhance the system's performance. In NLoS scenarios, the communications between the GBS and the GUs is assisted by a mobile UAV-RIS, modeled as a uniform planar array (UPA). The total number of reflective elements is  $F = F_x F_y$ , where  $F_x$  and  $F_y$  are the number of RIS reflecting elements along the X-axis and Y-axis, respectively. The RIS learns the optimal method to reflect incident signals by adjusting the phase shift, which helps improve the mmWave network's performance. In this model, we consider that a single UAV-RIS flies at a fixed altitude over a specific area and its position remains the same during one time slot and the flight is stable. Without loss of generality, the

coordinates of GBS, user location, and UAV-RIS can be defined as  $q_B = (x^{BS} = 0, y^{BS} = 0, h^{BS})$ ,  $q_U = (x^{GU}, y^{GU}, h^{GU})$ , and  $q_{UR} = (x^{UR}, y^{UR}, h^{UR})$ , respectively.

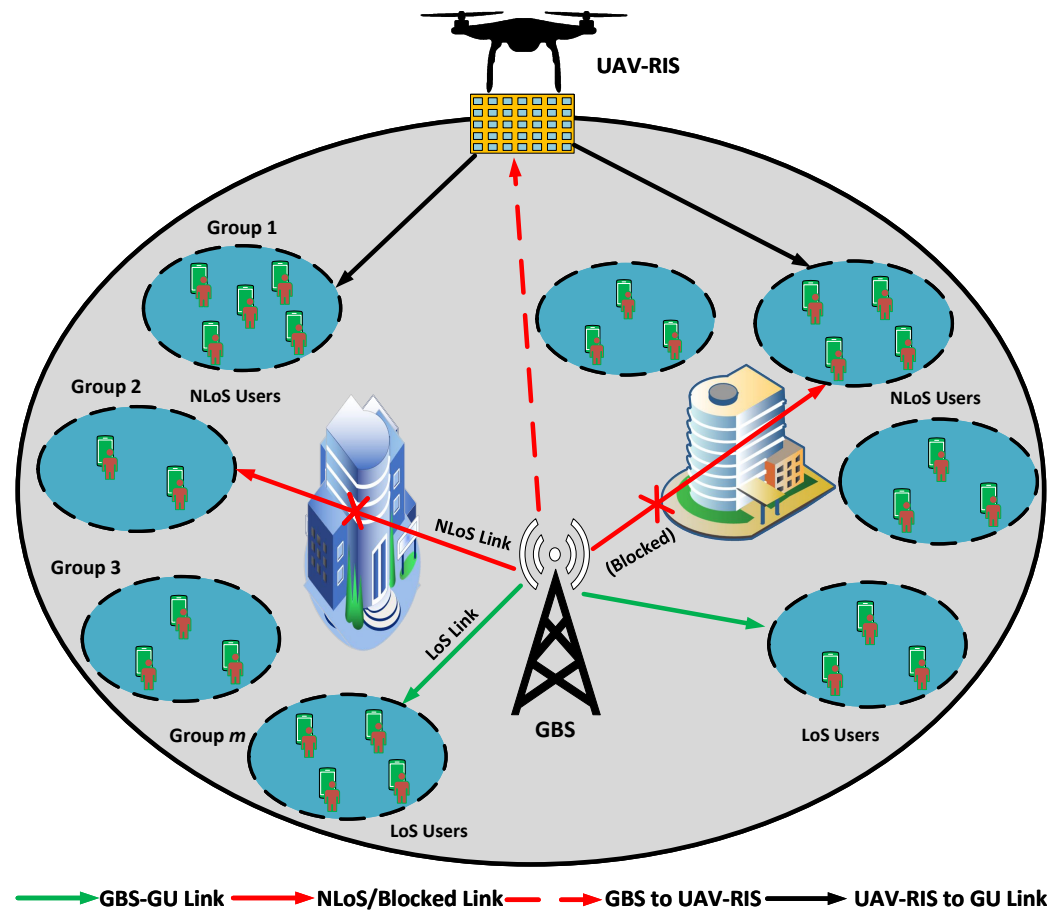


Figure 1. UAV-assisted RIS framework for GUs mobility management.

To facilitate the trajectory design, the total UAV-RIS flying time,  $T$ , is divided into  $N$  time slots within equal time intervals ( $T = N\delta_t$ ), where  $\delta_t$  is a single slot length. The movement of UAV-RIS satisfies the following mobility constraints:

$$\|q_{UR}[n+1] - q_{UR}[n]\|^2 \leq V^2, n = 1, \dots, N-1, \quad (1a)$$

$$\|q_{UR}[N] - q_F\|^2 \leq V^2, q_{UR}[1] = q_0, \quad (1b)$$

or can be written as follows:

$$\|q_{UR}[n+1] - q_{UR}[n]\|^2 \leq V^2, n = 1, \dots, N-1, \quad (2)$$

where  $q_F$  and  $q_0$  represent the final and initial horizontal positions of UAV-RIS, respectively. The values of  $q_0$  and  $q_F$  are determined by the centroids of the user groups obtained through the modified K-means clustering algorithm described in Section 3.1. This clustering process considers both user locations and SINR values to ensure optimal UAV-RIS positioning at the start and end of the mission within the feasible flight zone.  $V = v_{\max}\delta_t$  is the maximum horizontal distance that the UAV-RIS travels in a single time slot, and  $v_{\max}$  denotes the maximum UAV-RIS speed in m/s. The UAV operates at a fixed altitude within each time slot due to airspace safety regulations and energy efficiency considerations, while horizontal coordinates are dynamically optimized across slots.

The Euclidean distance from GBS to the UAV-RIS and from the UAV-RIS to the GUs is repressed by  $d_1$  and  $d_2$ , respectively, and can be calculated as

$$d_1 = \sqrt{(x^{\text{BS}} - x^{\text{UR}})^2 + (y^{\text{BS}} - y^{\text{UR}})^2 + (h^{\text{BS}} - h^{\text{UR}})^2} \quad (3)$$

$$d_2 = \sqrt{(x^{\text{UR}} - x^{\text{GU}})^2 + (y^{\text{UR}} - y^{\text{GU}})^2 + (h^{\text{UR}} - h^{\text{GU}})^2} \quad (4)$$

There may exist LoS (direct) and NLoS (indirect) links between UAV-RIS and GUs communication. These indirect links, caused by obstacles such as trees and high-rise buildings, introduce a multipath propagation effect that can degrade transmission signal quality and distort QoS in urban HetNet. To address this challenge, UAV-RIS is deployed to enhance the probability of LoS links, thus enabling high-speed data transmission with improved connectivity and HO performance. The probability of LoS between UAV-RIS and GUs is given as

$$p_{\text{LoS}}^{\text{UR}} = \frac{1}{1 + \zeta \exp\left[\Psi\left(\frac{180}{\pi} \tan^{-1} \frac{h^{\text{UR}}}{d_2}\right) - \zeta\right]} \quad (5)$$

Here,  $\Psi$  and  $\zeta$  are constant values representing the environmental influence such as urban, and suburban environments. The probability of NLoS between the UAV-RIS and GU  $u$  is

$$p_{\text{NLoS}}^{\text{UR}} = 1 - p_{\text{LoS}}^{\text{UR}} \quad (6)$$

Based on the above LoS and NLoS probability formulations, the path loss (PL) for the LoS and NLoS links between the UAV-RIS and the GU  $u$  are written [33] as

$$PL_{\text{LoS}}^{\text{UR}} = \delta_1 \left[ \frac{4\pi f_c h^{\text{UR}}}{c_l} \right]^\alpha p_{\text{LoS}}^{\text{UR}} \quad (7)$$

$$PL_{\text{NLoS}}^{\text{UR}} = \delta_2 \left[ \frac{4\pi f_c h^{\text{UR}}}{c_l} \right]^\alpha p_{\text{NLoS}}^{\text{UR}} \quad (8)$$

where  $\delta_1$  and  $\delta_2$  are the PL coefficients for LoS and NLoS links, respectively, and their values depend on the area type. The PL coefficients  $\delta_1$  and  $\delta_2$  vary according to the environment type: (i) 0.1 dB and 21 dB in suburban, (ii) 1 dB and 20 dB in urban, (iii) 1.6 dB and 23 dB in dense urban, and (iv) 2.3 dB and 34 dB in high-rise urban environments, respectively.

Moreover,  $f_c$  is the carrier frequency,  $\alpha$  is the PL exponent, and  $c_l$  is the speed of light. Furthermore, the average PL between UAV-RIS and GU  $u$  can be calculated by utilizing (7) and (8):

$$PL_{\text{avg}}^{\text{UR}} = \left[ \frac{4\pi f_c h^{\text{UR}}}{c_l} \right]^\alpha \left( \delta_1 p_{\text{LoS}}^{\text{UR}} + \delta_2 p_{\text{NLoS}}^{\text{UR}} \right). \quad (9)$$

We consider UAV-mounted RIS to facilitate GUs in dense urban areas, where obstacles block direct communication links between GBS and GUs. The channel gain of GBS-(UAV-RIS)-GUs is expressed [34] as

$$G(t) = g(t) \Phi v_g v_u \quad (10)$$

Here,  $\Phi$  is the RIS phase shifts matrix and is denoted by

$$\Phi = \text{diag} \left[ e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_F} \right] \quad (11)$$

where  $\theta_f \in [0, 2\pi]$ ,  $f \in F = \{1, 2, \dots, F\}$  is the phase shift of the  $F$ th reflecting element.



The channels between GUs and UAV-RIS contain both LoS and NLoS links. Using Rician channel modeling, these channel gains are modeled as

$$H = \sqrt{\eta d_2^{-\gamma}} \left( \sqrt{\frac{R}{R+1}} h_{\text{LoS}} + \sqrt{\frac{1}{R+1}} h_{\text{NLoS}} \right) \quad (12)$$

where  $R$  is the Rician fading factor, and  $h_{\text{LoS}}$  and  $h_{\text{NLoS}}$  are the fast fading components of LoS and NLoS channels between UAV-RIS and GUs, respectively.  $\gamma$ ,  $\eta$ , and  $d_2$  represent PL exponent, large-scale fading, and Euclidean distance among UAV-RIS and GUs, respectively.  $R$ -value decides whether the channel is a Rayleigh or Rician channel. When  $R = 0$ , then  $H$  is a Rayleigh channel; otherwise, it becomes a Rician channel. The Rician factor  $R = 2$  is selected to model realistic urban UAV-to-ground links where moderate LoS dominance is observed, balancing strong direct components with multipath scattering effects.  $h_{\text{NLoS}}$  is the non-deterministic LoS component and is modeled as complex Gaussian distributed with zero mean and unit variance. On the other hand,  $h_{\text{LoS}}$  is the deterministic LoS component and is expressed as

$$\begin{aligned} \mathbf{h}_{\text{LoS}} = & \left[ 1, \dots, \exp \left( -j \frac{2\pi}{\lambda} d (F_x - 1) \right. \right. \\ & \left. \left. \sin \phi_u \cos \varphi_u \right) \right] \\ & \otimes \left[ 1, \dots, \exp \left( -j \frac{2\pi}{\lambda} d (F_y - 1) \cos \phi_u \right) \right] \end{aligned}$$

where  $\sin \phi_u \cos \varphi_u = \frac{x^{\text{UR}} - x^{\text{GU}}}{d_2}$  and  $\cos \phi_u = \frac{h^{\text{UR}} - h^{\text{GU}}}{d_2}$  with  $\phi_u$  and  $\varphi_u$  represent the azimuth and elevation angles of arrival of the UAV-RIS and the GU  $u$ , respectively.  $\lambda$  and  $d$  denote the wavelength and antenna separation while  $F_x$  and  $F_y$  are the reflecting elements of RIS along the X-axis and Y-axis.

We assumed LoS links from GBS to UAV-RIS and from UAV-RIS to GUs;  $g(t)$  in (10) is the cascaded channel gain of GBS-(UAV-RIS)-GU and is calculated as

$$g(t) = h_1 \Phi H \quad (13)$$

where  $H$  is the UAV-RIS to GU channel gain and  $h_1$  is the GBS to UAV-RIS channel gain. Furthermore,  $v_g$  and  $v_u$  in (10), respectively, are the received array vector from GBS to the UAV-RIS and the transmit array vector from UAV-RIS to the  $k$ th GU at time  $t$ , and can be formulated [35] as

$$\begin{aligned} v_g &= \left[ e^{-j\theta_1}, e^{-j\theta_2}, \dots, e^{-j\theta_F} \right] \\ v_u &= \left[ e^{-j\beta_1}, e^{-j\beta_2}, \dots, e^{-j\beta_F} \right] \end{aligned}$$

Here, variable  $\theta$  denotes the relative phase difference between the received signal at the GBS and the first UAV-RIS element, while  $\beta$  represents the relative phase difference among the elements of the UAV-RIS reflected beams towards the GUs.

For the sake of fairness, every user in the group is allocated the same bandwidth ( $B$ ), and each UAV-RIS in the network utilizes the same frequency band simultaneously. By combining (4), (9), and (10), the signal-to-interference plus noise ratio (SINR) from each UAV-RIS to GUs in the HetNet during time instant  $t$  can be expressed as follows:

$$\omega = \frac{P_{\text{UR}} * G(t)}{\psi + B(d_2 * PL_{\text{avg}}^{\text{UR}} * \sigma^2)} \quad (14)$$

where  $P_{UR}$  is the transmit power of UAV-RIS to their GUs and  $B$  is the system bandwidth in the mmWave range.  $\psi$  and  $\sigma$  represent the network experiences interference and additive white Gaussian noise (AWGN), respectively. Based on Equation (14), the sum rate of GUs could be obtained as

$$\Theta_U = B \log_2(1 + \omega) \quad (15)$$

Here,  $\omega$  represents the instantaneous SINR for each ground user, and the sum rate expression follows the Shannon capacity formula.

### 3. Problem Formulation

In this article, our objective is to maximize the sum rate of GUs in urban areas by jointly optimizing the UAV trajectory, RIS phase shift, and transmit power from UAV-RIS to GUs. The sum rate maximization problem can be written as

$$\max_{\{q, \theta_n, p\}} \sum_{u=1}^U \Theta_u \quad (16)$$

$$\text{s.t.} : \Theta_u \geq \Theta_{\min} \forall t \quad (16a)$$

$$\sum_{u=1}^U P_u^{\text{UR}} \leq P_{\max}^{\text{UR}} \quad (16b)$$

$$0 \leq \theta_f \leq 2\pi; f = 1, \dots, F \quad (16c)$$

$$\begin{aligned} x_{\min} &\leq x^{\text{UR}} \leq x_{\max}; \\ y_{\min} &\leq y^{\text{UR}} \leq y_{\max}; \\ h_{\min} &\leq h^{\text{UR}} \leq h_{\max} \end{aligned} \quad (16d)$$

Constraint (16a) provides a minimum achievable rate for all GUs while guaranteeing their QoS. Constraint (16b) ensures that the overall power transmitted by the UAV-RIS does not exceed the maximum power transmission limit. Constraints in (16c) specify that the RIS reflecting matrices are phase shift matrices, ensuring that all transmitted signals are reflected without any power loss. Constraint (16d) limits UAV-RIS's capability to fly within a specific area, aiming to provide seamless connectivity to GUs and achieve a significant enhancement in throughput.

The optimization problem mentioned in (16) is non-convex and difficult to solve directly. Therefore, in the subsequent sections, we divide the original problem into two subproblems to determine the optimal user partitioning and the best locations of UAVs, along with the orientation of the RIS phase shift. The proposed methodology consists of two main stages. First, users are partitioned into clusters based on both distance and SINR levels using a modified K-means algorithm to ensure QoS satisfaction. Second, a DDPG-based reinforcement learning model is employed to jointly optimize UAV horizontal trajectory and RIS phase shifts, maximizing the system sum-rate while reducing HOFs and maintaining EE under user mobility.

#### 3.1. User Partitioning

User partitioning in mmWave and THz bands is an extremely challenging task for enhancing throughput performance. An iterative algorithm is utilized to address optimal user partitioning, where users are partitioned into groups based on their distances and SINR values. The proposed modified k-mean algorithm for user grouping includes vital modifications to the standard k-means algorithm such that it can incorporate the shortest Euclidean distance criteria (high SINR) and the data rate constraint mentioned in (16a) for each group independently. It offers the advantages of low complexity, easy implementation, and fast convergence. Algorithm 1 describes the user grouping process step by step that



may achieve faster convergence and enhance the overall mmWave system performance in terms of HO and mobility. The proposed modified K-means algorithm enables efficient user grouping by jointly considering both spatial distance and SINR levels, ensuring that each cluster maintains acceptable signal quality and proximity for effective UAV-RIS optimization. This two-stage clustering enhances the initial conditions for trajectory and phase shift optimization, reducing handover frequency and improving link stability. The primary advantage of this approach lies in its low computational complexity, fast convergence, and ability to incorporate QoS constraints. However, a potential limitation is that the initial cluster centroids may influence convergence to local optima, especially in highly dynamic user mobility scenarios.

---

**Algorithm 1:** Proposed SINR/Distance-based User Partitioning Scheme

---

**Input:** Number of GUs, SINR threshold, sum rate threshold

**Initialization:** Randomly generate each user's locations and initial data rates.

- 1: **for** every user  $u = 1, \dots, U$  **do**
- 2:   Calculate Euclidean distance between the user and neighboring users.
- 3:   Randomly assign data rate to each user using random distribution techniques.
- 4:   Assign the user to the group with the highest SINR, minimum distance, and maximum data rate.
- 5:   Repeat steps 2–4 for all users in the network until convergence.
- 6: **end for**

**Until** Groups achieve minimal inter-user distances, highest SINR, and maximum data rates.

**Output:** Optimal User Groups

---

The communication coverage area depends on UAV-RIS altitude; the higher the UAV-RIS altitude, the larger the coverage area, and vice versa. The optimal coverage radius, based on UAV-RIS altitude and beam angle in the mmWave network, can be expressed as

$$\text{radius}(r) = h^{\text{UR}} * \tan(\vartheta) \quad (17)$$

where  $\vartheta$  is the half beam angle in the mmWave band and  $h^{\text{UR}}$  is the UAV-RIS height.

### 3.2. Joint Optimization of UAV Location and RIS Phase Shifts

To jointly optimize the UAV location and RIS phase shift in mmWave networks, a DRL-DDPG algorithm is proposed to address the formulated problem of finding the optimal UAV location and adjusting RIS phase shifts. The proposed DRL framework consists of a state set  $s(t)$ , an action set  $a(t)$ , a reward set  $r(t)$ , and UAV-RIS as an agent by the DDPG algorithm [36].  $s(t)$ ,  $a(t)$ , and  $r(t)$  are clearly described as follows:

**State Space:** The set of spaces, including the UAV's optimal location and RIS phase shifts, at time  $t - 1$  is described as

$$s_t = \left[ \underbrace{\theta_1^{(t-1)}, \dots, \theta_{2F}^{(t-1)}}_{\text{RIS phase shift}} \mid \underbrace{x^{(t-1)}, y^{(t-1)}, h^{(t-1)}}_{\text{UAV optimal location}} \right] \quad (18)$$

**Action Space:** The action space includes a UAV movement and RIS phase shifts when transitioning from the current to the next state. The suggested approach enables the agent to continuously determine the optimal movement while considering the long-term reward and identify the optimal phase shift for each time instance. The agent (UAV-RIS) inputs the state  $s_t$  at time step  $t$  to determine the appropriate action based on the current environment,

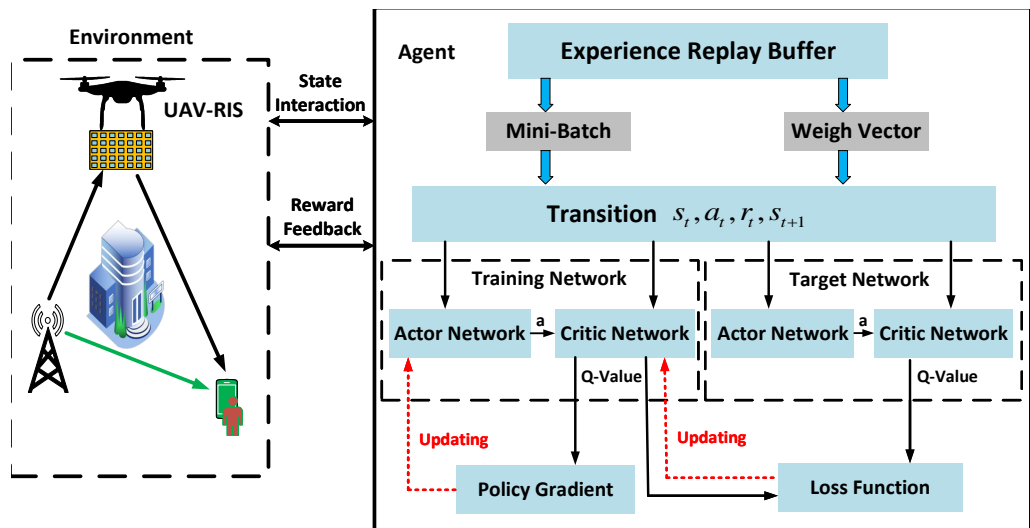
resulting in the optimal UAV horizontal location and updated RIS phase shift to improve connectivity and mobility issues in urban environments. The action space is expressed as

$$\mathbf{a}_t = [\theta_1^{(t)}, \dots, \theta_{2F}^{(t)}, x^{(t)}, y^{(t)}, h^{(t)}] \quad (19)$$

**Reward:** After performing action  $\mathbf{a}_t$  in state  $s_t$  at time  $t$ , the agent obtains a reward  $r_t(s_t, \mathbf{a}_t)$ . Based on the objective of the paper, the sum rate per user group describes the reward and can be written as

$$r_t : \Theta_{sum}^{(t)} = \sum_{u=1}^U \Theta_u^{(t)}, u = 1, \dots, U \quad (20)$$

The DDPG algorithm aims to determine the optimal action that maximizes the Q-value, assessing state–action pair quality to maximize the expected commutative reward under an optimal policy  $\pi$ . This approach facilitates the assessment of the agent’s actions and state transitions based on the given state ( $s_t$ ), action ( $\mathbf{a}_t$ ), and reward ( $r_t$ ). The proposed DDPG algorithm adaptively optimizes the UAV location and RIS phase shifts (actions) in dynamic mmWave urban dense environments for achieving improved signal quality (rewards). This is accomplished by training two main networks, i.e., actor network and critic network. The actor network suggests actions, while the critic network evaluates those actions to enhance overall network efficiency and user experience in FWNs. Algorithm 2 summarizes the proposed DDPD algorithm, while Figure 2 depicts the flow diagram of the proposed scheme for optimizing the UAV location and RIS phase shift to achieve improved network performance. The flow diagram illustrates the interaction between the UAV-RIS system and the DDPG learning agent. The UAV-RIS observes environmental states, including user positions and channel conditions, and takes joint actions of UAV positioning and RIS phase adjustments. These actions yield corresponding rewards based on system performance. The agent stores transition tuples  $(s_t, \mathbf{a}_t, r_t, s_{t+1})$  into the experience replay buffer, which are sampled as mini-batches to train the actor and critic networks. The critic network evaluates the Q-value of actions, while the actor network updates its policy through gradient ascent. Target networks are softly updated to stabilize learning.



**Figure 2.** Flow diagram of the proposed DDPG algorithm.

The step-by-step explanation of the Algorithm 2 is as follows:

The algorithm initializes several key components, including a replay buffer  $D$  for storing learning experiences, a discount factor  $\gamma$  to weigh future rewards in the learning process, and a soft update coefficient  $\tau$  for updating target networks, and it sets the minibatch size  $N_B$  for determining sample transitions from  $D$  during training. Additionally,

actor and critic networks  $\mu(s|\theta^\mu)$  and  $Q(s, a|\theta^Q)$  are initialized with their respective weights  $\theta^\mu$  and  $\theta^Q$ , respectively. To facilitate stable training, the target networks  $\mu'$  and  $Q'$  are initialized with weights  $\theta^{\mu'} \leftarrow \theta^\mu$  and  $\theta^{Q'} \leftarrow \theta^Q$ .

---

**Algorithm 2:** Proposed DDPG Algorithm for Joint Optimization of UAV Location and RIS Phase Shifts

---

**Initialization:**

- Replay buffer  $D$ , discount factor  $\gamma$ , soft update coefficient  $\tau$ , and the minibatch size  $N_B$
- Actor network  $\mu(s|\theta^\mu)$  with weights  $\theta^\mu$ , critic network  $Q(s, a|\theta^Q)$  with weights  $\theta^Q$ , target networks  $\mu'$  and  $Q'$  with weights  $\theta^{\mu'} \leftarrow \theta^\mu$  and  $\theta^{Q'} \leftarrow \theta^Q$

1: **for** episode  $m, \dots, M$  **do**

2:   Get the initial observed state  $s_t$  (18)

3:   Initialize the random process  $\mathcal{N}$  for action exploration

4:   **for each** time step  $t = 1, \dots, T$  **do**

5:     Select action  $a_t = \mu(s_t|\theta^\mu) + \mathcal{N}_t$  based on the current policy and exploration noise

6:     Execute the actions  $a_t$  and observe reward  $r_t$  and new state  $s_{t+1}$

7:     Store transition  $(s_t, a_t, r_t, s_{t+1})$  in replay buffer  $D$

8:     Sample the random  $N_B$  mini-batch transitions  $(s_t, a_t, r_t, s_{t+1})$  from  $D$

9:     Calculate target Q-value by  $z_t = r_t(s_t, a_t) + \gamma Q'(s_{t+1}, \mu'(s_{t+1}|\theta^{\mu'}))|\theta^{Q'}$

10:    Update critic by minimizing the loss:

$$L = \frac{1}{N_B} \sum_t^{N_B} (z_t - Q(s_t, a_t|\theta^Q))^2$$

11:    Update the actor policy using the sampled policy gradient:

$$\nabla_{\theta^\mu} M \approx \frac{1}{N_B} \sum \nabla_a Q(s, a|\theta^Q)|_{s=s_t, a=\mu(s_t)} \nabla_{\theta^\mu} \mu(s|\theta^\mu)|_{s_t}$$

12:    Update the target networks:

$$\theta^{Q'} \leftarrow \tau_{tc} \theta^Q + (1 - \tau_{tc}) \theta^{Q'}$$

$$\theta^{\mu'} \leftarrow \tau_{ta} \theta^\mu + (1 - \tau_{ta}) \theta^{\mu'}$$

13:     $s_t \leftarrow s_{t+1}$

14:   **end for**

15: **end for**

---

In step 2, the algorithm obtains the initial observed state  $s_t$  of the environment by utilizing equation (18). In step 3, the DDPG algorithm initializes noise  $\mathcal{N}$  to facilitate the agent in achieving a better action and state exploration. In step 5, the agent selects an action  $a_t$  based on the current policy, represented by actor network  $\mu(s_t|\theta^\mu)$ , while adding exploration noise  $\mathcal{N}$  to encourage a broader exploration of the action space. This approach allows the agent to discover a wider range of actions during training, thereby improving its ability to learn an optimal policy. Next, in step 6, the agent performs the selected action  $a_t$  in the environment, observes the resulting reward  $r_t$ , and transitions to the next state  $s_{t+1}$ . This step is important for learning from environmental feedback and refining the agent's policy.

In step 7, the algorithm stores the current state, action taken, reward received, and next state observed in the replay buffer  $D$  to retain experiences for learning. To train the neural networks, step 8 involves randomly selecting a mini-batch of  $N_B$  transitions from  $D$ , based on the current state, action, reward, and next state.

In step 9, the algorithm calculates the target Q-value ( $z_t$ ) using the Bellman equation. Here,  $r_t(s_t, a_t)$  is the observed reward after taking  $a_t$  from  $s_t$ , while  $Q'(s_{t+1}, \mu'(s_{t+1}|\theta^{\mu'}))|\theta^{Q'}$  denotes the target Q-value for next state  $s_{t+1}$ . Moving forward, step 10 computes the loss

function  $L$ , which measures the difference between the target Q-value ( $z_t$ ) and training Q-value  $Q(s, a|\theta^Q)$  [37]. This ensures an accurate Q-value estimation and maintains stability throughout the training process. In step 11, the actor policy is updated using a sampled policy gradient, which links the action-value function  $Q(s, a|\theta^Q)$  to the actor's actions  $\mu(s|\theta^\mu)$ , aiming to maximize expected returns computed across a mini-batch of transitions for stable learning. Subsequently, in step 12, the target critic network  $\theta^{Q'}$  and target actor network  $\theta^{\mu'}$  are updated by adjusting their weights towards the main networks  $\theta^Q$  and  $\theta^\mu$ , respectively.  $\tau_{tc}$  and  $\tau_{ta}$  represent the learning rates for the target critic and actor networks' soft updating coefficient, respectively, where  $\tau_{tc}, \tau_{ta} \ll 1$ . The utilization of soft updating prevents the instability and divergence issues typically associated with Q-learning.

Finally, in step 13, the algorithm assigns the next state value  $s_{t+1}$  to the current state  $s_t$  at each time step of an episode during training, effectively updating the current state variable. Steps 5 to 13 are repeated for all time steps  $t = 1, \dots, T$  within an episode to ensure effective learning through experience replay and updates to both the actor and critic networks. The proposed DDPG-based algorithm efficiently handles the continuous action space of joint UAV trajectory and RIS phase shift optimization in dynamic urban environments. Its primary advantage is the ability to learn optimal policies through continuous interaction with the environment without requiring a full system model. This allows for adaptive and real-time optimization under complex mobility and blockage conditions. Additionally, DDPG stabilizes learning using experience replay and target networks, ensuring robust convergence. However, its limitations include a higher computational complexity compared to classical optimization methods, sensitivity to hyperparameter tuning, and potential convergence instability under extremely sparse or highly dynamic training scenarios.

## 4. Numerical Results and Analysis

In this section, we evaluate the effectiveness of the proposed UAV-RIS framework and the DDPG scheme through extensive simulation in terms of EE, HOF, throughput, OP, and LoS probability in an mmWave network. The proposed scheme classifies users according to distances and SINR values while jointly optimizing the UAV trajectory and RIS phase shifts to ensure uninterrupted connectivity for GUs in urban environments. To validate the proposed scheme, we compared it with UAV-only, RIS-only, and without UAV-RIS deployments.

### 4.1. Scenario Setup

Simulations were conducted using MATLAB R2022b on an X-64 processing system with an Intel(R) Core(TM) i7-10510U CPU running at 2.30 GHz and 12 GB RAM. The height, speed, carrier frequency, and transmit power of the UAV-RIS were chosen based on operational standards and 3GPP recommendations. These selections aimed to ensure the relevance of the simulation results to urban scenarios. The urban environment was modeled with varying building area ratio (0.1 to 0.5), number of buildings (750/km<sup>2</sup> to 300/km<sup>2</sup>), and average building height (8 m to 50 m), impacting the signal blockage and LoS availability. The simulation considered different street widths, building distributions, and RIS configurations to reflect real-world challenges such as mobility and dynamic obstacles. The modified K-means algorithm incorporates both spatial proximity and SINR constraints. After initial spatial clustering based on Euclidean distance, SINR and achievable rate thresholds are applied within each cluster to maintain service quality. The algorithm iteratively refines cluster centroids until the cluster memberships stabilize with minimal changes, ensuring convergence. Table 2 summarizes the simulation parameters used in the study.

**Table 2.** Simulation Parameters.

Parameter	Value
Maximum UAV-RIS speed	72 km/h
Maximum UAV-RIS height	500 m
Carrier frequency	100 GHz
Bandwidth	10 GHz
No. of randomly distributed GUs	50
GUs speed	3 km/h
GBS to GU transmit power	40 dBm
UAV-RIS to GU transmit power	30 dBm
PL exponent for LoS and NLoS links	2, 3
GBS antenna spacing	$d = \frac{\lambda}{2}$
Rician factor (R)	2
Discount factor	0.9
Hidden layers	2
Training networks' learning rate	0.001
Target networks' learning rate	0.001
Number of RIS elements	100, 200, 300
RIS elements spacing (vertical and horizontal)	$\frac{\lambda}{4} = 0.75$ mm
Number of episodes	6000
Number of steps per episode	2000
Experience replay buffer size	150,000
Mini-batch size	128

#### 4.2. Simulation Discussion and Comparison

Figure 3 compares the performance of the proposed UAV-RIS framework with three scenarios in terms of EE. The simulation results depict that the proposed framework is more energy-efficient than other benchmarks, as it reduces the energy required for signal transmission and enhances the overall performance of the FWNs. The UAV-only deployment consumes more energy than the RIS-only deployment at all altitudes due to the additional energy required for flight, mobility, and communication. The detailed analysis of the UAV-RIS framework reveals that its EE improves with increasing altitude. This is due to the higher elevation angle between the UAV-RIS and GUs, which reduces obstructions from buildings and other obstacles. Consequently, the LoS probability and the Rician factor increase, strengthening the channel gain and improving SINR. However, EE decreases when the UAV-RIS exceeds a certain altitude, as the increased distance to GUs leads to a higher path loss, ultimately degrading system performance.

The impact of the number of GUs on EE is shown in Figure 4. The simulation outcomes clearly indicate that the system's EE increases with the number of GUs. The proposed DDPG-based UAV-RIS system outperforms the other three setups, primarily due to the joint optimization of UAV mobility and RIS phase shift adjustments, which enhance signal quality and reduce interference, particularly in dense urban environments. For instance, when there are 30 users, the UAV-RIS framework achieves gains of 40%, 31.4%, and 20% compared to the system without UAV-RIS, the RIS-only, and the UAV-only deployment, respectively.

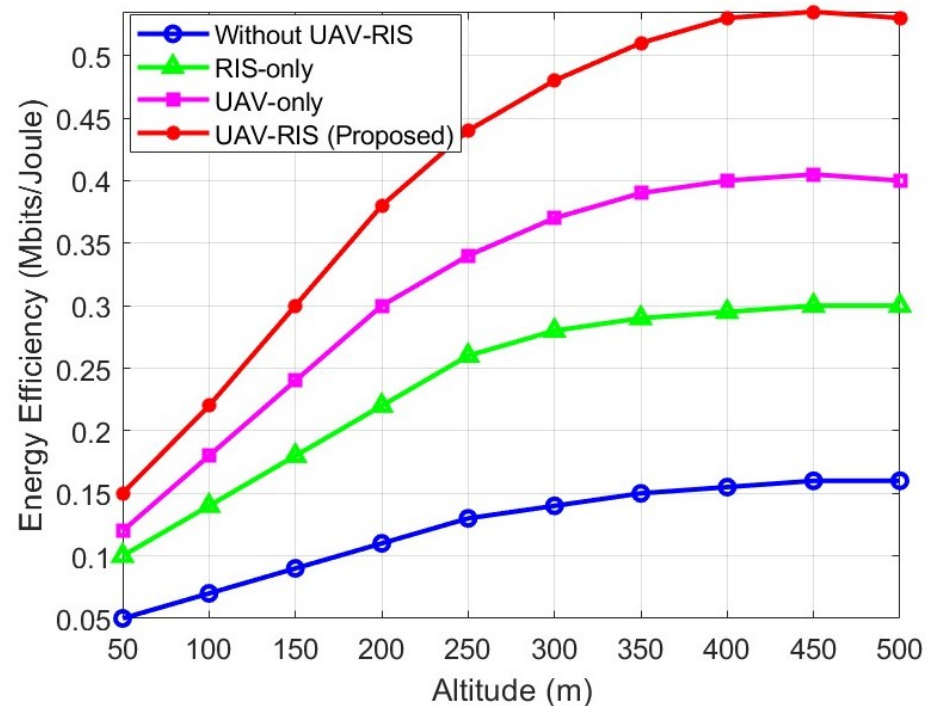


Figure 3. EE comparison of UAV-RIS scheme with benchmarks.

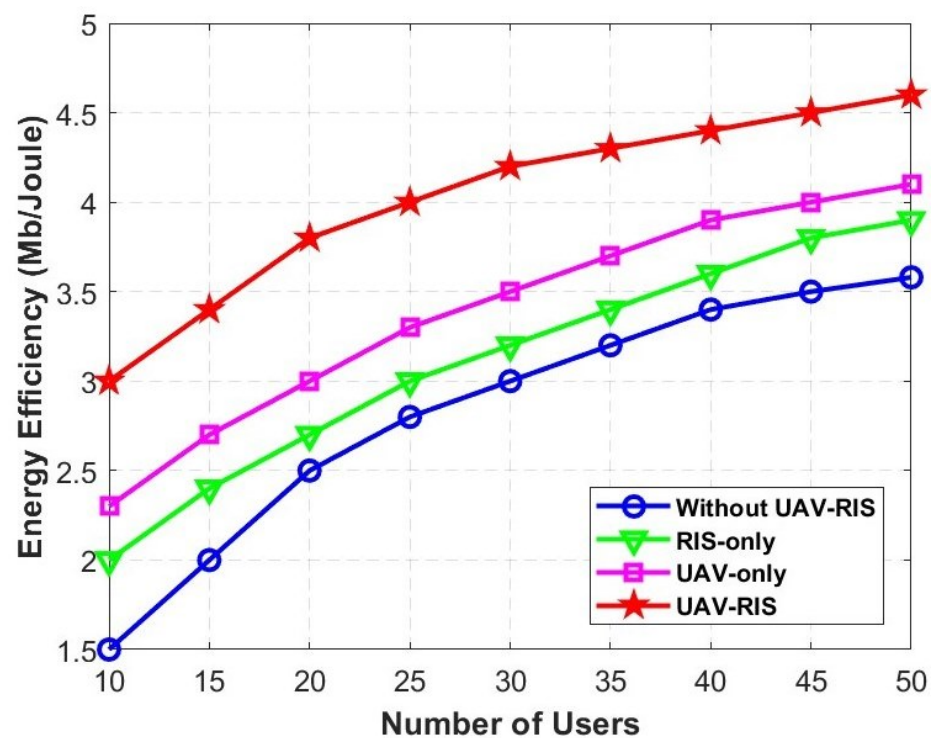
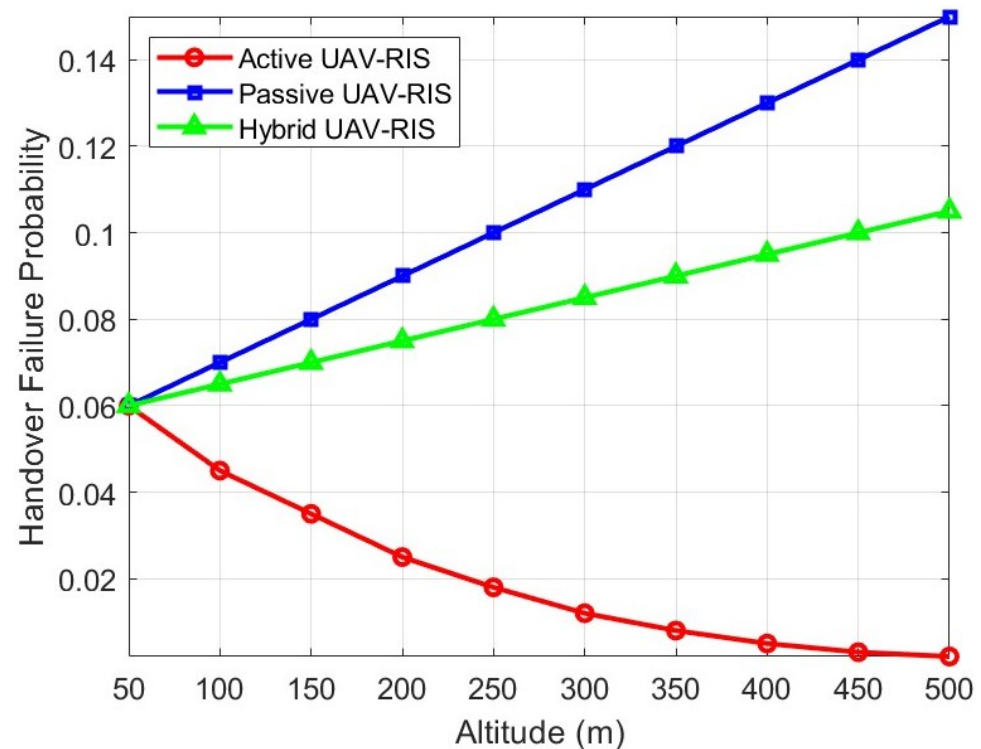


Figure 4. EE vs. number of users.

The hybrid UAV-RIS deployment is optimal for achieving superior system performance in terms of HOF and EE in urban environments. Figure 5 illustrates the effect of various altitudes on HOF for the three configurations: active UAV-RIS, passive UAV-RIS, and hybrid UAV-RIS. It is obvious from the simulation outcomes that at higher altitudes, the active UAV-RIS achieves better performance in reducing HOF because of its ability to amplify and direct the signals to the desired GUs, maintaining reliable connectivity. In contrast, passive UAV-RIS tends to increase the HOF probability at higher heights, as it has the property of reflection only without dynamic adjustment or amplification. This



results in a poor signal strength over greater distances between the UAV-RIS and GUs. Finally, the performance of the hybrid UAV-RIS, which falls between active and passive configurations is analyzed, resulting in a moderate HOF rate across different altitudes. Furthermore, Figure 5 demonstrates that the HOF performance of the hybrid UAV-RIS improves as the number of active elements increases. It reaches the performance level of a fully active RIS-UAV when the number of active elements in the hybrid UAV-RIS equals the total number of elements. Simulation results indicate the superior efficiency of the active UAV-RIS among all configurations in HOF, especially at higher altitudes.



**Figure 5.** HOF vs. UAV-RIS configurations.

Figure 6 shows the throughput comparison between the proposed framework and other benchmarks. The simulation results reveal that the proposed UAV-RIS framework outperforms all others, even at higher altitudes, by dynamically optimizing signal paths and enhancing coverage through UAV positioning and RIS adjustment. The GBS (without UAV-RIS system) delivers satisfactory throughput at lower altitudes but becomes less effective at higher altitudes, making it unsuitable for high-altitude scenarios without UAVs or RIS support. For instance, at 300m height, the proposed UAV-RIS framework achieved 34.48% and 92.8% gains in throughput compared to UAV-only and RIS-only setups, respectively. In contrast, without the UAV-RIS framework, the throughput loss was approximately 66.67%. Overall, the suggested UAV-RIS framework provides the most robust solution to maintain improved throughput across various altitudes.

Increasing the number of GUs in urban environments requires careful optimization of the UAV's altitude and RIS elements. Higher UAV altitudes improve LoS but increase path loss, requiring more RIS elements to maintain signal quality. Figure 7 shows the impact of different UAV altitudes and RIS elements on HO outage probability. Each RIS element actively reflects incident signals with varying phase shifts, and its performance depends on the relative positions of the UAV and the GUs. The simulation results demonstrate that dynamic optimization through the proposed UAV-RIS framework, utilizing the DDPG algorithm, effectively balances connectivity and HO performance in urban dense envi-

ronments. For instance, at a UAV altitude of 350 m with 300 RIS elements, the proposed scheme reduces the OP by 20% compared to the scenario without DDPG.

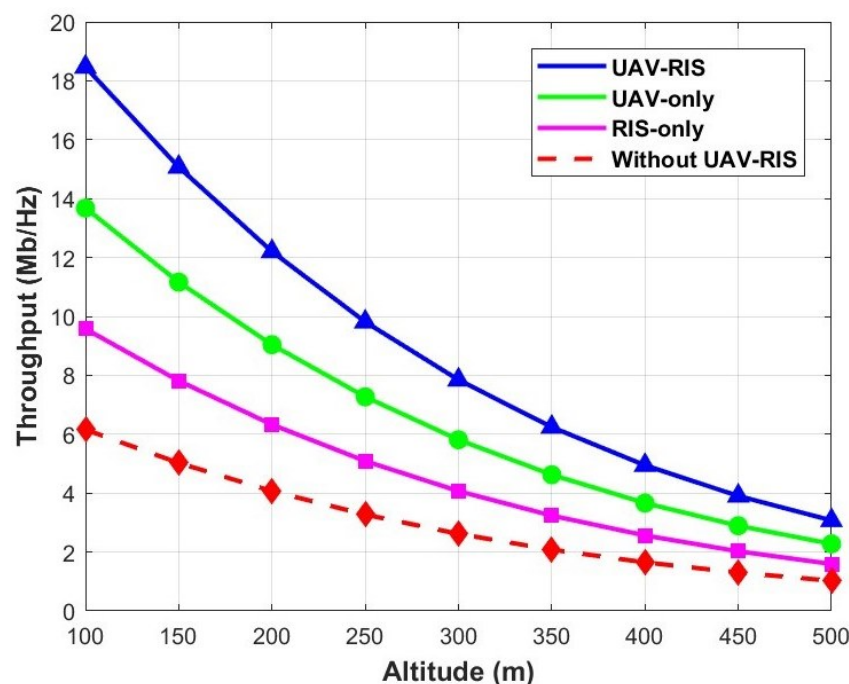


Figure 6. Throughput vs. altitude.

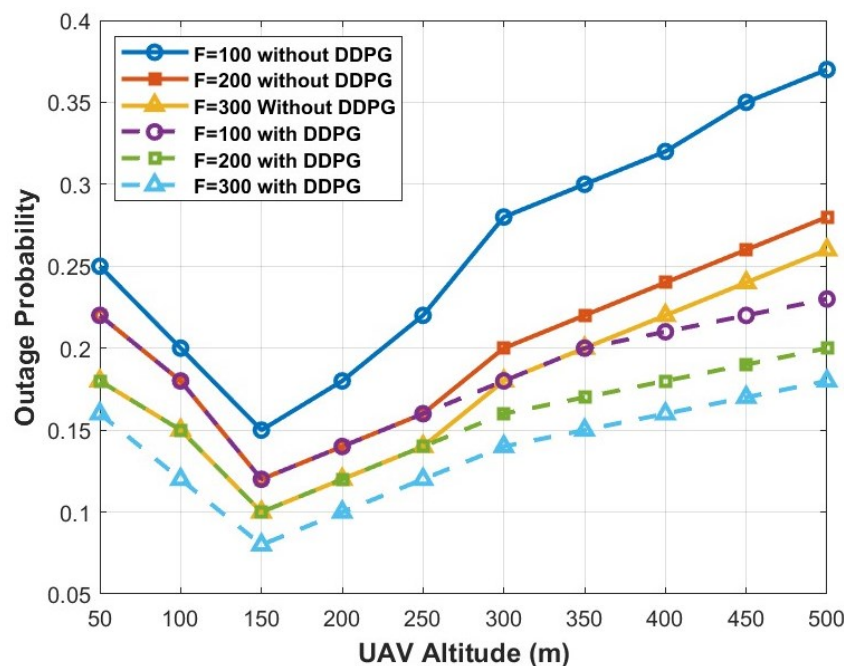
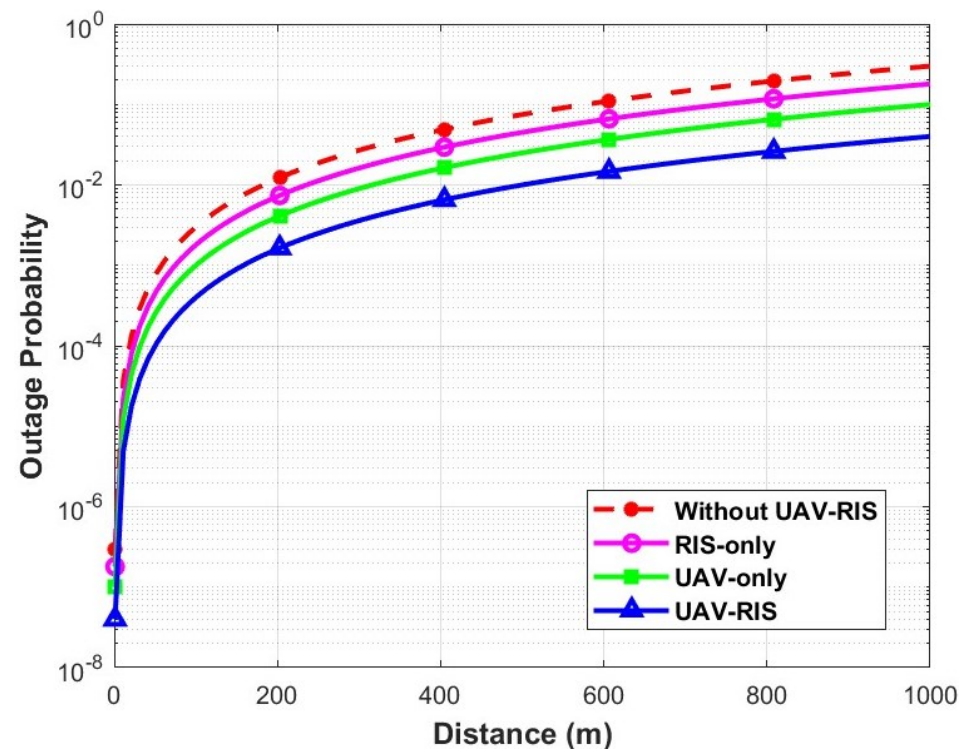


Figure 7. Impact of UAV altitude and RIS elements on OP.

Figure 8 illustrates the OP as a function of distance for UAV-to-ground communication. In the without-UAV-RIS configuration, the OP is relatively high, especially as the distance increases. The RIS-only configuration achieves a moderate reduction in OP by reflecting signals to improve connectivity in LoS-blocked areas, but its effectiveness is limited over longer distances. The UAV-only configuration further reduces the OP by dynamically adjusting the UAV positions to maintain LoS links. However, challenges in HOF and mobility management persist due to UAV movement. The UAV-RIS configuration demonstrates superior performance in reducing OP by integrating UAV mobility with RIS optimization.

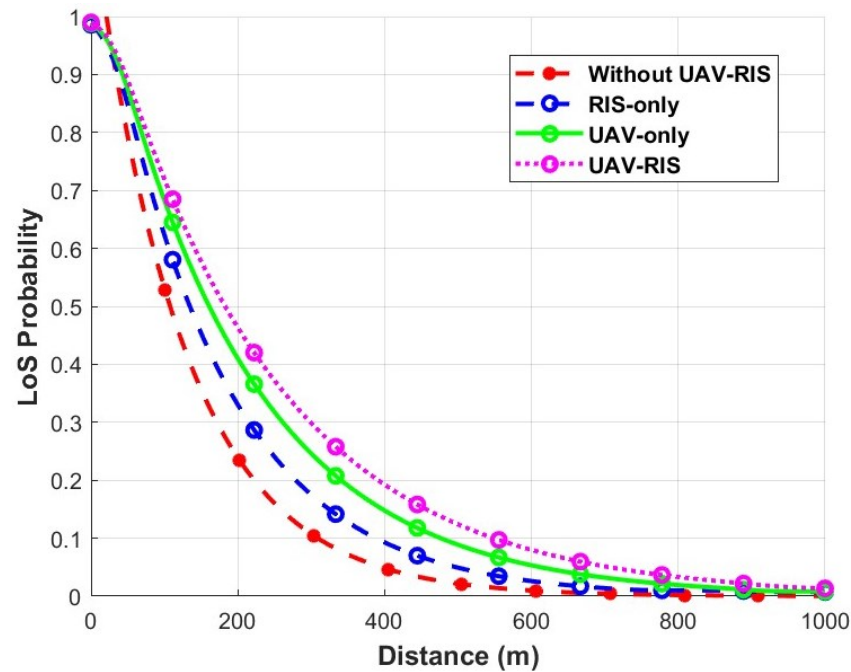
The proposed UAV-RIS system dynamically adjusts the UAV positioning and RIS phase shifts, leading to significant improvements in throughput, EE, and LoS probability.



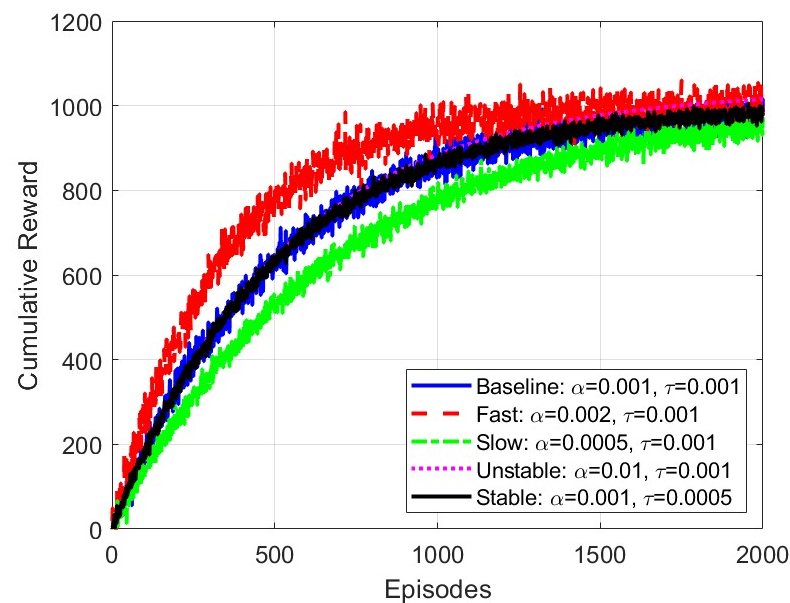
**Figure 8.** OP vs. distance for UAV-to-ground communication.

Figure 9 shows the LoS probability as a function of distance for UAV-to-ground communication in an urban environment, comparing four configurations: without UAV-RIS, RIS-only, UAV-only, and UAV-RIS. The without-UAV-RIS setup experiences a rapid decrease in LoS probability as distance increases. The RIS-only scenario improves performance but still shows a decreasing trend with distance. The UAV-only deployment, compared to the RIS-only and without-UAV-RIS scenarios, provides a better LoS probability due to the UAV's dynamic positioning. It is evident from the simulation outcomes that the proposed UAV-RIS setup achieves the highest LoS probability by combining UAV mobility and RIS, ensuring improved connectivity across all distances. This highlights the effectiveness of UAV-RIS in maintaining seamless connectivity and overcoming LoS blockages, leading to enhanced throughput in urban environments.

The convergence and stability of the proposed DDPG-based UAV-RIS framework are critical for reliable deployment in dense urban environments. Figure 10 shows the cumulative reward convergence under varying hyperparameters. The baseline setting ( $\kappa = 0.001$ ,  $\tau = 0.001$ ) achieved stable convergence after 1500–2000 episodes. Increasing the learning rate ( $\kappa = 0.002$ ) accelerated the convergence but increased the reward variance due to aggressive updates. Lowering the learning rate ( $\kappa = 0.0005$ ) yielded slower but smoother convergence. A high learning rate ( $\kappa = 0.01$ ) resulted in instability and oscillations. Reducing the soft update coefficient ( $\tau = 0.0005$ ) further stabilized convergence by smoothing target network updates. In contrast to the above works, the proposed DDPG-based UAV-RIS framework jointly optimized the UAV trajectory, RIS phase shifts, and user mobility management via modified K-means clustering to address HOFs, LoS connectivity, and energy-efficient deployment in dense urban scenarios.



**Figure 9.** LoS probability vs. distance for UAV-to-ground communication.



**Figure 10.** Convergence behavior of the proposed DDPG-based UAV-RIS optimization framework under different learning rates  $\kappa$  and soft update coefficients  $\tau$ .

The proposed DDPG-based UAV-RIS framework demonstrated performance across multiple key metrics. Specifically, the integration of UAV mobility with RIS phase adaptation significantly reduced HOF probability and OP while improving throughput and energy efficiency in dense urban environments. The simulation results confirm the robustness of the proposed approach in dynamically handling user mobility, LoS blockages, and urban propagation impairments, validating the effectiveness of joint UAV trajectory and RIS phase shift optimization under practical deployment conditions.

## 5. Conclusions and Future Work

The integration of RIS and UAVs is anticipated to enhance wireless network coverage, create additional propagation paths around obstacles, and establish LoS links with distant

GUs, making it a promising technology for FWNs. This study investigated the performance of various frameworks and demonstrated that the proposed UAV-RIS framework, utilizing the DDPG scheme, outperformed the benchmarks in improving connectivity and throughput for GUs. The integration of UAV-RIS with the DDPG-based scheme leads to a significant improvement in addressing the challenges posed by high-mobility scenarios in FWNs. The proposed framework effectively improves LoS probability to GUs by reducing path loss, which is crucial for maintaining reliable communication in urban dense environments. Furthermore, the proposed DDPG algorithm, through the training of two main networks (actor network and critic network), adaptively optimizes the UAV trajectory and RIS phase shifts for enhancing the LoS link and achieving improved signal quality for GUs while transitioning between cells. The simulation results demonstrated that the proposed DDPG-based UAV-RIS framework achieved significant performance improvements. Specifically, the proposed scheme reduced HOF probability by up to 15% compared to a UAV-only deployment. The EE was improved by approximately 40% over baseline scenarios with increasing user densities. Throughput gains of 34.5% and 92.8% were observed compared to UAV-only and RIS-only systems, respectively, at a 300 m altitude. Additionally, the OP was reduced by up to 20% at higher RIS element configurations. These quantitative improvements validate the effectiveness of the proposed joint optimization approach for urban dense UAV-RIS deployments.

#### Future Work

The limited battery life of UAVs and the number of RIS elements can affect the communication performance of GUs in urban scenarios. Additionally, the computational complexity of the DDPG algorithm and reliance on UAVs may limit real-time operations and EE. Future work will address these challenges by optimizing UAV lifetime and the number of active RIS elements to improve EE and reduce HOF through strong LoS connectivity. Furthermore, we will explore multiple UAV-RIS systems and investigate the potential of simultaneously transmitting and reflecting reconfigurable intelligent surfaces (STAR-RIS) to enhance GU mobility in dense urban areas.

**Author Contributions:** Y.U. prepared the original draft of the manuscript and visualized the findings. M.R. and M.A.I. supervised the study, provided funding, and offered critical feedback on the manuscript. The manuscript was reviewed and edited by Y.U., I.O.A., F.A., S.A., A.F.O. and F.Z.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported and funded by a Telekom Malaysia Research & Development (TMR&D) grant, RDTC/241134, MMUE/240095, Malaysia.

**Data Availability Statement:** All the data are available in the paper.

**Acknowledgments:** The authors would like to acknowledge the Center for Wireless Technology, Faculty of Artificial Intelligence and Engineering, Multimedia University Cyberjaya, for providing essential resources, support, and facilities for this study.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Redondi, A.E.; Innamorati, C.; Gallucci, S.; Fiocchi, S.; Matera, F. A Survey on Future Millimeter-Wave Communication Applications. *IEEE Access* **2024**, *12*, 133165–133182. [\[CrossRef\]](#)
2. Roy, S.; Tiang, J.J.; Roslee, M.B.; Ahmed, M.T.; Mahmud, M.A.P. A Quad-Band Stacked Hybrid Ambient RF-Solar Energy Harvester with Higher RF-to-DC Rectification Efficiency. *IEEE Access* **2021**, *9*, 39303–39321. [\[CrossRef\]](#)
3. Roslee, M.B.; Abdullah, R.S.A.R.; Shafr, H.Z. Road pavement density analysis using a new non-destructive ground penetrating radar system. *Prog. Electromagn. Res. B* **2010**, *21*, 399–417. [\[CrossRef\]](#)



4. Banafaa, M.; Shayea, I.; Din, J.; Azmi, M.H.; Alashbi, A.; Daradkeh, Y.I.; Alhammadi, A. 6G mobile communication technology: Requirements, targets, applications, challenges, advantages, and opportunities. *Alex. Eng. J.* **2023**, *64*, 245–274. [[CrossRef](#)]
5. Khan, S.A.; Shayea, I.; Ergen, M.; El-Saleh, A.A.; Roslee, M. An Improved Handover Decision Algorithm for 5G Heterogeneous Networks. In Proceedings of the 2021 IEEE 15th Malaysia International Conference on Communication (MICC), Kuala Lumpur, Malaysia, 1–2 December 2021; pp. 25–30.
6. Li, B.; Fei, Z.; Zhang, Y. UAV communications for 5G and beyond: Recent advances and future trends. *IEEE Internet Things J.* **2018**, *6*, 2241–2263. [[CrossRef](#)]
7. Ullah, Y.; Roslee, M.B.; Mitani, S.M.; Khan, S.A.; Jusoh, M.H. A survey on handover and mobility management in 5G HetNets: Current state, challenges, and future directions. *Sensors* **2023**, *23*, 5081. [[CrossRef](#)]
8. Chen, P.; Luo, L.; Guo, D.; Tang, G.; Zhao, B.; Li, Y.; Luo, X. Why and How Lasagna Works: A New Design of Air-Ground Integrated Infrastructure. *IEEE Netw.* **2024**, *38*, 132–140. [[CrossRef](#)]
9. Elnabty, I.A.; Fahmy, Y.; Kafafy, M. A survey on UAV placement optimization for UAV-assisted communication in 5G and beyond networks. *Phys. Commun.* **2022**, *51*, 101564. [[CrossRef](#)]
10. Dai, M.; Sun, G.; Yu, H.; Wang, S.; Niyato, D. User Association and Channel Allocation in 5G Mobile Asymmetric Multi-Band Heterogeneous Networks. *IEEE Trans. Mob. Comput.* **2025**, *24*, 3092–3109. [[CrossRef](#)]
11. Ullah, Y.; Roslee, M.; Mitani, S.M.; Sheraz, M.; Ali, F.; Aurangzeb, K.; Osman, A.F.; Ali, F.Z. A survey on AI-enabled mobility and handover management in future wireless networks: Key technologies, use cases, and challenges. *J. King Saud Univ. Comput. Inf. Sci.* **2025**, *37*, 1–37. [[CrossRef](#)]
12. Chu, H.; Pan, X.; Jiang, J.; Li, X.; Zheng, L. Adaptive and Robust Channel Estimation for IRS-Aided Millimeter-Wave Communications. *IEEE Trans. Veh. Technol.* **2024**, *73*, 9411–9423. [[CrossRef](#)]
13. Renzo, M.D.; Debbah, M.; Phan-Huy, D.T.; Zappone, A.; Alouini, M.S.; Yuen, C.; Sciancalepore, V.; Alexandropoulos, G.C.; Hoydis, J.; Gacanin, H.; et al. Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come. *EURASIP J. Wirel. Commun. Netw.* **2019**, *2019*, 1–20. [[CrossRef](#)]
14. Huang, S.; Sun, C.; Pompili, D. Meta-ETI: Meta-Reinforcement Learning with Explicit Task Inference for UAV-IoT Coverage. *IEEE Internet Things J.* **2025**.
15. Di Renzo, M.; Zappone, A.; Debbah, M.; Alouini, M.S.; Yuen, C.; De Rosny, J.; Tretjakov, S. Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2450–2525. [[CrossRef](#)]
16. Chen, J.; Cao, K.; Ding, H.; Lv, L.; Ye, Y.; Chi, H.; Yang, L. Double-RIS Enabled Physical Layer Security for Wireless-Powered Communication Systems Over Rayleigh Fading Channels. *IEEE Trans. Commun.* **2025**.
17. Liu, X.; Liu, Y.; Chen, Y. Machine learning empowered trajectory and passive beamforming design in UAV-RIS wireless networks. *IEEE J. Sel. Areas Commun.* **2020**, *39*, 2042–2055. [[CrossRef](#)]
18. Ullah, Y.; Roslee, M.; Mitani, S.M.; Sheraz, M.; Ali, F.; Osman, A.F.; Jusoh, M.H.; Sudhamani, C. Reinforcement learning-based unmanned aerial vehicle trajectory planning for ground users' mobility management in heterogeneous networks. *J. King Saud Univ.-Comput. Inf. Sci.* **2024**, *36*, 102052. [[CrossRef](#)]
19. Liu, Y.; Liu, K.; Han, J.; Zhu, L.; Xiao, Z.; Xia, X.G. Resource allocation and 3-D placement for UAV-enabled energy-efficient IoT communications. *IEEE Internet Things J.* **2020**, *8*, 1322–1333. [[CrossRef](#)]
20. Azari, A.; Ghavimi, F.; Ozger, M.; Jantti, R.; Cavdar, C. Machine learning assisted handover and resource management for cellular connected drones. In Proceedings of the 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, 25–28 May 2020; IEEE: New York, NY, USA, 2020; pp. 1–7.
21. Jiao, L.; Wang, P.; Alipour-Fanid, A.; Zeng, H.; Zeng, K. Enabling efficient blockage-aware handover in RIS-assisted mmWave cellular networks. *IEEE Trans. Wirel. Commun.* **2021**, *21*, 2243–2257. [[CrossRef](#)]
22. Li, S.; Duo, B.; Yuan, X.; Liang, Y.C.; Di Renzo, M. Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming. *IEEE Wirel. Commun. Lett.* **2020**, *9*, 716–720. [[CrossRef](#)]
23. Ma, D.; Ding, M.; Hassan, M. Enhancing cellular communications for UAVs via intelligent reflective surface. In Proceedings of the 2020 IEEE Wireless Communications and Networking Conference (WCNC), Seoul, Republic of Korea, 25–28 May 2020; IEEE: New York, NY, USA, 2020; pp. 1–6.
24. Joshi, N.; Budhiraja, I.; Garg, D.; Garg, S.; Choi, B.J.; Alrashoud, M. Deep reinforcement learning based rate enhancement scheme for RIS assisted mobile users underlaying UAV. *Alex. Eng. J.* **2024**, *91*, 1–11. [[CrossRef](#)]
25. Wei, Z.; Cai, Y.; Sun, Z.; Ng, D.W.K.; Yuan, J.; Zhou, M.; Sun, L. Sum-rate maximization for IRS-assisted UAV OFDMA communication systems. *IEEE Trans. Wirel. Commun.* **2020**, *20*, 2530–2550. [[CrossRef](#)]
26. Yao, Y.; Lv, K.; Huang, S.; Li, X.; Xiang, W. UAV Trajectory and Energy Efficiency Optimization in RIS-Assisted Multi-User Air-to-Ground Communications Networks. *Drones* **2023**, *7*, 272. [[CrossRef](#)]



27. Saif, M.; Valaee, S. Improving Connectivity of RIS-Assisted UAV Networks using RIS Partitioning and Deployment. In Proceedings of the 2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall), Washington, DC, USA, 7–10 October 2024; IEEE: New York, NY, USA, 2024; pp. 1–6.
28. Zhao, H.; Sun, W.; Ni, Y.; Xia, W.; Gui, G.; Zhu, C. Deep Deterministic Policy Gradient-Based Rate Maximization for RIS-UAV-Assisted Vehicular Communication Networks. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 15732–15744. [\[CrossRef\]](#)
29. Jiao, S.; Xie, X.; Ding, Z. Deep reinforcement learning based optimization for IRS based UAV-NOMA downlink networks. *arXiv* **2021**, arXiv:2106.09616.
30. Peng, H.; Wang, L.C.; Li, G.Y.; Tsai, A.H. Long-lasting UAV-aided RIS communications based on SWIPT. In Proceedings of the 2022 IEEE Wireless Communications and Networking Conference (WCNC), Austin, TX, USA, 10–13 April 2022; IEEE: New York, NY, USA, 2022; pp. 1844–1849.
31. Liu, X.; Yu, Y.; Peng, B.; Zhai, X.B.; Zhu, Q.; Leung, V.C.M. RIS-UAV Enabled Worst-Case Downlink Secrecy Rate Maximization for Mobile Vehicles. *IEEE Trans. Veh. Technol.* **2023**, *72*, 6129–6141. [\[CrossRef\]](#)
32. Zhang, X.; Zhang, H.; Liu, L.; Han, Z.; Poor, H.V.; Di, B. Target Detection and Positioning Aided by Reconfigurable Surfaces: Reflective or Holographic? *IEEE Trans. Wirel. Commun.* **2024**, *23*, 19215–19230. [\[CrossRef\]](#)
33. Jiang, F.; Li, T.; Lv, X.; Rui, H.; Jin, D. Physics-Informed Neural Networks for Path Loss Estimation by Solving Electromagnetic Integral Equations. *IEEE Trans. Wirel. Commun.* **2024**, *23*, 15380–15393. [\[CrossRef\]](#)
34. Xu, F.; Duo, B.; Xie, Y.; Pan, G.; Yang, Y.; Zhang, L.; Wang, Y. Multi-UAV Assisted Mixed FSO/RF Communication Network for Urgent Tasks: Fairness Oriented Design With DRL. *IEEE Trans. Veh. Technol.* **2025**, *74*, 1736–1741. [\[CrossRef\]](#)
35. Abbas, Y.; Alarfaj, A.A.; Alabdulqader, E.A.; Algarni, A.; Jalal, A.; Liu, H. Drone-Based Public Surveillance Using 3D Point Clouds and Neuro-Fuzzy Classifier. *Comput. Mater. Contin.* **2025**, *82*, 4759–4776. [\[CrossRef\]](#)
36. Alshehri, M.; Zahoor, L.; AlQahtani, Y.; Alshahrani, A.; AlHammadi, D.; Jalal, A.; Liu, H. Unmanned aerial vehicle based multi-person detection via deep neural network models. *Front. Neurobot.* **2025**, *19*, 1582995. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Xu, K.; Wei, A.; Zhang, C.; Chen, Z.; Lu, K.; Hu, W.; Lu, F. HiFusion: An Unsupervised Infrared and Visible Image Fusion Framework With a Hierarchical Loss Function. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 5015616. [\[CrossRef\]](#)

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.