# Roads Digital Twin: Predictive Situational Awareness using 360° Video Streaming and Graph Neural Networks

Sotirios Messinis, Oussama El Marai, Nicholas E. Protonotarios, Tarik Taleb, *Senior Member, IEEE,* and Nikolaos Doulamis, *Member, IEEE*

***Abstract*—As vehicle technologies rapidly advance, video streaming capabilities emerge as a significant feature of modern on-board vehicular systems. The integration of 360° video streaming in vehicles is expected to enhance road situational awareness, by providing panoramic live streaming and recording capabilities. This paper introduces the concept of *roads digital twin*, combining, for the first time, 360° video streaming with graph neural networks, in order to enhance predictive situational awareness in road environments. To this end, we have developed eGAT, a selective edge-enhanced graph attention network architecture, that uses graph neural networks with attention mechanisms. eGAT is capable of effectively predicting road coverage, considering uplink bandwidth limitations that may affect video streaming quality and user quality of experience (QoE). For the evaluation of our novel method, we utilized four different datasets, considering several vehicular scenarios. For performance comparison purposes, we employed three metrics, namely the overall percentage of the covered region, the normalized mutual information (NMI), and the precision-recall scores. In terms of overall coverage percentage, eGAT provided superior coverage performance against similar studies in all nine scenarios and for all numbers of streaming vehicles investigated, reaching an increase of up to 32.3%. In terms of NMI score, for low values of prediction horizon, eGAT outperformed similar attention-based algorithms, with an increase of up to 29.6%, whereas for larger prediction horizons, eGAT presented lower, however comparable, coverage performance. Our results suggest that eGAT is a promising solution for scenarios involving streaming vehicles, showcasing its potential for applications such as autonomous vehicles and traffic management systems.**

***Index Terms*—graph neural networks, bandwidth allocation, optimal area coverage, 360° video streaming, situational awareness, digital twins**

## I. Introduction

VEHICLE technologies are evolving rapidly with growing complexity, encompassing several new onboard products and infrastructure-based services [1]. These advancements are expected to significantly transform vehicle ecosystems in the future. Among the critical technologies they need to incorporate into their functionalities, video cameras have been the front-runner for several promising applications. In particular, 360° video streaming enables vehicles to live stream or record video content, offering panoramic views, thus enhancing their road awareness [2].

In this direction, the integration of 360° video recording capabilities into vehicles is significantly enhancing road awareness and functionality. Several of the so-called "dash cameras" are equipped with these capabilities. Mounted on the vehicle's dashboard or windshield, these cameras can capture video footage from all angles. The recorded video can be used for insurance claims, documenting road trips, or sharing driving experiences [3]. In the context of autonomous vehicles, omnidirectional video streaming can be used for surveillance and remote control [4]. Multiple cameras placed around the vehicle may capture a complete overview of the environment, allowing operators or remote monitoring systems to assess the driving environment in real-time [5]. Future vehicles may also use 360° video streaming to help drivers maneuver in tight spaces, thus enhancing safety. By accurately indicating blind spots and obstacles, these video cameras may help prevent accidents and improve overall situational awareness [6].

360° dash cameras play an active role in the emergence of the concept of smart cities, which is expected to improve quality of life. The concept of the digital twin for smart cities has emerged quite recently, combining the advantages of relevant state-of-the-art technologies with specific functionalities of smart cities [7]. To this end, a road digital twin is a dynamic virtual replica of the physical road network that integrates real-time data from various sensors, including 360° cameras on vehicles, traffic cameras, weather stations, and GPS devices [8]. Continuously updated, it reflects current road conditions, traffic flow, and environmental factors in a three-dimensional, high-resolution model. Furthermore, road digital twins enhance situational awareness and decision-making by providing real-time monitoring and simulation capabilities, aiding in traffic management and emergency response. In smart cities, road digital twins support urban planning and integrate with intelligent transportation systems (ITS) for adaptive traffic control and automated incident detection, thus improving traffic efficiency and safety [9].

Creating real-time smart city digital twins and enhancing situational awareness (SA) can be achieved through frameworks utilizing 360° omnidirectional cameras mounted on vehicles, as demonstrated by the authors of [10].The cameras are included in digital twin boxes (DTBs), covering a region as they pass through a road network. DTBs create a digital twin of the physical road asset by constantly sending real-time data to the edge or cloud, including the 360° live stream, GPS location, and measurements of the temperature, and humidity, see Figure 1. Optimal area coverage is required

This work was partially conducted at ICTFICIAL Oy. It is partially supported by the European Union's HE research and innovation program HORIZON-JUSNS-2023 under the 6G-Path project (Grant No. 101139172). The paper reflects only the authors' views, and the European Commission bears no responsibility for any utilization of the information contained herein.

Sotirios Messinis and Nikolaos Doulamis are with the School of Rural, Surveying and Geoinformatics Engineering, National Technical University of Athens, Greece.

Oussama El Marai is with the Department of Communications and Networking, School of Electrical Engineering, Aalto University, Finland.

Nicholas E. Protonotarios is with the Mathematics Research Center, Academy of Athens, Greece.

Tarik Taleb is with the Faculty of Electrical Engineering and Information Technology, Ruhr University Bochum, Germany.

in several applications, including surveillance and monitoring [11], search and rescue operations [12] and infrastructure inspection [13]. The DTBs aim to enable efficient coverage of large areas while reducing costs, optimizing resource usage and considering the limited network bandwidth in the area under investigation. SA involves the optimal selection of a subset of vehicles that ensures the maximum area coverage, while preventing the degradation of the user's Quality of Experience (QoE) [14].

In this paper, we solve the problem of area coverage under certain bandwidth allocation constraints by employing an application-oriented neural network architecture. In particular, we employ graph neural networks (GNNs) in order to provide optimal area coverage while considering optimal bandwidth allocation. Bandwidth allocation corresponds to specific video qualities in the predicted trajectories of the vehicles. Indicative vehicle trajectories are illustrated in Figure 2. The key innovation of our study lies in the development of road digital twins with predictive situational awareness capabilities based on distributed 360° video streaming sources for optimal area coverage considering video quality requirements. Optimal area coverage is necessary for maintaining the integrity of digital twin services, while the allocated video quality directly impacts their performance and enables additional functionalities such as object detection tasks. We use GNNs in order to effectively address the corresponding trade-offs between optimal area coverage and bandwidth allocation. Unlike previous works that focus solely on video streaming or GNNs, our work is based on the initial work of [10], in the framework of the project 6Genesis [15], aiming to develop a roads infrastructure digital twin. While our GNN-based architecture does not contribute to graph neural network theory or standalone video streaming technologies, however it offers a clear and innovative contribution to the creation of enhanced digital twin models within the framework of smart cities and future urban networks. By integrating several optimally selected 360° video data sources, our architecture provides richer situational awareness and more accurate predictions than methods relying on fewer one-dimensional data sources.

To the best of our knowledge, in the context of 360° video streaming, there is no existing solution that employs area coverage and bandwidth allocation combined. In this direction, we aim to enhance predictive situational awareness through the digital twinning of smart cities and their inclusive connected mobility. The rest of this paper is organized as follows. In Section II, we provide the relevant literature work on bandwidth allocation and area coverage. In Section III, we present our system model and its corresponding formulation. In section IV, we introduce and present our suggested architecture. Finally, in section V, our proposed solution is demonstrated with specific vehicular scenarios and our experimental results are discussed. In Section VI we present our concluding remarks.

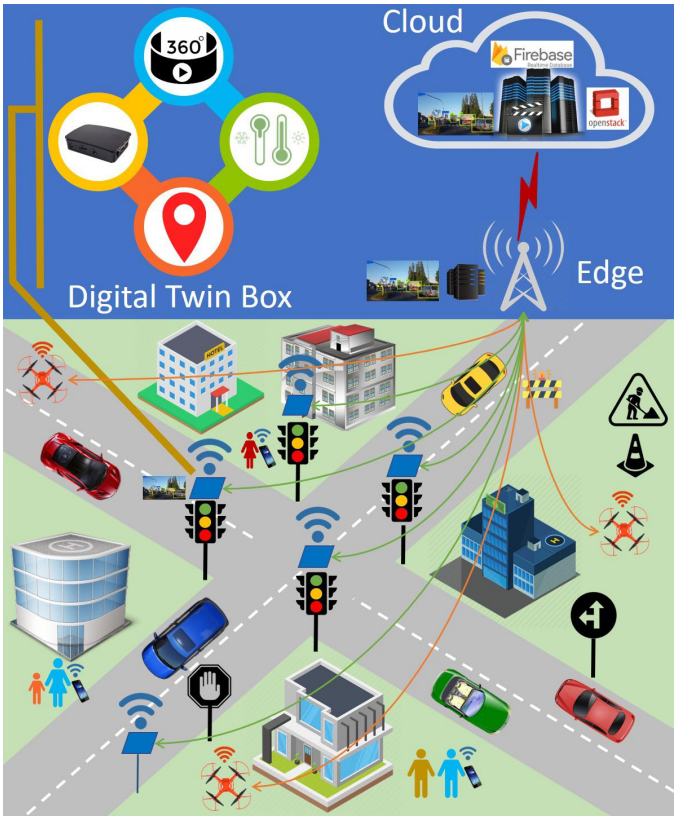


Fig. 1. DTB global system architecture: DTBs create digital twins of roads and send real-time data to the edge or cloud, including 360° live streams, GPS locations, and measurements of temperature and humidity [10].

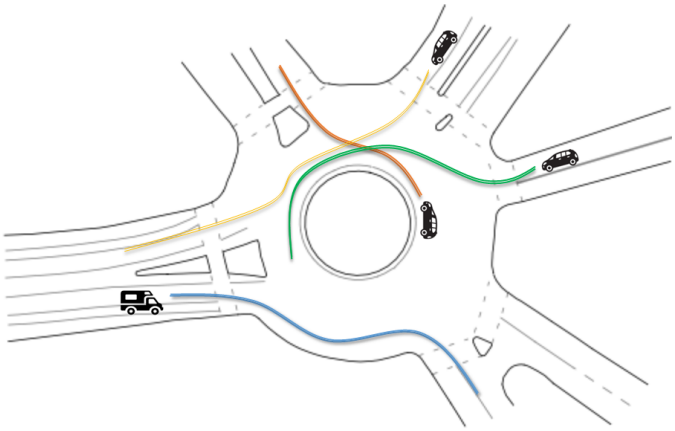

Fig. 2. Vehicles' trajectories.

## II. Related Works

Recent advancements in vehicular networks have extensively explored bandwidth allocation, trajectory prediction, and resource allocation, often leveraging deep learning. Bandwidth allocation studies have focused on optimizing network resources using blockchain technology and reinforcement learning to address privacy, authenticity, and efficiency issues in the Internet of Vehicles (IoV). Research in area coverage has focused on developing algorithms that enable fleets of vehicles or drones to efficiently cover designated areas, aiming to maximize coverage while minimizing energy consumption or total distance traveled. Trajectory prediction has benefited from GNNs in modeling interactions among vehicles for precise forecasting in autonomous driving scenarios. Resource allocation with GNNs has proposed decentralized solutions for wireless networks and IoV, optimizing communication,

computing, and caching resources. However, a gap exists in integrating these areas to jointly address area coverage and bandwidth allocation, particularly for enhancing road infrastructure digital twins. This paper addresses this gap by proposing a GNN-based approach that simultaneously optimizes area coverage and bandwidth allocation constraints, thereby improving situational awareness and efficiency in vehicular networks.

### *A. Bandwidth Allocation*

In terms of network resource deployment, the authors of [16] integrate blockchain technology to address privacy and authenticity issues in the IoV. They propose an intelligent resource allocation framework, modeled as a Markov decision process and optimized using the asynchronous advantage actor-critic approach. In [17], the authors address the problem of supporting multimedia services with ultra-low latency, under extensive computation requirements on resource-constrained end-user devices.

From a machine learning point of view, the authors of [18] address the problem of transmission resource allocation in connected vehicles by proposing a deep reinforcement learning algorithm. This specific algorithm enables them to minimize the energy consumption of roadside sensors and ensure the freshness of high-definition map content in vehicles. A systematic mapping study regarding the deployment of 5G in modern smart cities is presented in [19], where several metrics for the evaluation of bandwidth allocation to specific assets are considered. In addition, in [20], in order to maximize the communication capacity of the unmanned aerial vehicle (UAV) edge computing network, the authors investigate bandwidth allocation and trajectory control by utilizing multiagent reinforcement learning algorithms. The problem of resource allocation in the IoV is the main focus of [21], where deep reinforcement learning (DRL) is deployed to address the corresponding optimization problem. Furthermore, to address the bandwidth allocation sub-problem of caching popular content at the wireless edge, an innovative DRL-based vehicle mobility model is introduced in [22]. In [23], a novel channel bandwidth allocation strategy for wireless ad hoc network systems is proposed; this strategy aligns the number of control channel time slots with the number of on-board units (OBUs) to reduce congestion of channel access demand information. The bandwidth segments are dynamically adjusted for quickly arriving and leaving OBUs in the coverage of roadside base stations.

### *B. Area Coverage*

Ongoing research focuses on the development of algorithms aimed at efficiently covering a given area with a fleet of vehicles, including drones, robots, and autonomous vehicles [24]. In multi-agent area coverage, the goal is to find the optimal trajectories for the vehicles to ensure maximum coverage while minimizing redundant or overlapping efforts [25]. Allocating tasks among multiple vehicles is another important aspect of determining which areas each vehicle must cover, considering factors like proximity, coverage history, and workload balancing [26]. In a similar context, coverage problems are associated with several conflicting objectives, including the need to maximize coverage while minimizing either energy consumption or distance [27]. Researchers are employing a variety of algorithms to address the area coverage challenges, including graph theory, reinforcement learning, and several other heuristic approaches [28].

The authors of [29] introduce a distributed and online heuristic strategy for solving the problem of collaborative coverage using multiple UAVs in target search scenarios. Furthermore, in [30], the authors focus on the coverage path planning problem for heterogeneous UAVs in large-scale cooperative search systems with multiple separated regions. The authors of [31] propose a novel approach for online coverage path planning in unknown environments using cooperative multi-robotic agents. The method accelerates coverage by optimizing distributed multi-agent planning with dynamic programming. In another study, the authors address the area coverage problem for multiple capacity-constrained robots [32].

### *C. Trajectory Prediction with GNNs*

Trajectory prediction is a thoroughly researched field that has recently been enhanced significantly by the employment of deep learning algorithms. In this direction, GNNs have proven to be a suitable approach for this type of problem, as they take into account the graph-related attributes resulting from the interactions between vehicles and other entities. In [33], the authors develop GroupNet, a novel multiscale hypergraph neural network that facilitates comprehensive modeling of interactions for precise trajectory prediction. The authors of [34] introduce CRAT-Pred, a novel, map-free trajectory prediction model for autonomous vehicles. By leveraging a graph convolution method from material science and integrating multi-head self-attention, CRAT-Pred effectively models social interactions among vehicles. In [35], the authors introduce a graph-based spatial-temporal convolutional network, tailored for forecasting the future trajectory distributions of neighboring vehicles using historical trajectories. The model employs a gated recurrent unit network to encode and decode the spatial-temporal features and generate future trajectory distributions.

The authors of [36] address the challenge of predicting the motion trajectories of moving agents in complex traffic scenarios with the introduction of a directed graph convolutional neural network. The study [37] presents a novel and scalable approach (SCALE-Net) based on an edge-enhance graph convolutional neural network to address the challenges associated with predicting the trajectories of surrounding vehicles in autonomous driving scenarios. In [38], the authors introduce a three-channel framework incorporating a novel heterogeneous edge-enhanced graph attention network (HEAT), in order to discern distinct motion patterns among agents and their interdependencies with surrounding agents and traffic structures. To the same extent, the authors of [39] focus on the attention-based interaction-aware trajectory prediction (AI-TP) for self-driving vehicles. AI-TP employs an encoder-decoder architecture with graph attention networks (GAT) and convolutional gated recurrent units to predict the trajectories of traffic agents.

In [40], the authors present GATraj that utilizes attention mechanisms to capture spatial-temporal dynamics of agents and employs a graph convolutional network to model their interactions achieving state-of-the-art prediction performance.

### *D. Resource Allocation with GNNs*

In terms of resource allocation, the authors of [41] propose a GNN approach for addressing decentralized resource allocation challenges in wireless networks. In a similar context, the authors of [42] utilized GNNs to learn the low-dimensional feature of each node based on the graph information. According to the learned feature, multi-agent reinforcement learning is then used to make spectrum allocation in a vehicle-to-everything (V2X) network. In [43], the authors introduce a cache-aided multi-access edge computing (MEC) offloading framework that optimizes communication, computing, and caching (3C) resources within the MEC-enabled IoT. The authors of [44] propose a regularized unsupervised learning framework that solves wireless resource allocation problems by utilizing regularization techniques to minimize the risk of constraint violations during training. In addition, the authors of [45] have developed a distributed power allocation scheme for interference-limited wireless networks with the use of GNNs. Furthermore, the authors of [46] propose an unsupervised resource allocation approach, based on GNNs. Acknowledging the significance of interactions among allocation requirements, they introduce a method to optimize a global utility function by autonomously determining how resources should be optimally allocated. In this paper, we adopt the approach of [46] for the required bandwidth allocation to the selected vehicles for video streaming.

## III. System Model and Formulation

The objective of this work is to maximize the SA of a specific region denoted by $\mathcal{R}_\alpha$, $\alpha = 1, 2, \ldots, A$. Each such region contains a random number of vehicles, $n$, that might be equipped with mobile DTBs [10]. We denote by $\mathcal{D}_\alpha$,

$$\mathcal{D}_\alpha = \{d_\alpha^{(1)}, d_\alpha^{(2)}, \ldots, d_\alpha^{(M)}\}, \tag{1}$$

the set of available DTBs in the region $\mathcal{R}_\alpha$, where $M$ denotes the number of DTBs in that region. Each DTB, $d_\alpha^{(m)}$, with $\alpha = 1, 2, \ldots, A$, and $m = 1, 2, 3, \ldots, M$, has a specific GPS coordinate $g_\alpha^{(m)}$. The set of GPS coordinates of a given snapshot of the system, corresponding to the set of DTBs $\mathcal{D}_\alpha$ co-located in the same region $\mathcal{R}_\alpha$, is denoted by $\mathcal{G}_\alpha$, namely:

$$\mathcal{G}_\alpha = \{g_\alpha^{(1)}, g_\alpha^{(2)}, \ldots, g_\alpha^{(M)}\}. \tag{2}$$

The input of the model consists of the historical states, $\mathcal{S}_t$, of the vehicles, and a map. At time $t$, each vehicle's historical states are represented by $\mathcal{S}_t$, i.e.,

$$\mathcal{S}_t = \{\mathcal{S}_t^{(1)}, \mathcal{S}_t^{(2)}, \ldots, \mathcal{S}_t^{(N)}\}. \tag{3}$$

Historical states of vehicle $i$ at time $t$ are represented by:

$$\mathcal{S}_t^{(i)} = \left\{ s_{t-T_h+1}^{(i)}, s_{t-T_h+2}^{(i)}, \ldots, s_t^{(i)} \right\}, \tag{4}$$

with $T_h$ denoting the traceback horizon. The state $s_t^{(i)}$ is the vehicle $i$'s position and velocity, i.e.,

$$s_t^{(i)} = \left( x_t^{(i)}, y_t^{(i)}, u_{x,t}^{(i)}, u_{y,t}^{(i)} \right), \tag{5}$$

where $(x_t^{(i)}, y_t^{(i)})$ and $(u_{x,t}^{(i)}, u_{y,t}^{(i)})$ denote the usual Cartesian coordinates and velocities in the $x$ and $y$ directions, respectively, computed at time $t$ for the vehicle $i$. The expected output, denoted by $\mathcal{F}_t$, will generate predicted trajectories of $l \leqslant M$ vehicles, namely:

$$\mathcal{F}_t = \left\{ \mathcal{F}_t^{(1)}, \mathcal{F}_t^{(2)}, \ldots, \mathcal{F}_t^{(l)} \right\}, \tag{6}$$

where $\mathcal{F}_t^{(i)}$ represents a sequence of the predicted 2D coordinates, $(x_i, y_i)$, of vehicle $i$ over a prediction horizon $T_f$, i.e.,

$$\mathcal{F}_t^{(i)} = \left\{ \left( x_{t+1}^{(i)}, y_{t+1}^{(i)} \right), \ldots, \left( x_{t+T_f}^{(i)}, y_{t+T_f}^{(i)} \right) \right\}, \tag{7}$$

from time $t+1$ to $t+T_f$.

We consider 360° video cameras capable of streaming a spherical view of the environment. For computational and demonstration reasons, we assume all cameras have the same lens focal length (LFL). The cameras are able to stream live using DASH technique at $K$ predefined video qualities, namely

$$\mathcal{Q} = \{q_1, q_2, \ldots, q_K\}, \tag{8}$$

where $q_1$ and $q_K$ represent the lowest and highest video qualities, respectively. It is worth mentioning that we denote by $B$ the total uplink bandwidth available in the whole region.

Furthermore, we aim to enhance the SA in a region by covering the area through the field-of-view of the DTBs' cameras, while taking into account the number of video sources and the user's QoE. Due to the limited uplink capacity $B$, and in order to avoid QoE degradation, only a subset of DTBs, namely $\mathcal{L}_\alpha \subseteq \mathcal{D}_\alpha$, referred to as *active* DTBs, are allowed to stream. The DTB selection process is followed by a video quality identification process, performed in order to assign a video quality $q_i \in \mathcal{Q}$, $i = 1, 2, \ldots, K$, to each DTB $d_\alpha^{(m)} \in \mathcal{L}_\alpha$.

Our study implements graph neural networks in area coverage and bandwidth allocation. Graph neural networks are a class of neural network models developed for processing structured data in the form of graphs. Graphs consist of nodes (vertices) and edges (links), which connect pairs of nodes, as in Figure 3. The main idea behind GNNs is to learn embeddings for each node in a graph by aggregating information from its neighbors [47], [48]. A graph $G$ is represented as

$$G = (V, E), \quad \text{with} \quad E \subseteq V \times V, \tag{9}$$

where $V$ and $E$ denote the sets of $N$ nodes and of $L$ edges of the graph, i.e.,

$$V = \{v_1, v_2, \ldots, v_N\}, \quad E = \{e_1, e_2, \ldots, e_L\}. \tag{10}$$

We note that, for computational convenience, instead of the notation provided by the second of equations (10), we will denote an edge connecting nodes $i$ and $j$ of the graph by $e_{ij}$.

The connections of the graph are represented as an adjacency matrix $A \in \mathbb{R}^{N\times N}$, where $A_{ij} \neq 0$ if there exists an edge from node $v_i$ to node $v_j$ and $A_{ij} = 0$ otherwise. The basic building block of GNNs is the graph convolutional neural network (GCN). GCNs define a message-passing scheme, where each node aggregates information from its neighbors and updates its representation. The aggregation step typically involves computing a weighted sum or applying a pooling operation over the neighbor representations. Then, a node-specific transformation is applied to the aggregated information, followed by non-linear activation. In our work, the nodes represent the vehicles moving in the area, and the edges are attributed to the distances between them. Subsequently, the adjacency matrix $A_{ij}$ represents the distances among all the vehicles at a specific time $t$.

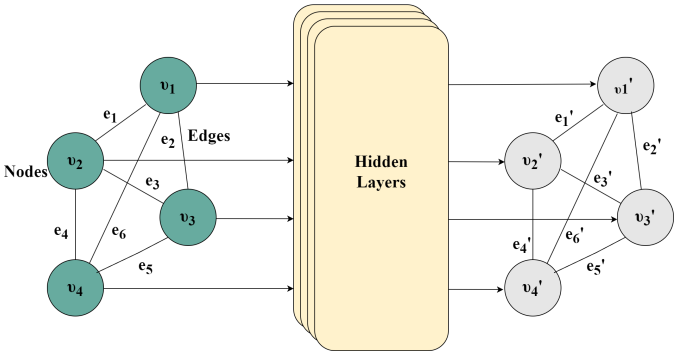


Fig. 3. A graph neural network (GNN), with nodes and edges denoted by $v_n$ $(n = 1, \ldots, N)$ and $e_\ell$ $(\ell = 1, \ldots, L)$, respectively.

The availability of labeled nodes and edges in the graph is crucial for supervised learning tasks. GNNs can also be used for unsupervised and semi-supervised learning by leveraging graph structure and node attributes [49]. Among several variants and extensions of GNNs that have been proposed, including GraphSAGE [50] and graph isomorphism networks (GINs) [51], we have selected graph attention networks (GATs) [52] as the most appropriate for our system model, due to their proven applicability in vehicle trajectory prediction problems [37]. Furthermore, GATs often introduce additional mechanisms in order to enhance the information aggregation process. In particular, to highlight the importance of different nodes in the graph during information aggregation, GATs utilize attention mechanisms. The key idea of GATs is to assign attention coefficients to the neighboring nodes of the target node, indicating their relative importance for information aggregation, see Figure 4. For a given target node, the model computes attention coefficients that are typically calculated using a shared attention mechanism, which involves a learning-oriented attention weight matrix and a non-linear activation function, such as LeakyReLU [53], softmax, and certain others [54].

For the purposes of our study, we model the vehicles as nodes in a graph, with edges representing the distances between them. By interconnecting all neighboring vehicles through edges, we construct a graph that serves as input to our system model.

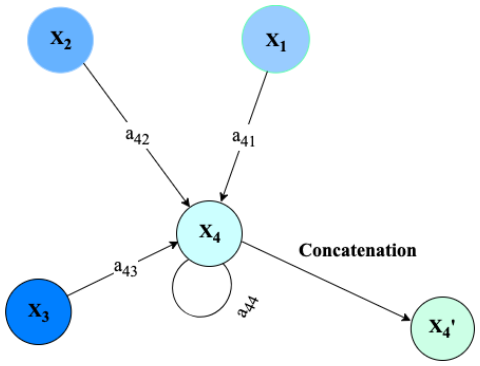


Fig. 4. A graph attention network (GAT), assigning attention coefficients to the neighboring nodes of the target node.

## IV. Predictive Situational Awareness with GNN and 360°video streaming

In this section, we describe our proposed GNN architecture for maximizing SA, based on the predicted trajectories of the vehicles, and on the selected DTBs along with their 360° video streaming capability. Our architecture is able to select the DTBs that enable wider video coverage of a particular area, maintaining high levels of QoE. It consists of two types of sequential neural networks that give the trajectory prediction of the vehicles, and their corresponding coverage with selected video streaming qualities, respectively, as shown in Figure 5. It is worth mentioning that the prediction mechanism is deployed on the server infrastructure and is applied to a specific area for its digital twin implementation. The corresponding architecture consists of two phases, namely, one addressing coverage, and one addressing bandwidth allocation.

TABLE I
eGAT architecture notation

| | |
|---|---|
| $r_t^{(i)}$ | dynamics feature of vehicle $i$ at time $t$ |
| $\mathcal{R}_t$ | set of dynamics features of all vehicles at time $t$ |
| $s_t^{(i)}$ | historical states of vehicle $i$ at time $t$ |
| $\mathcal{U}_t$ | set of interaction features of all vehicles at time $t$ |
| $u_t^{(i)}$ | interaction feature of vehicle $i$ at time $t$ |
| $E_t$ | edge set of the graph of vehicles at time $t$ |
| $e_{ij}$ | edge attribute between nodes $i$ and $j$ |
| $e_{ij}^{+}$ | concatenated edge feature |
| $f_t^{(i)}$ | trajectory features of vehicle $i$ at time $t$ |
| $m_t^{(i)}$ | map selector of vehicle $i$ at time $t$ |
| $h_i$ | features of node $i$ |
| $\mathcal{H}$ | set of all nodes features $i$ |
| $\mathbf{a}$ | attention mechanism |
| $a_{ij}$ | attention coefficient from node $j$ to target node $i$ |
| $W_h$ | weighted sum of node features $h$ |
| $h_i^{new}$ | updated output feature |
| $J_{enc}^{u}$ | encoding of the video covered surfaces |
| $J_{enc}^{e}$ | encoding of the distance between vehicles |
| $J_{dec}^{u}$ | final state decoding |

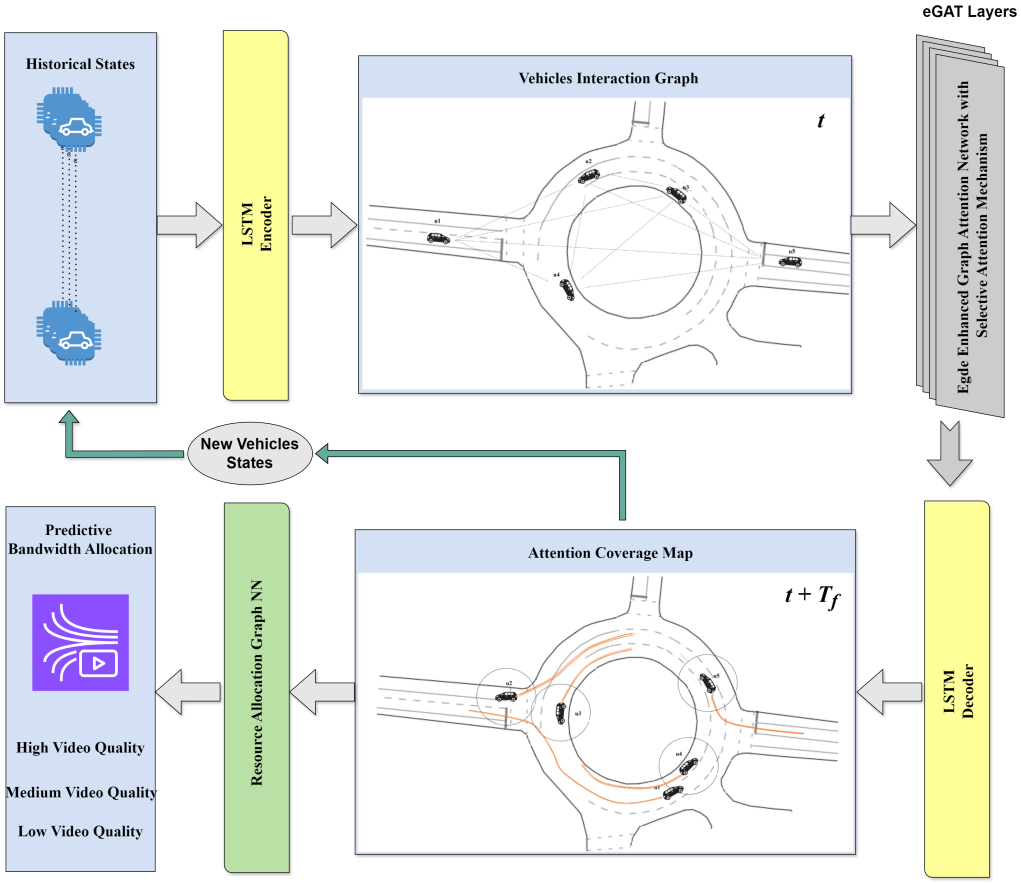


Fig. 5. eGAT architecture for predictive situational awareness using 360° video streaming, consisting of two types of sequential neural networks that give the trajectory prediction of the vehicles, and their corresponding coverage with selected video streaming qualities, respectively.

### A. Coverage Prediction

The first phase of the architecture relies on a multi-agent trajectory prediction approach, where the vehicles leverage both their historical states and the layout of the road infrastructure considering an adaptive map selector. As the vehicle's historical states are represented by a temporal sequence, several neural network encoders may be used for the underlying sequence modeling, namely recurrent neural networks (RNNs), long short-term memory (LSTM) and gated recurrent units [55]. In this work, we deploy LSTM as the encoder of choice for the historical states of vehicles. The dynamics feature of vehicle $i$ at time $t$, denoted by $r_t^i$, is calculated by

$$r_t^{(i)} = \text{LSTM}\left(s_t^{(i)}\right), \tag{11}$$

with LSTM representing the encoder of its corresponding historical states. The dynamics features of all vehicles will then be gathered as:

$$\mathcal{R}_t = \left\{r_t^{(i)}, \ldots, r_t^{(N)}\right\}. \tag{12}$$

In order to consider the interactions among the vehicles in an area, we propose a selective edge-enhanced graph attention network (eGAT) to extract interaction features from the graph representation [56]. We denote by $\mathcal{U}_t$ the inter-vehicle interaction in global coordinate systems with the dynamics features ($\mathcal{R}_t$) of the vehicles to be embedded in the corresponding graph nodes, i.e.,

$$\mathcal{U}_t = \left\{u_t^{(0)}, u_t^{(1)}, ..., u_t^{(N)}\right\}, \tag{13}$$

where $u_t^{(i)}$ represents the interaction feature of agent $i$ at time $t$. eGAT models the interaction features of all the vehicles simultaneously, in the sense that:

$$\mathcal{U}_t = \text{eGAT}(\mathcal{R}_t, E_t), \tag{14}$$

where $E_t$ represents the set that contains the edge indexes and attributes at time $t$.

Having defined the encoding and interaction processes, we proceed with the trajectory prediction and the decoding process by jointly considering the vehicles' dynamics features $\mathcal{R}_t$, the interaction features $\mathcal{U}_t$, and the road infrastructure, represented by a map selector $m_t^{(i)}$. To overcome the multi-agent trajectory prediction constraints on local maps, we use $m_t^{(i)}$ as an adaptive map selector that is computed based on a shared map across all vehicles in correspondence with their current positions and velocities.

Considering the vehicle's speed and the need to cover the dynamics of all vehicles, our multi-agent prediction method will share the same map feature across all target vehicles ignoring the fact that different target agents are affected by different

parts of the map. For these reasons, we adopt the approach of [38] for selective map sharing based on a convolutional neural network map feature extraction, exploiting the LSTM gates of the decoder. Furthermore, we employ an LSTM as the decoder, in the following manner:

$$f_t^{(i)} = \mathrm{LSTM}\left(\left[r_t^{(i)}\right]\left[u_t^{(i)}\right]\left[m_t^{(i)}\right]\right), \tag{15}$$

with $[r_t^{(i)}][u_t^{(i)}][m_t^{(i)}]$ representing the concatenation of the corresponding features. The interaction representation may be approached with a GNN capable of handling the nodes, the directed edges, and the continuous edge attributes. By aggregating information from neighborhoods, several eGAT layers are employed with each layer updating node features, denoted by $h_i$, $i = 1, 2, \ldots, N$, i.e.,

$$\mathcal{H} = \{h_1, h_2, \ldots, h_N\}, \tag{16}$$

where $\mathcal{H}$ represents the set of all nodes features. A layer first transforms node and edge features accordingly, and then it aggregates neighboring node features with a multi-head attention mechanism. This combination allows the eGAT layers to capture spatial dependencies and the LSTM to model temporal dependencies. Considering each node's recent edge features and the corresponding edge attributes, eGAT gives updated node features based on the information received by its neighborhood.

Edge attributes represent the distances between the nodes of the graph. Regarding the edge-enhanced attention, a concatenated feature vector, $e_{ij}^{+}$, defined by

$$e_{ij}^{+} = [e_{ij}][h_j], \tag{17}$$

represents the edge feature of node $j$ from node $i$'s point of view, which is then forwarded to a shared attention mechanism. The attention mechanism $\mathbf{a} = (a_{ij})$ is a single-layer, feed-forward neural network with LeakyRelu, softmax linearization, and non-linearity attributes. Its attention coefficient, $a_{ij}$, indicates the importance of the node $j$ to node $i$, while jointly considering node and edge features. Traditionally, a GAT layer applies attention over the neighborhood of node $i$ utilizing only the structural information of the graph, while casting away the edges' features [52]. To this end, and based on [38], we include the edge features in the computation of the attention coefficients, namely,

$$a_{ij} = \frac{\exp\left(\mathrm{LeakyReLU}\left(\mathbf{a}^T[h_i][e_{ij}^{+}]\right)\right)}{\sum\limits_{k\in N_i}\exp\left(\mathrm{LeakyReLU}\left(\mathbf{a}^T[h_i][e_{ik}^{+}]\right)\right)}, \tag{18}$$

where $N_i$ is the neighborhood of vehicle $i$. In order to stabilize the self-attention mechanism, we implement independent attention mechanisms per node. Specifically, the multi-head attention mechanism is shaped by $k$ independent attention mechanisms with their features to be concatenated, resulting in the following output feature representation:

$$h_i^{new} = \sigma\left(\sum_{j\in N_i} a_{ij}^k W_h^k e_{ij}^{+}\right), \tag{19}$$

with $\sigma$, $W$ and $k$ representing the sigmoid function, the weighted sum of node features over its neighborhood, and the number of attention heads, respectively. For the minimization of the overlapping effects at the interaction among vehicles and their 360° video streaming capability, we further adapt the multi-head attention mechanism considering a minimum threshold in the distances between the neighboring vehicles, namely,

$$e_{ij} \geqslant \lambda\,\mathrm{LFL}, \tag{20}$$

where $\lambda$ denotes the variable that defines the minimum acceptable distance in proportion to the LFL of the 360° cameras.

### B. Bandwidth Allocation

The second phase of our proposed architecture allocates the actual bandwidth to the selected vehicles from the previous phase. The intermediate inputs at this phase are the predicted coordinates of the selected vehicles, the fixed surface they cover through 360° video streaming, and the Euclidean distances between the chosen vehicles that are given in the form of an adjacency matrix.

We employ a full graph neural network block containing a global block, a node block, and an edge block. Inspired by the work of [46], and adopting a set-to-set mapping approach, we define the resource allocation GNN, denoted by $\mathrm{GNN}_{RA}$, as:

$$\mathrm{GNN}_{RA} = \{J_{enc}^u, J_{enc}^e, J_{dec}^u\}, \tag{21}$$

where we adopted the following notation: $J_{enc}^u$: $\mathbb{R}^2 \rightarrow \mathbb{R}^{n_u}$ represents the encoding rates of the video covered surface, $J_{enc}^e$: $\mathbb{R}^2 \rightarrow \mathbb{R}^{n_e}$ the encoding rates of the distances between the vehicles, $J_{dec}^u$: $\mathbb{R}^{n_u} \rightarrow \mathbb{R}$ the decoding of the final state for the assignment of the bandwidth resource, and $(J^e, J^u, J^b)$: $\mathbb{R}^{n_u+n_e+n_b} \rightarrow \mathbb{R}^{n_e}$, $\mathbb{R}^{n_u}$ or $\mathbb{R}^{n_b}$ represent the updates of the graph's edges, nodes, and global values, respectively. Furthermore, $n_u$, $n_e$ and $n_b$ denote the hyperparameters that represent the size of the node encodings, the edges and the final resource assignments, respectively. All aggregation operations are implemented via sum-pooling. The notation involved in our architecture is presented in Table I.

In our approach, we consider a distance-oriented resource allocation with vehicles that are far distant from the rest receiving adequate bandwidth for video streaming. As a consequence, our neural network architecture, illustrated in Figure 5, provides an end-to-end solution for the trajectory-based predictive SA with optimal coverage-wise bandwidth allocation guarantees. Details of the eGAT architecture are provided in Algorithm 1.

## V. Experiments

### A. Simulation Set Up

For the investigation of the performance of our proposed framework, we trained and validated our neural network architecture through extensive simulations, under several urban vehicular scenarios, with the use of four datasets, namely the Interaction dataset [57], the Road Vehicle Localization dataset

**Algorithm 1** Enhanced Graph Attention Network (eGAT)

**Input:** Historical states $s_t^{(i)}$ for all vehicles, edge set $E_t$, map selector $m_t^{(i)}$, hyperparameters $\lambda$ and LFL
**Phase 1: Coverage Prediction**
**for** each vehicle $i$ **do**
  Compute dynamics feature $r_t^{(i)} = \text{LSTM}(s_t^{(i)})$
  Gather all dynamics features $\mathcal{R}_t = \{r_t^{(1)}, \ldots, r_t^{(N)}\}$
  Compute interaction features $\mathcal{U}_t = \text{eGAT}(\mathcal{R}_t, E_t)$
  Apply adaptive map selection $m_t^{(i)}$
  Compute trajectory features $\text{LSTM}([r_t^{(i)}][u_t^{(i)}], [m_t^{(i)}])$
**end for**
**for** each vehicle $i$ **do**
  Update vehicle features $h_i^{new}$ with multi-head attention mechanisms
  **for** each neighboring vehicle $j$ **do**
    Apply minimum distance constraint $e_{ij} \geqslant \lambda$LFL
    Compute concatenated edge feature $e_{ij}^{+} = [e_{ij}, h_j]$
    Calculate attention coefficient $a_{ij}$
    Update vehicle feature $h_i^{new} = \sigma(\sum_{j \in N_i} a_{ij} W_h e_{ij}^{+})$
  **end for**
**end for**
**Phase 2: Bandwidth Allocation**
Encode video coverage $J_{enc}^{u}$ and distance $J_{enc}^{e}$
Apply $\text{GNN}_{RA} = \{J_{enc}^{u}, J_{enc}^{e}, J_{dec}^{u}\}$ for bandwidth allocation
Set video qualities $\mathcal{Q} = \{q_1, q_2, \ldots, q_K\}$ based on vehicle and edge updates
**Output:** Predicted trajectories and bandwidth allocation for selected vehicles

[58], the Next Generation Simulation (NGSIM) dataset [59], and the Highway Traffic Data of the California Department of Transportation [60]. All datasets were pre-processed in order to be consistent as inputs in the GNN models. For all four datasets investigated, we utilized the layouts of the Interaction dataset, which includes various real-world vehicular scenarios with diverse traffic conditions. A typical intersection layout with the expected 360° vehicular video coverage is depicted in Figure 6.

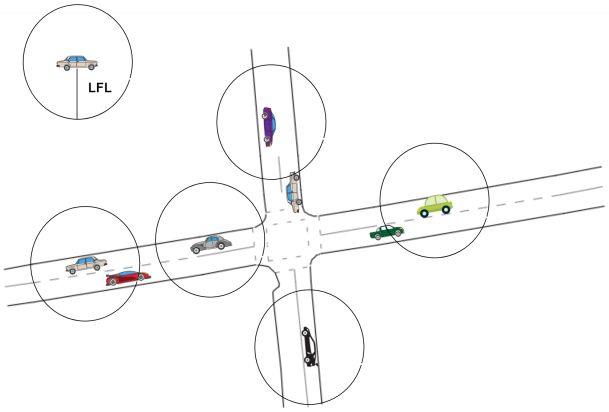


Fig. 6. Intersection layout: vehicles and their 360° video coverage.

We investigated an intersection, a roundabout and a map-free layout approach. More particularly, in order to evaluate the scalability of the proposed framework, we examined several simulation scenarios, where we varied the number of vehicles in the region. For smaller numbers of vehicles (up to 30), we performed our experiments on map layouts of the dataset, namely the DR_USA_INTERSECTION_GL for the intersection layout case, and the DR_USA_ROUNDABOUT_FT for the roundabout layout case, whereas for larger scales with more vehicles we considered a map-free approach. We implemented both the map-based and map-free experiments and, based on our results, we concluded that the performance of eGAT is map-independent. The set of video qualities is defined as $Q = \{6, 8, 10\}$ Mbps, with 10 Mbps representing the highest possible allocated video quality, 8 Mbps a medium video quality and 6 Mbps the lowest acceptable case. Considering critical situations, where the uplink capacity is limited while the number of vehicles in a certain region is large, we set the uplink capacity $B$ accordingly, in order to test the bandwidth allocation efficiency. The initial parameters of our architecture are presented in Table II.

The evaluation process was conducted on a DELL Inspiron 3847 with a processor Intel Dual Core i3 1600MHz, 8GB RAM. Regarding software configuration, we ran our experiments on a Windows 10 OS with Python 3.7.13 using the NumPy 1.23.0 library and PyTorch 2.0 framework. For

TABLE II
eGAT PARAMETERS

| | | |
|---|---|---|
| Coverage | Input size | 2 and 720 |
| | Input embedding size | 64 |
| | Encoder size | 128 |
| | Decoder size | 64 |
| | Output length | Number of Vehicles |
| | In channels node | 1 |
| | In channels edge attributes | 2 |
| | Output channels | 2 |
| | Heads | 2 |
| | Concatenation | True |
| | Edge attribute size | 2 |
| | Node size | 2 |
| | Distance threshold ($\lambda$ = 1 , LFL = 30) | 30 meters |
| Bandwidth Allocation | Number of input features for node encoding | 2 |
| | Number of output features for decoding | 2 |
| | Decoding on nodes | True |
| | Number of message-passing steps | 5 |
| | Layer Normalization | False |
| | Number of vertices | Number of Vehicles |
| | Number of edges | 45 |
| | Dimension of the input features | 3 |
| | Number of hidden layers | 100 |
| | Use of the edge model | False |

the evaluation of the performance of our architecture, we employed three metrics, namely the overall percentage of the covered region, the normalized mutual information (NMI) score and the precision-recall. The overall percentage of the covered region, $E_P$, is calculated by:

$$E_p = \frac{E}{E_R} \times 100, \tag{22}$$

where $E_R$ represents the total area surface, and $E$ represents the covered region, namely:

$$E = \sum_{m=1}^{M} E_m - \sum_{\xi=1}^{M} \sum_{\eta=\xi+1}^{M} \Theta_{\xi\eta}, \tag{23}$$

where $E_m$ denotes the virtual circular surfaces of the active DTBs, $\mathcal{L}_\alpha$, and $\Theta_{\xi\eta}$ denotes the overlapping area between $E_\xi$ and $E_\eta$. The NMI metric, utilized for the comparison of model performance, is defined by

$$NMI(\mathcal{L}_\alpha, \widehat{\mathcal{L}}_\alpha) = \frac{2I(\mathcal{L}_\alpha, \widehat{\mathcal{L}}_\alpha)}{H(\mathcal{L}_\alpha) + H(\widehat{\mathcal{L}}_\alpha)}, \tag{24}$$

where $\mathcal{L}_\alpha$ represents the set of active DTBs in a region $\mathcal{R}_\alpha$, $\widehat{\mathcal{L}}_\alpha$ denotes the ground truth set of selected DTBs in a region $\mathcal{R}_\alpha$, $I(\mathcal{L}_\alpha, \widehat{\mathcal{L}}_\alpha)$ represents the mutual information between $\mathcal{L}_\alpha$ and $\widehat{\mathcal{L}}_\alpha$, and $H$ denotes entropy [61]. NMI values range between 0 and 1; values closer to 1 demonstrate higher model performance.

In order to test the sensitivity of our model to other GPS coordinate datasets and their corresponding vehicle speeds per timestep, we applied the precision metric, i.e., the proportion of correct positive predictions made by the model, namely:

$$\text{Precision} = \frac{TP}{TP + FP}, \tag{25}$$

where $TP$ and $FP$ denote true and false positives, respectively. Furthermore, we employed the recall metric, also referred to as sensitivity or true positive rate. Recall quantifies the proportion of actual positive cases correctly identified by the model and is calculated by:

$$\text{Recall} = \frac{TP}{TP + FN}, \tag{26}$$

where $FN$ denotes false negatives.

### B. Predictive Video Coverage

To demonstrate the coverage performance and its corresponding metrics, we consider different vehicular densities, in both small and large scales. In Table III, we present our video coverage results under specific bandwidth values and selected vehicle numbers. The algorithm automatically determines the number of vehicles selected.

TABLE III
eGAT AREA COVERAGE ($E_p$) FOR PREDICTION HORIZON $t = 5$ S (SMALL SCALE)

| Bandwidth (Mbps) | Vehicles | Selected Vehicles | Intersection Coverage (%) | Roundabout Coverage (%) |
|---|---|---|---|---|
| 50 | 10 | 5 | 28.2 | 26.7 |
| 120 | 20 | 7 | 63.2 | 65.6 |
| 200 | 30 | 14 | 81.6 | 80.9 |

Table IV demonstrates the scalability of our approach in terms of area coverage, as the number of vehicles increases considerably in a map-free. The two selected map layouts are presented in Figure 7, with blue dots representing the vehicles in a certain indicative prediction horizon, and their surrounding circles representing the 360° coverage of the chosen vehicles.

TABLE IV
eGAT AREA COVERAGE ($E_p$) FOR PREDICTION HORIZON $t = 5$ S (LARGE SCALE, MAP-FREE)

| Bandwidth (Mbps) | Vehicles | Selected Vehicles | Coverage(%) |
|---|---|---|---|
| 250 | 50 | 23 | 62.3 |
| 500 | 100 | 47 | 73.6 |
| 900 | 300 | 109 | 81.2 |
| 1500 | 500 | 179 | 87.7 |
| 3000 | 800 | 198 | 92.8 |
| 5000 | 1200 | 210 | 93.3 |
| 7000 | 2000 | 234 | 95.1 |
| 9000 | 3000 | 241 | 97.6 |

### C. Predictive Bandwidth Allocation

Following the video coverage results, and aiming to allocate bandwidth as effectively as possible, we consider the same scenarios as in the previous subsection. In order to test the feasibility of our architecture in critical situations, where the bandwidth values are limited and disproportionate to the number of the selected vehicles, we computed the allocated bandwidth in certain video qualities. Bandwidth allocation and the corresponding video qualities per vehicle are presented in Table V. Furthermore, by testing scenarios with a larger number of vehicles and with several bandwidth values in the global map, we were able to demonstrate the scalability of eGAT, see Table VI.

TABLE V
eGAT BANDWIDTH ALLOCATION AND VIDEO QUALITIES FOR PREDICTION HORIZON $t = 5$ S (SMALL SCALE)

| Intersection ($n$=10 Vehicles) | | | | | |
|---|---|---|---|---|---|
| Bandwidth | Selected | Streaming | Vehicles per Video Quality | | |
| (Mbps) | Vehicles | Vehicles | High | Medium | Low |
| 50 | 5 | 5 | 5 | 0 | 0 |
| 40 | 5 | 5 | 1 | 3 | 1 |
| 30 | 5 | 3 | 1 | 2 | 0 |
| **Roundabout ($n$=10 Vehicles)** | | | | | |
| Bandwidth | Selected | Streaming | Vehicles per Video Quality | | |
| (Mbps) | Vehicles | Vehicles | High | Medium | Low |
| 50 | 5 | 5 | 5 | 0 | 0 |
| 40 | 5 | 5 | 2 | 1 | 2 |
| 30 | 5 | 3 | 2 | 1 | 0 |

TABLE VI
eGAT BANDWIDTH ALLOCATION AND VIDEO QUALITIES FOR PREDICTION HORIZON $t = 5$ S (LARGE SCALE, MAP-FREE)

| Bandwidth (Mbps) | Vehicles | Streaming Vehicles | Vehicles per Video Quality | | | Residual Bandwidth |
|---|---|---|---|---|---|---|
| | | | High | Medium | Low | |
| 250 | 50 | 23 | 9 | 8 | 6 | 40 |
| 500 | 100 | 47 | 22 | 17 | 8 | 54 |
| 900 | 300 | 109 | 37 | 44 | 28 | 10 |
| 1500 | 500 | 179 | 72 | 60 | 47 | 18 |
| 3000 | 800 | 198 | 150 | 28 | 28 | 1108 |
| 5000 | 1200 | 210 | 158 | 30 | 22 | 3048 |
| 7000 | 2000 | 234 | 131 | 83 | 20 | 4906 |
| 9000 | 3000 | 241 | 169 | 26 | 46 | 6826 |

### D. Performance Comparisons

The performance of our architecture is compared against similar studies, presented in [62] and [63], considering a pre-

diction horizon of $t = 5$ seconds. The results of the coverage performance are shown in Table VII. In [62], the authors focus on deploying sensors to enhance connectivity among robots navigating within a designated area. In our experiments, the cameras are used in a manner similar to the one the authors of [62] used sensors attached to robots for optimal area coverage. Furthermore, the authors of [63] proposed an improved flower pollination algorithm to guarantee the coverage and connectivity requirements in a specific region. We suitably adapted their experimental setup for our study to enable performance comparisons. Specifically, we assumed that the experiment's target points correspond to the centers of our areas, considering the sensor nodes as vehicles in our scenarios. The algorithm's output, representing the optimal node locations, is then adjoining with the nearest randomly positioned vehicles in our scenario to guide the selection for video coverage. Our approach allocates bandwidth based on distance, ensuring distant vehicles receive a sufficient amount of resources for video streaming. To the best of our knowledge, no other studies in the recent literature have applied similar methods to the problem of bandwidth allocation.

TABLE VII
COVERAGE PERCENTAGE ($E_P$) COMPARISONS

| Streaming Vehicles | Coverage (%) | | |
|---|---|---|---|
| | eGAT | [62] | [63] |
| 5 | **28.2** | 27.8 | 25.6 |
| 23 | **67.2** | 65.3 | 66.8 |
| 47 | **73.6** | 67.4 | 66.3 |
| 109 | **81.2** | 67.8 | 67.9 |
| 179 | **87.7** | 68.1 | 70.0 |
| 198 | **92.8** | 70.1 | 72.2 |
| 210 | **93.3** | 71.9 | 72.8 |
| 234 | **95.1** | 75.3 | 79.2 |
| 241 | **97.6** | 87.4 | 87.5 |

Finally, to demonstrate the combined effectiveness of the GNN and GAT concepts in our architecture, we conducted experiments in terms of coverage and bandwidth allocation, as compared to other state-of-the-art multi-agent attention-based trajectory algorithms, namely SCALE-Net [37], HEAT [38], and AI-TN [39]. In these algorithms, the attention mechanism is adjusted according to the distance threshold, set by the parameter $\lambda$ in equation (20), to yield comparable results for selecting video streaming vehicles within the specified coverage prediction horizons. The NMI score results are presented in Tables VIII and IX. The calculation of the NMI score is based on the comparison between the expected ground truth and the actual selection of video streaming vehicles in certain prediction horizons for optimal video coverage. Considering the other datasets, we observe that the precision and recall values maintain the same trend as the coverage prediction horizon increases, with minor fluctuations among the datasets for each model case, see Table X.

### *E. Discussion*

Integrating 360° video streaming with GNNs offers significant benefits over traditional approaches by providing comprehensive scene understanding through full-surround views and improved contextual awareness. GNNs' ability to capture complex dependencies enables more accurate predictive analytics and modeling of both the spatial and temporal aspects of 360° video, surpassing methods that treat these dimensions separately. This allows for efficient data processing and scalability, ensuring robust performance even with large datasets, thus facilitating the creation of dynamic digital twins for road environments.

TABLE VIII
NMI SCORE COMPARISONS FOR $n$= 100 VEHICLES (LARGE SCALE)

| State-of-the-art Models | Coverage Prediction Horizons | | | | | |
|---|---|---|---|---|---|---|
| | 1 sec | 5 secs | 7 secs | 10 secs | 15 secs | 20 secs |
| eGAT | **0.83** | **0.72** | **0.73** | 0.68 | 0.61 | 0.57 |
| SCALE-Net [37] | 0.64 | 0.64 | 0.70 | 0.61 | 0.58 | 0.51 |
| HEAT [38] | 0.79 | 0.70 | 0.72 | 0.70 | **0.64** | 0.62 |
| AI-TN [39] | 0.76 | 0.69 | 0.70 | **0.71** | 0.63 | **0.63** |

TABLE IX
$e$GAT BANDWIDTH ALLOCATION: EFFECT OF SCALABILITY IN CRITICAL SITUATIONS FOR PREDICTION HORIZON T = 5S

| Streaming Vehicles ($n$) | 5 | 23 | 47 | 109 | 179 | 198 |
|---|---|---|---|---|---|---|
| Bandwidth Values | 40 | 200 | 400 | 800 | 1500 | 1700 |
| NMI Score | 0.73 | 0.63 | 0.67 | 0.58 | 0.56 | 0.47 |

The results of the implementation of our architecture demonstrate the feasibility and high-level performance of predictive situational awareness in smart cities. The eGAT architecture shows flexibility in providing area coverage for different scales and under various bandwidth and vehicle configurations. Furthermore, it efficiently allocates bandwidth, adjusting the number of the chosen streaming vehicles to maintain video quality across different scenarios, as indicated in Table VI. The architecture demonstrates scalability, as it performs effectively with an increasing number of vehicles and bandwidth, achieving high area coverage and maintaining video quality. Table VII emphasizes the efficacy of the eGAT architecture in achieving high area coverage compared to certain state-of-the-art architectures. In particular, in all 9 scenarios investigated, eGAT consistently outperformed its counterparts achieving higher coverage percentages. The increase in coverage percentage ranged from $0.5\%$, for $n = 23$ streaming vehicles, to $32.3\%$, for $n = 198$ streaming vehicles, thus indicating the superiority of eGAT over the ones employed in [62] and [63].

Furthermore, our NMI score results indicate that our method provides comparable, or, in cases with low values of coverage prediction horizons, superior performance over several state-of-the-art graph attention methods, namely [37], [38], and [39], as indicated in Table VIII. In particular, our proposed method outperformed its counterparts for $T_h$ = 1, 5, and 7 seconds, with an increase in the NMI score ranging from $2.8\%$, for $T_h$ = 5 seconds, to $29.6\%$, for $T_h$ = 1 second. This indicates the superiority of eGAT against the other attention-based GNN models in low prediction horizons. For larger values of prediction horizons, namely, $T_h$ = 10, 15, and 20 seconds, eGAT presents lower, however comparable, coverage performance.

Our results suggest that eGAT is a promising solution

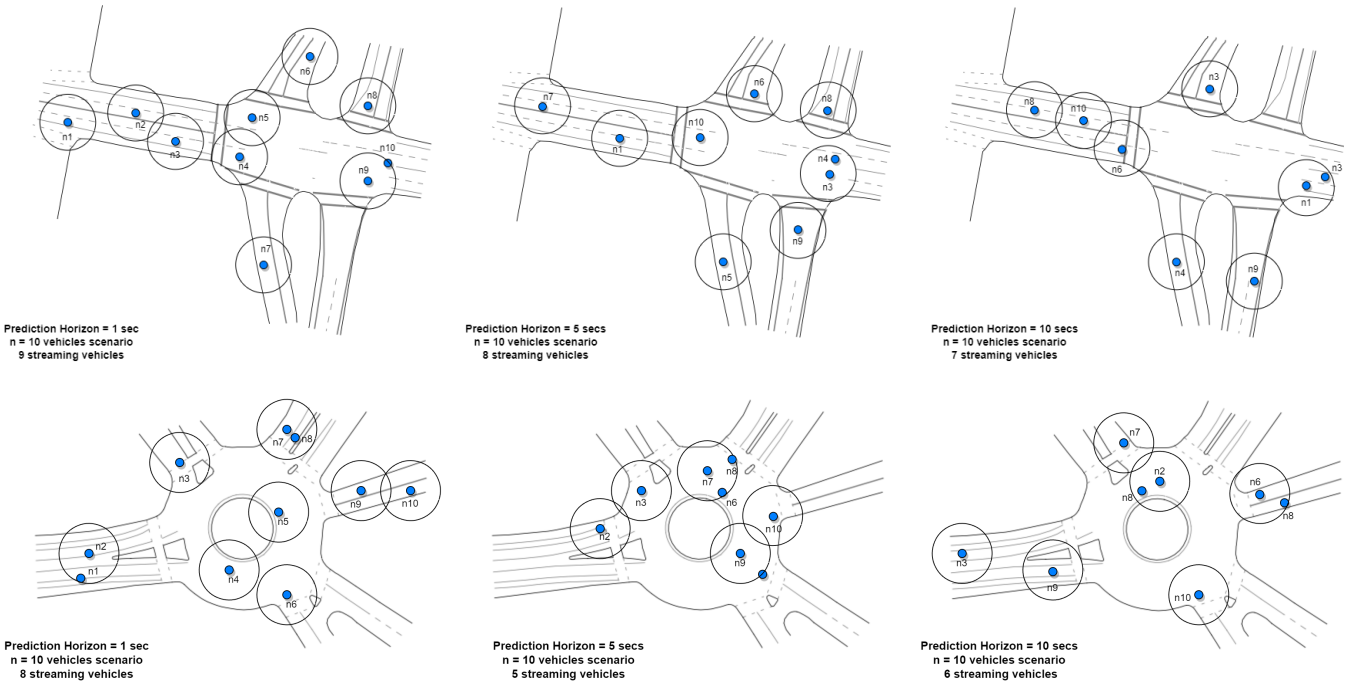


Fig. 7. Visualisation of eGAT predicted coverage on two separate driving scenarios for three prediction horizons, namely, $T_h$ = 1, 5 and 10 seconds, provided by the Interaction dataset [57]. First row: predictive area coverage via the DR_USA_Intersection_GL. Second row: predictive area coverage via the DR_USA_Roundabout_FT. Dark blue dots represent the vehicles' current positions, with their IDs indicated by $n_1$ to $n_{10}$, whereas surrounding circles indicate streaming vehicles.

TABLE X
PRECISION (P) AND RECALL (R) VALUES IN FOUR COVERAGE PREDICTION HORIZONS USING FOUR DIFFERENT DATASETS

| Models | Datasets | 1 sec | | 5 secs | | 10 secs | | 20 secs | |
|---|---|---|---|---|---|---|---|---|---|
| | | P | R | P | R | P | R | P | R |
| eGAT | [57] | 0.931 | 0.903 | 0.892 | 0.874 | 0.871 | 0.847 | 0.853 | 0.828 |
| | [58] | 0.915 | 0.894 | 0.881 | 0.861 | 0.862 | 0.835 | 0.841 | 0.812 |
| | [59] | 0.926 | 0.913 | 0.904 | 0.886 | 0.884 | 0.859 | 0.867 | 0.837 |
| | [60] | 0.944 | 0.928 | 0.915 | 0.897 | 0.896 | 0.873 | 0.879 | 0.849 |
| SCALE-Net | [57] | 0.921 | 0.905 | 0.890 | 0.878 | 0.876 | 0.849 | 0.858 | 0.833 |
| | [58] | 0.918 | 0.892 | 0.882 | 0.869 | 0.865 | 0.839 | 0.845 | 0.824 |
| | [59] | 0.933 | 0.919 | 0.907 | 0.889 | 0.887 | 0.857 | 0.871 | 0.846 |
| | [60] | 0.945 | 0.929 | 0.914 | 0.894 | 0.899 | 0.869 | 0.882 | 0.854 |
| HEAT | [57] | 0.929 | 0.906 | 0.892 | 0.876 | 0.871 | 0.843 | 0.855 | 0.834 |
| | [58] | 0.915 | 0.892 | 0.879 | 0.864 | 0.858 | 0.834 | 0.844 | 0.819 |
| | [59] | 0.924 | 0.914 | 0.899 | 0.884 | 0.878 | 0.855 | 0.863 | 0.848 |
| | [60] | 0.937 | 0.922 | 0.909 | 0.891 | 0.885 | 0.866 | 0.872 | 0.857 |
| AI-TN | [57] | 0.924 | 0.908 | 0.889 | 0.873 | 0.869 | 0.841 | 0.849 | 0.825 |
| | [58] | 0.912 | 0.896 | 0.881 | 0.865 | 0.861 | 0.832 | 0.836 | 0.815 |
| | [59] | 0.929 | 0.917 | 0.902 | 0.882 | 0.880 | 0.855 | 0.860 | 0.838 |
| | [60] | 0.941 | 0.927 | 0.911 | 0.893 | 0.891 | 0.862 | 0.873 | 0.847 |

for scenarios involving streaming vehicles, showcasing its potential for applications such as autonomous vehicles or traffic management systems. However, a critical limitation of our study that requires further explanation is the residual bandwidth allocation per area. As indicated in Table VI, in scenarios with high bandwidth availability, there appears an unexpected inconsistency between the actual and the expected allocated video qualities with the selected vehicles to stream at lower qualities. Our contribution differs in that our GNN architecture allocates bandwidth based on distance, focusing on resource efficiency while meeting coverage requirements. The latter highlights the effectiveness of the algorithm, where bandwidth availability is disproportionate to the number of vehicles. Especially in critical situations, vehicle selection may be compromised due to their imbalanced distribution in the area of interest. To this end, the effect of scalability on bandwidth allocation is presented in Table IX.

The primary limitations of our study include reliance on simulation data, which may not capture all real-world complexities, and the use of fixed video quality levels, that could be further optimized. Furthermore, our approach requires significant computational resources, potentially limiting deployment in resource-constrained environments. Future work should focus on integrating additional data sources and improving system robustness in diverse urban scenarios. Additionally, data privacy and security must be ensured by implementing end-to-end encryption and anonymizing video streams. To address ethical considerations, real-time video processing techniques should be used to blur faces, license plates, and other identifiable features before any data is stored or further processed, or to mask sensitive areas to prevent the identification of individuals. Integrating the eGAT architecture with existing infrastructure for real-time 360° video streaming involves implementing edge processing to perform local anonymization and dynamically manage bandwidth, thereby anonymizing sensitive data at the source.

In computational complexity terms, the time complexity of multiple ($k$) GAT attention heads computing $W$ features may be expressed as $k \cdot O(|M|WW_{in} + |M|W)$, where $W_{in}$, $|N|$ and $|M|$ denote the number of input features, of nodes, and

of edges in the graph, respectively [52]. Implementing $k$-head attention increases the storage and parameter requirements by a factor of $k$, which, in turn, may increase the response time of our proposed scheme. However, employing certain hardware acceleration techniques and model optimization algorithms is expected to reduce response delays. Our architecture demonstrates computational efficiency through several key aspects. GATs leverage multi-head attention mechanisms to efficiently aggregate information from neighboring vehicles, focusing on the most relevant parts of the graph and reducing computational overhead compared to traditional convolutional networks. By learning efficient node embeddings, we can quickly determine the optimal bandwidth allocation for each vehicle, achieving near real-time performance. Designed for high scalability, our system is able to accommodate a growing number of vehicles and increasing data demands in dense urban areas. Our extensive simulations highlighted the scalability of eGAT architecture under varying vehicular densities.

A promising direction for future research involves the weighted selection of vehicles in a specific area, tailored to meet certain QoE requirements. These requirements originate from the spatial distribution of the vehicles and the necessity for high-quality video streaming at specific locations, where only one vehicle is capable of streaming. Further research is required in order to understand the trade-off between achieving targeted coverage and ensuring sufficient bandwidth allocation. This trade-off impacts the vehicle's QoE and the efficient coverage of neighboring areas sharing the same uplink capacity. Integrating fairness-aware resource allocation strategies [64] and distributed resource allocation over weighted balanced graphs [65] represent two promising GNN-based approaches that may yield insightful results. Furthermore, future research could focus on enhancing data collection by integrating multimodal data sources such as LIDAR and on employing crowdsourced data from connected vehicles and smartphones. Additionally, developing advanced GNN architectures tailored for 360° video data, including spatio-temporal GNNs and employing transfer learning, could improve efficiency and generalization.

## VI. Conclusion

In this work, we present a novel GNN architecture integrating, for the first time, 360° video streaming into the digital twins of smart cities. Our application-oriented architecture, eGAT, makes a distinct contribution by optimally balancing area coverage with guaranteed bandwidth allocation, thus aligning with predictive situational awareness. By integrating real-time 360° video data with advanced GNNs, our system develops a comprehensive predictive model of road conditions and traffic patterns. This enhances the accuracy and reliability of situational awareness, enabling proactive decision-making for traffic management and road safety. The experimental results underscore the effectiveness of our approach in coverage performance and in bandwidth allocation across various critical scenarios. Furthermore, our results demonstrate the capability of eGAT to predict the area coverage in smart cities in several scenarios. The innovative nature of our method enhances real-time informational capabilities, addressing key coverage and resource allocation challenges in 360° video streaming for future vehicle locations. Future work will focus on expanding the dataset to include diverse weather conditions, integrating additional data sources like sensor networks, and deploying the system in real-world scenarios for validation. Furthermore, we plan to further improve our simulation models to better capture real-world complexities and validate our approach through collaborations with industry partners.

## References


1 Sehar, N. u., Khalid, O., Khan, I. A., Rehman, F., Fayyaz, M. A., Ansari, A. R., and Nawaz, R., "Blockchain enabled data security in vehicular networks," *Scientific Reports*, vol. 13, no. 1, p. 4412, 2023.

2 Aslam, A. M., Chaudhary, R., Bhardwaj, A., Budhiraja, I., Kumar, N., and Zeadally, S., "Metaverse for 6g and beyond: the next revolution and deployment challenges," *IEEE Internet of Things Magazine*, vol. 6, no. 1, pp. 32–39, 2023.

3 Akilan, T., Wu, Q. M. J., and Zhang, W., "Video foreground extraction using multi-view receptive field and encoder–decoder dcnn for traffic and surveillance applications," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 9478–9493, 2019.

4 Chen, L., Teng, S., Li, B., Na, X., Li, Y., Li, Z., Wang, J., Cao, D., Zheng, N., and Wang, F.-Y., "Milestones in autonomous driving and intelligent vehicles—part ii: Perception and planning," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 10, pp. 6401–6415, 2023.

5 Agudo, I., Montenegro-Gómez, M., and Lopez, J., "A blockchain approach for decentralized v2x (d-v2x)," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 4001–4010, 2021.

6 Li, X., Cao, H., Zhao, S., Li, J., Zhang, L., and Raj, B., "Panoramic video salient object detection with ambisonic audio guidance," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, pp. 1424–1432, 06 2023.

7 Utku, D. H., Catak, F. O., Kuzlu, M., Sarp, S., Jovanovic, V., Cali, U., and Zohrabi, N., "Digital twin applications for smart and connected cities," in *Digital Twin Driven Intelligent Systems and Emerging Metaverse*. Springer, 2023, pp. 141–154.

8 Choi, C.-S., "Modeling and analysis of priority-based distributed access control in vehicular networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 6, pp. 6848–6852, 2022.

9 Hui, Y., Zhao, G., Yin, Z., Cheng, N., and Luan, T. H., "Digital twin enabled multi-task federated learning in heterogeneous vehicular networks," in *2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, 2022, pp. 1–5.

10 Marai, O. E., Taleb, T., and Song, J., "Roads infrastructure digital twin: A step toward smarter cities realization," *IEEE Network*, vol. 35, no. 2, pp. 136–143, 2021.

11 Rahmani, A. M., Ali, S., Malik, M. H., Yousefpoor, E., Yousefpoor, M. S., Mousavi, A., Khan, F., and Hosseinzadeh, M., "An energy-aware and q-learning-based area coverage for oil pipeline monitoring systems using sensors and internet of things," *Scientific Reports*, vol. 12, no. 1, p. 9638, 2022.

12 Wu, J., Sun, Y., Li, D., Shi, J., Li, X., Gao, L., Yu, L., Han, G., and Wu, J., "An adaptive conversion speed q-learning algorithm for search and rescue uav path planning in unknown environments," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 12, pp. 15 391–15 404, 2023.

13 Nath, A., Rather, Z., Mitra, I., and L., S., "Multi-criteria approach for identification and ranking of key interventions for seamless adoption of electric vehicle charging infrastructure," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 7, pp. 8697–8708, 2023.

14 Burhanuddin, L. A. b., Liu, X., Deng, Y., Challita, U., and Zahemszky, A., "Qoe optimization for live video streaming in uav-to-uav communications via deep reinforcement learning," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 5, pp. 5358–5370, 2022.

15 Katz, M., Matinmikko-Blue, M., and Latva-Aho, M., "6genesis flagship program: Building the bridges towards 6g-enabled wireless smart society and ecosystem," in *2018 IEEE 10th Latin-American Conference on Communications (LATINCOM)*. IEEE, 2018, pp. 1–9.

16 Ye, X., Li, M., Si, P., Yang, R., Wang, Z., and Zhang, Y., "Collaborative and intelligent resource optimization for computing and caching in iov with blockchain and mec using a3c approach," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 2, pp. 1449–1463, 2023.

17 Tilahun, F. D., Abebe, A. T., and Kang, C. G., "Multi-agent reinforcement learning for distributed resource allocation in cell-free massive mimo-enabled mobile edge computing network," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 12, pp. 16 454–16 468, 2023.

18 Hong, G., Yang, B., Su, W., Li, H., Huang, Z., and Taleb, T., "Joint content update and transmission resource allocation for energy-efficient edge caching of high definition map," *IEEE Transactions on Vehicular Technology*, pp. 1–13, 2023.

19 Yang, C., Liang, P., Fu, L., Cui, G., Huang, F., Teng, F., and Bangash, Y. A., "Using 5g in smart cities: A systematic mapping study," *Intelligent Systems with Applications*, vol. 14, p. 200065, 2022.

20 Wang, J., Zhang, X., He, X., and Sun, Y., "Bandwidth allocation and trajectory control in uav-assisted iov edge computing using multiagent reinforcement learning," *IEEE Transactions on Reliability*, vol. 72, no. 2, pp. 599–608, 2023.

21 Wu, W., Dong, J., Sun, Y., and Yu, F. R., "Heterogeneous markov decision process model for joint resource allocation and task scheduling in network slicing enabled internet of vehicles," *IEEE Wireless Communications Letters*, vol. 11, no. 6, pp. 1118–1122, 2022.

22 Ma, T., Chen, X., Ma, Z., and Jiao, L., "Deep reinforcement learning based dynamic content placement and bandwidth allocation in internet of vehicles," in *International Conference on Wireless Algorithms, Systems, and Applications*. Springer, 2021, pp. 244–253.

23 Jiang, X., Wu, H., Jiang, H., Du, X., and Fang, J., "Co-hcca: Bandwidth allocation strategy in internet of vehicles with dynamically segmented congestion control," *Journal of Communications and Information Networks*, vol. 6, no. 2, pp. 175–183, 2021.

24 Kumar, K. and Kumar, N., "Region coverage-aware path planning for unmanned aerial vehicles: A systematic review," *Physical Communication*, vol. 59, p. 102073, 2023.

25 Yi, L., Wan, A. Y. S., Le, A. V., Hayat, A. A., Tang, Q., and Mohan, R. E., "Complete coverage path planning for reconfigurable omni-directional mobile robots with varying width using gbnn(n)," *Expert Systems with Applications*, vol. 228, p. 120349, 2023.

26 Guerna, A., Bitam, S., and Calafate, C. T., "Roadside unit deployment in internet of vehicles systems: A survey," *Sensors*, vol. 22, no. 9, 2022.

27 Asudeh, A., Berger-Wolf, T., DasGupta, B., and Sidiropoulos, A., "Maximizing coverage while ensuring fairness: A tale of conflicting objectives," *Algorithmica*, vol. 85, no. 5, pp. 1287–1331, 2023.

28 Zhou, Y., "Learning-driven optimization approaches for combinatorial search problems," Ph.D. dissertation, Angers, 2017.

29 Zhu, K., Han, B., and Zhang, T., "Multi-uav distributed collaborative coverage for target search using heuristic strategy," *Guidance, Navigation and Control*, vol. 1, no. 01, p. 2150002, 2021.

30 Mukhamediev, R. I., Yakunin, K., Aubakirov, M., Assanov, I., Kuchin, Y., Symagulov, A., Levashenko, V., Zaitseva, E., Sokolov, D., and Amirgaliyev, Y., "Coverage path planning optimization of heterogeneous uavs group for precision agriculture," *IEEE Access*, vol. 11, pp. 5789–5803, 2023.

31 Sadek, M. G., El-Garhy, A. M., and Mohamed, A. E., "A dynamic cooperative multi-agent online coverage path planning algorithm," in *2021 16th International Conference on Computer Engineering and Systems (ICCES)*, 2021, pp. 1–9.

32 Agarwal, S. and Akella, S., "Area coverage with multiple capacity-constrained robots," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 3734–3741, 2022.

33 Xu, C., Li, M., Ni, Z., Zhang, Y., and Chen, S., "Groupnet: Multiscale hypergraph neural networks for trajectory prediction with relational reasoning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022, pp. 6498–6507.

34 Schmidt, J., Jordan, J., Gritschneder, F., and Dietmayer, K., "Crat-pred: Vehicle trajectory prediction with crystal graph convolutional neural networks and multi-head self-attention," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 7799–7805.

35 Sheng, Z., Xu, Y., Xue, S., and Li, D., "Graph-based spatial-temporal convolutional network for vehicle trajectory prediction in autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 10, pp. 17 654–17 665, 2022.

36 Su, Y., Du, J., Li, Y., Li, X., Liang, R., Hua, Z., and Zhou, J., "Trajectory forecasting based on prior-aware directed graph convolutional neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 16 773–16 785, 2022.

37 Jeon, H., Choi, J., and Kum, D., "Scale-net: Scalable vehicle trajectory prediction network under random number of interacting vehicles via edge-enhanced graph convolutional neural network," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE Press, 2020, p. 2095–2102. [Online]. Available: 10.1109/IROS45743.2020.9341288

38 Mo, X., Huang, Z., Xing, Y., and Lv, C., "Multi-agent trajectory prediction with heterogeneous edge-enhanced graph attention network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 9554–9567, 2022.

39 Zhang, K., Zhao, L., Dong, C., Wu, L., and Zheng, L., "Ai-tp: Attention-based interaction-aware trajectory prediction for autonomous driving," *IEEE Transactions on Intelligent Vehicles*, vol. 8, no. 1, pp. 73–83, 2023.

40 Cheng, H., Liu, M., Chen, L., Broszio, H., Sester, M., and Yang, M. Y., "Gatraj: A graph- and attention-based multi-agent trajectory prediction model," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 205, pp. 163–175, 2023.

41 Wang, Z., Eisen, M., and Ribeiro, A., "Learning decentralized wireless resource allocations with graph neural networks," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1850–1863, 2022.

42 He, Z., Wang, L., Ye, H., Li, G. Y., and Juang, B.-H. F., "Resource allocation based on graph neural networks in vehicular communications," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*. IEEE, 2020, pp. 1–5.

43 Wang, D., Bai, Y., Huang, G., Song, B., and Yu, F. R., "Cache-aided mec for iot: Resource allocation using deep graph reinforcement learning," *IEEE Internet of Things Journal*, vol. 10, no. 13, pp. 11 486–11 496, 2023.

44 Huang, H., Lin, Y., Gui, G., Gacanin, H., Sari, H., and Adachi, F., "Regularization strategy aided robust unsupervised learning for wireless resource allocation," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 7, pp. 9647–9652, 2023.

45 Gu, Y., She, C., Quan, Z., Qiu, C., and Xu, X., "Graph neural networks for distributed power allocation in wireless networks: Aggregation over-the-air," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2023.

46 Cranmer, M., Melchior, P., and Nord, B., "Unsupervised resource allocation with graph neural networks," in *NeurIPS 2020 Workshop on Preregistration in Machine Learning*. PMLR, 2021, pp. 272–284.

47 Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M., "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, 2020.

48 Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Philip, S. Y., "A comprehensive survey on graph neural networks," *IEEE transactions on neural networks and learning systems*, vol. 32, no. 1, pp. 4–24, 2020.

49 Chen, L., Li, J., Peng, J., Xie, T., Cao, Z., Xu, K., He, X., Zheng, Z., and Wu, B., "A survey of adversarial learning on graphs," *arXiv preprint arXiv:2003.05730*, 2020.

50 Hamilton, W. L., Ying, R., and Leskovec, J., "Inductive representation learning on large graphs," *Proceedings of the 31st International Conference on Neural Information Processing Systems*, p. 1025–1035, 2017.

51 Kim, B.-H. and Ye, J. C., "Understanding graph isomorphism network for rs-fmri functional connectivity analysis," *Frontiers in neuroscience*, p. 630, 2020.

52 Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., and Bengio, Y., "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

53 He, K., Zhang, X., Ren, S., and Sun, J., "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1026–1034.

54 Protonotarios, N. E., Fokas, A. S., Kastis, G. A., and Dikaios, N., "Sigmoid and beyond: Algebraic activation functions for artificial neural networks based on solutions of a riccati equation," *IT Professional*, vol. 24, no. 5, pp. 30–36, 2022.

55 Khan, A. and Sarfaraz, A., "Rnn-lstm-gru based language transformation," *Soft Computing*, vol. 23, no. 24, pp. 13 007–13 024, 2019.

56 Jeon, H., Choi, J., and Kum, D., "Scale-net: Scalable vehicle trajectory prediction network under random number of interacting vehicles via edge-enhanced graph convolutional neural network," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 2095–2102.

57 Zhan, W., Sun, L., Wang, D., Shi, H., Clausse, A., Naumann, M., Kümmerle, J., Königshof, H., Stiller, C., de La Fortelle, A., and Tomizuka, M., "INTERACTION Dataset: An INTERnational, Adversarial and Cooperative moTION Dataset in Interactive Driving Scenarios with Semantic Maps," *arXiv:1910.03088 [cs, eess]*, 2019.

58 B. Cruz, S. and Aguiar, A., "Road vehicle localization dataset," 2020. [Online]. Available: https://dx.doi.org/10.21227/92ya-v240

59 Ammourah, R., Beigi, P., Fan, B., Hamdar, S. H., Hourdos, J., Hsiao, C.-C., James, R., Khajeh-Hosseini, M., Mahmassani, H. S., Monzer, D. *et al.*, "Introduction to the third generation simulation dataset: Data

collection and trajectory extraction," *Transportation Research Record*, p. 03611981241257257, 2024.

60 California Department of Transportation (Caltrans), "Performance measurement system (pems)," 2023, accessed: 2024-08-19. [Online]. Available: http://pems.dot.ca.gov/

61 Estevez, P. A., Tesmer, M., Perez, C. A., and Zurada, J. M., "Normalized mutual information feature selection," *IEEE Transactions on Neural Networks*, vol. 20, no. 2, pp. 189–201, 2009.

62 Tirandazi, P., Rahiminasab, A., and Ebadi, M., "An efficient coverage and connectivity algorithm based on mobile robots for wireless sensor networks," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–23, 2022.

63 Jiao, W., Tang, R., and Xu, Y., "A coverage optimization algorithm for the wireless sensor network with random deployment by using an improved flower pollination algorithm," *Forests*, vol. 13, no. 10, 2022.

64 Agarwal, B., Togou, M. A., Ruffini, M., and Muntean, G.-M., "A fairness-driven resource allocation scheme based on a weighted interference graph in hetnets," in *2021 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2021, pp. 1–6.

65 Liang, S., Zeng, X., and Hong, Y., "Distributed sub-optimal resource allocation over weight-balanced graph via singular perturbation," *Automatica*, vol. 95, pp. 222–228, 2018.