Deep Neural Networks Meet Computation Offloading in Mobile Edge Networks: Applications, Taxonomy, and Open Issues

Ehzaz Mustafa[1*], Junaid Shuja[2], Faisal Rehman[1], Ahsan Riaz[1], Mohammed Maray[3], Muhammad Bilal[4] and Muhammad Khurram Khan[5]

[1*]Department of Computer Science, Comsats University Islamabad, Abbottabad, Campus, 22060, Pakistan.
[2]Department of Computer and Information Sciences, Universiti Teknologi PETRONAS, Seri Iskandar,Malaysia.
[3]Department of Information Systems, King Khalid University,Saudi Arabia.
[4]School of Computing and Communications, Lancaster University,United Kingdom.
[5]King Saud University, Saudi Arabia.

*Corresponding author(s). E-mail(s): aizazmustafa@cuiatd.edu.pk;

Abstract

Mobile Edge Computing (MEC) is a modern paradigm that involves moving computing and storage resources closer to the network edge, reducing latency, and enabling innovative, delay-sensitive applications. Within MEC, computation offloading refers to the process of transferring computationally intensive tasks or processes from mobile devices to edge servers, optimizing the performance of mobile applications. Traditional numerical optimization methods for computation offloading often necessitate numerous iterations to attain optimal solutions. In this paper, we provide a tutorial on how Deep Neural Networks (DNNs) resolve the challenges of computation offloading. The article explores various applications of DNNs in computation offloading, encompassing channel estimation, caching, AR and VR applications, resource allocation, mode selection, unmanned aerial vehicles (UAVs), and vehicle management. We present a comprehensive taxonomy that categorizes these applications, and offer an overview of existing schemes, comparing their effectiveness. Additionally, we outline the open research issues that can be addressed through the application of DNNs in MEC offloading. We also highlight specific challenges related to DNN utilization in computation offloading. In conclusion, we affirm that DNNs are widely acknowledged as invaluable tools for optimizing computation offloading in MEC.

Keywords: Mobile edge computing, Computation Offloading, Deep Neural Networks, Supervised Learning, Unsupervised Learning, Reinforcement Learning

1 Introduction

Mobile Edge Computing (MEC) has emerged as an efficient paradigm, offering enough storage and computational resources at the network edge [1]. These network edges encompass a spectrum of devices, including eNBs, routers, switches, and other edge infrastructure. However, it is worth noting that end devices often face constraints in terms of computational capacity and storage [2]. Edge servers, strategically deployed at the network edge, play a pivotal role in bridging these resource disparities while providing diverse capabilities, such as computational power, network resources, and storage facilities with minimum latency. Beyond reducing latency, MEC processes data locally to provide network efficiency by freeing up bandwidth on congested cores. Additionally, MEC enhances scalability by distributing processing power and potentially reduces costs by offloading tasks from the cloud. This translates to a better user experience, especially in areas with limited coverage, by enabling faster response times and reliable connectivity for applications like AR/VR and real-time gaming.

One of the significant aspects of edge computing is computation offloading, a practice that unlocks the potential for innovative applications while minimizing transmission latency [4]. This is particularly critical in scenarios where nodes within the edge network demand ultra-reliable communication with minimal latency. Consider, for instance, the imperative need for minimal latency in the operation of connected vehicles and Unmanned Aerial Vehicles (UAVs) within real-time environments [7]. Moreover, the impending surge in Internet of Things (IoT) devices within future-edge networks will bring forth a substantial influx of data, compelling the need for streamlined processing capabilities at the edge servers. This transformation of data promises to reshape the networking landscape and mandates a comprehensive reevaluation of edge network modeling, analysis, design, and optimization strategies [9].

In MEC, some computation offloading problems pose significant challenges due to their NP-hard nature. These challenges primarily include the dynamic nature of wireless channels, which influence the decision between local and remote computation. Prior efforts in the domain of computation offloading optimization have grappled with solving these NP-hard problems. However, there exists an avenue for enhancing these solutions through the integration of DNNs. DNNs have shown remarkable promise in elevating the performance of MEC-based offloading techniques when compared to conventional numerical methods [10].

For instance, the deployment of artificial intelligence and DNN techniques holds the potential to significantly enhance a multitude of applications within connected vehicles and transportation systems [7]. DNNs can be harnessed to detect obstacles and estimate distances in autonomous vehicles, thereby contributing to safer and more efficient transportation [12]. Additionally, DNNs find relevance in various other domains within MEC, including resource allocation, channel estimation, and mode selection. For instance, DNNs can be employed to determine optimal offloading actions, while Convolutional Neural Networks (CNNs) prove effective in channel estimation tasks [8]-[13]-[14].

To facilitate a more comprehensive grasp of the application of DNNs in computation offloading, we introduce a scenario featuring a fleet of self-governing delivery drones conducting operations in an urban setting. These drones are tasked with efficiently and autonomously transporting packages to their designated destinations. To achieve this, they rely on real-time decision-making regarding computation offloading. Traditional offloading approaches, such as simplistic rule-based strategies advocating for either complete local computation or offloading all tasks to the edge server, prove suboptimal in light of the dynamic and resource-constrained nature of drone operations.

Here, DNNs can play a vital role. In this scenario, a sophisticated DNN model takes the role of optimizing real-time computation offloading decisions for each individual drone. The DNN takes into account a multitude of critical factors, including the drone's current geographical coordinates, the status of its onboard battery, the availability of computational resources, prevailing network conditions, and the computational complexity inherent to the tasks at hand. This extensive array of input data undergoes intricate processing within the DNN's deep neural layers, enabling it to discern intricate patterns and correlations.

The DNN model, trained using a combination of supervised and reinforcement learning methods, makes decisions regarding whether to perform computing tasks locally or remotely. Consider a scenario where a drone operates in an area with robust network connectivity and abundant nearby computational resources; in such cases, the DNN may prioritize remote computation to optimize resource management and preserve the drone's battery life. Conversely, when the drone encounters network congestion or deals with tasks of lower computational complexity, the DNN may recommend local computation to minimize latency and reduce reliance on network resources.

This DNN-driven approach results in an adaptive and exceedingly efficient offloading strategy, whereby each drone optimizes its computational burden by dynamically balancing local processing with offloading tasks to the edge servers. Consequently, the drones are empowered to make real-time decisions, navigate intricate urban landscapes, and execute their delivery duties with minimal latency, ensuring timely and efficient package deliveries. Moreover, this approach maximizes the utilization of edge computing resources. This illustrative example underscores the pivotal role played by DNNs in transforming computation offloading practices within MEC, harnessing their ability to intelligently optimize decisions in response to a multitude of dynamic and interconnected variables.

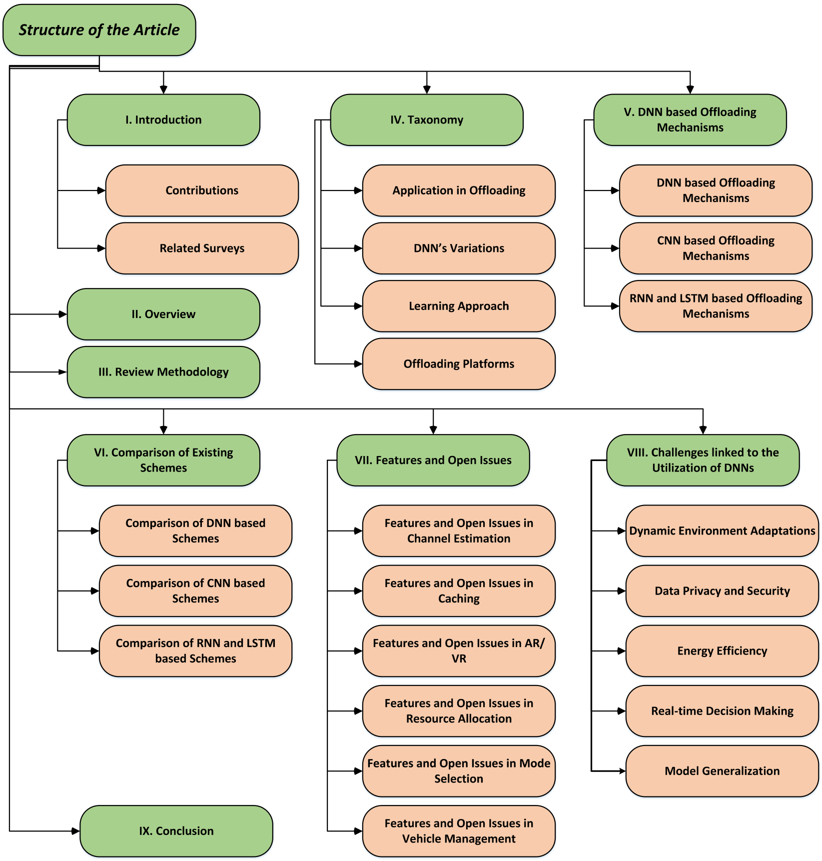


Fig. 1: Structure of the Article

1.1 Related Surveys

Authors of [15] reviewed the studies of machine learning-based offloading approaches including supervised, unsupervised, and reinforcement learning. The author provided a classical taxonomy and showed which studies improve the offloading process with different intelligent approaches. The authors of [16] reviewed the studies of stochastic-based offloading approaches and provided a classical taxonomy. Artificial neural networks based wireless networks were explored in [9]. In [5], the authors reviewed the recent literature on computation offloading empowered with wireless power transfer. A mobility-aware computation offloading survey is provided in [6]. In [167]. the authors provided a detailed review of optimization methods such as Convex, and Lyapunov etc., for computation offloading. In both [168] and [169], authors reviewed the current studies based on reinforcement learning for computation offloading in MEC.

Table 1: Significance of the Proposed Survey: In this context, AO represents Applications in Offloading, SL represents Supervised Learning, UL represents Unsupervised Learning, RL represents Reinforcement Learning, and GT represents Generic Taxonomy

Article	Focus	MEC	DNNs	AO	SL	UL	RL	GT
[16]	Offloading approaches based on Machine Learning	✓	✗	✗	✓	✓	✓	✗
[15]	Stochastic based offloading approaches	✓	✗	✗	✗	✗	✗	✗
[9]	DNNs based review in wireless networks	✗	✓	✓	✓	✓	✗	✗
[4]	Review of computation offloading in MEC/MCC	✓	✗	✗	✗	✗	✗	✗
[5]	Review of Joint Wireless power transfer and offloading	✓	✗	✗	✗	✗	✗	✗
[6]	Mobility aware computation offloading	✓	✗	✗	✗	✗	✗	✗
[165]	A review of optimization methods for computation offloading	✓	✗	✗	✗	✗	✗	✗
[166]	A survey of reinforcement learning for computation offloading	✓	✗	✗	✗	✗	✓	✗
[167]	Reinforcement learning methods for computation offloading	✗	✗	✗	✗	✗	✓	✗
[170]	Reinforcement learning methods for computation offloading	✗	✗	✗	✗	✗	✓	✗
[171]	Video data offloading in MEC	✓	✗	✗	✗	✗	✗	✗
This work	DNNs for computation offloading with applications in MEC	✓	✓	✓	✓	✓	✓	✓

To the best of the author's knowledge, none of the previous works have provided a comprehensive and systematic review within the MEC domain that adequately addresses the importance of employing DNNs for intelligent offloading. Unlike prior research efforts, our study encompasses DNNs for computation offloading, covering all major learning approaches, including supervised, unsupervised, and reinforcement learning aspects. Additionally, we present an inclusive and detailed taxonomy that takes into account numerous parameters related to DNN-based offloading, offering comprehensive discussions and referencing relevant sources. Table 1 provides the significance of the proposed survey.

1.2 Contributions

Following are the novel contributions of the proposed survey.

- We present a detailed taxonomy categorized into a) applications of DNNs in computation offloading, b) DNNs variations, c) Learning approaches, and d) Offloading platforms.
- We present various applications of DNNs that boost the process of computation offloading and improve the performance of MEC.
- We present an overview of DNNs-based computation offloading and provide a summary of existing works. We also provide a comparison of state-of-the-art techniques based on various parameters presented in taxonomy.
- We provide open research issues regarding the application of DNNs in MEC computation offloading.
- We also provide the challenges linked to the utilization of DNNs for computation offloading.

Figure 1 provides the overall structure of the article. Section 2 provides an overview of DNNs in computation offloading. Section 3 provides the review methodology. Section 4 provides the comprehensive taxonomy. A detailed overview of DNN-based computation offloading schemes is provided in Section 5. In Section 6, we provide a comparison of existing schemes that consider DNNs for computation offloading. Features and open issues are discussed in section 7. Section 8 presents the challenges linked to the utilization of DNNs for computation offloading. Finally, we conclude in section 9.

2 Overview of DNNs in Computation Offloading

Given their ability to handle complex tasks efficiently and adapt to dynamic environments, DNNs play a crucial role in enabling computation offloading. Figure 2 presents an overview of DNNs for computation offloading in MEC and cloud. Despite our focus which is on edge MEC, we provide a detailed representation of the use of DNNs in MEC and cloud layers. Here, we can see that the fundamental layers, characterized by their minimal computational demands or lightweight DNN configurations, are well-suited for execution on low-power terminal devices. In parallel, intermediate layers, which may encompass certain resource-intensive DNN components, find optimal execution on edge servers. Finally, the uppermost layers or complex DNN models, characterized by their computational intensity, are best suited for execution within the high-capacity cloud layer. This hierarchical distribution of layers ensures efficient utilization of computational resources across the spectrum from edge to cloud. For example, it is not only possible to dynamically identify the optimal partition point for DNN models but also to leverage complex network theory to determine the most efficient task offloading assignment, thus mitigating routing congestion when transmitting tasks in D2D multi-hop networks as presented in figure 3. Moreover, we delve into the significance, necessity, and nuances of applying DNNs to address key MEC applications, including channel estimation, caching, AR/VR, resource allocation, mode selection, UAV management, and vehicle management. Precise channel estimation is paramount in MEC to ensure reliable and efficient communication. DNNs are instrumental in this domain, as they possess the capacity to learn intricate patterns from data. By leveraging historical channel information and real-time data, DNNs enhance the accuracy of channel estimation. This is particularly crucial in scenarios where real-time, low-latency communication is imperative, such as in remote surgery or autonomous vehicle control. DNNs ensure that communication channels are optimized to meet stringent requirements. Similarly, DNNs play a pivotal role in caching by optimizing content caching strategies. They learn from usage patterns and adapt content placement decisions dynamically. This not only reduces latency but also alleviates network congestion. In applications like Augmented Reality (AR) and Virtual Reality (VR), where rapid data retrieval is essential for seamless user experiences, DNN-driven caching strategies are indispensable.

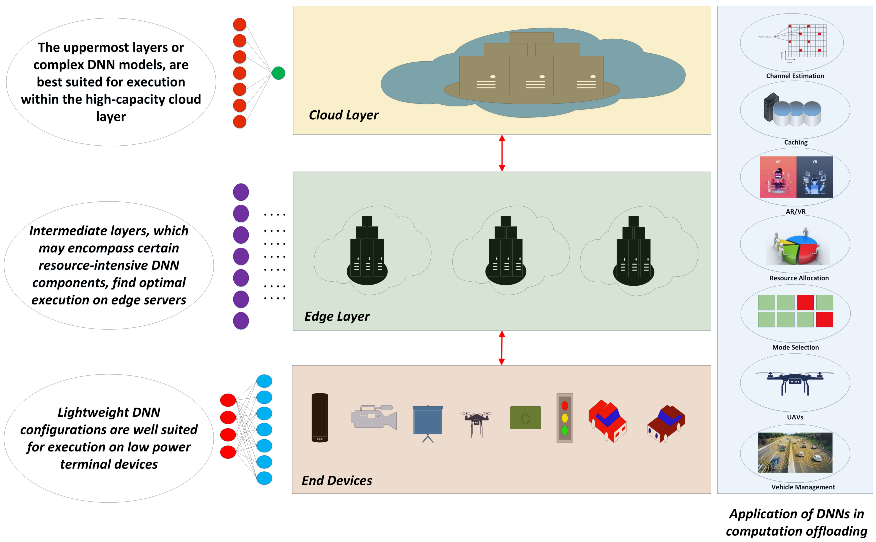


Fig. 2: Overview of DNN's application in Computation Offloading

In the context of resource allocation, efficient resource allocation in MEC involves judiciously assigning computational resources to tasks [25]-[26]. DNNs excel in this context by dynamically allocating resources based on task requirements, device capabilities, and network conditions. This optimization enhances the performance of applications like video streaming, gaming, and IoT analytics while ensuring the efficient utilization of edge resources. Moreover, mode selection, whether to perform computations locally or remotely, is a complex decision-making process in MEC. DNNs adapt to dynamic conditions by analyzing factors like task complexity, device capabilities, and network quality. They determine the optimal mode to align with the varying needs of applications. This adaptability is crucial in scenarios where responsiveness and energy efficiency are paramount, such as in healthcare applications or smart manufacturing.

UAV Management is another crucial application that holds immense potential in MEC, and DNNs are instrumental in optimizing their management. DNNs enable UAVs to navigate complex environments, avoid obstacles, and perform real-time data analysis. This enhances applications such as surveillance, inspection, and delivery operations, where precision, autonomy, and safety are critical. In the realm of connected vehicles, DNNs revolutionize decision-making processes. They optimize traffic routing, reduce congestion, and enhance autonomous driving capabilities. This not only ensures safer transportation but also reduces environmental impact by optimizing fuel consumption and emissions. DNN-driven vehicle management is fundamental for the realization of smart and efficient transportation systems.

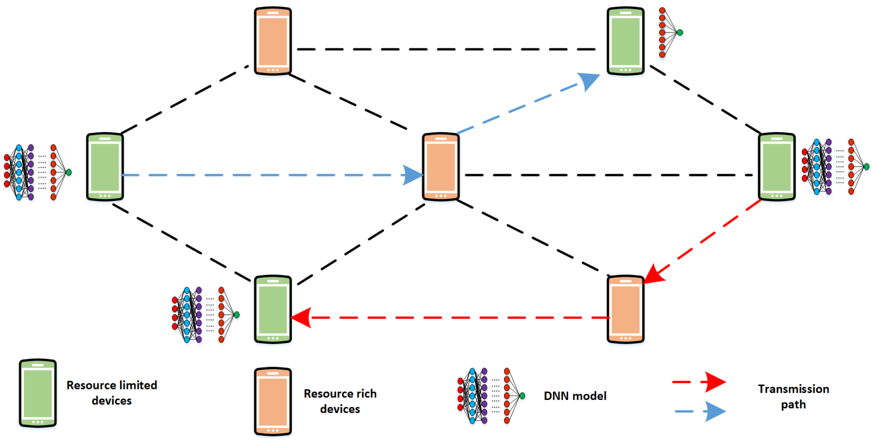


Fig. 3: DNN partitioning while task offloading in D2D networks

3 Review Methodology

The primary aim of this study is to conduct a comprehensive survey of the role of DNNs in computation offloading, encompassing supervised, unsupervised, and RL approaches rather than a systematic literature review (SLR). However, we use a systematic process to extract the relevant studies from databases and for the review. This section encompasses the formulation of research questions, the search process, the establishment of inclusion/exclusion criteria, and quality assessment. To begin the review process, it is crucial to identify Research Questions (RQ) as they define our motivation. This study is formulated by eight fundamental questions that serve as the basis for the literature review search strategy. The main research questions and the reasons for formulating them are outlined below:

RQ1: How do DNNs address and resolve challenges in computation offloading, surpassing the limitations of traditional numerical optimization methods? The expected outcome of this investigation is a comprehensive understanding of the specific challenges in computation offloading that DNNs effectively address. The findings can potentially highlight the advantages of DNNs over traditional numerical optimization methods, leading to insights that contribute to the optimization of computation offloading in MEC environments.

RQ2: How can DNNs enhance channel estimation accuracy in computation offloading? Answering this question help is the identification of optimal DNN architectures and learning approaches for precise channel estimation, potentially leading to advancements in reliable communication between devices and edge servers.

RQ3: What caching strategies leveraging DNN variations are most effective for minimizing data retrieval latency in computation offloading? Answering this question helps in evaluating DNN-driven caching mechanisms, providing insights into strategies that optimize content availability, reduce latency, and enhance overall system performance.

RQ4: How do DNNs contribute to the optimization of AR and VR applications in computation offloading? This question examines the impact of DNNs on improving the efficiency of AR and VR applications in MEC. Resource Allocation:

RQ5: In what ways can DNNs optimize resource allocation for computation offloading in MEC? This question helps in the identification of DNN-driven strategies for resource allocation, providing insights into efficient utilization of computing and storage resources at the network edge, and potential advancements in dynamic resource management.

RQ6: How can DNNs employing various learning approaches, facilitate optimal mode selection including partial or binary offloading in MEC? This question examines DNN-driven algorithms for mode selection, offering recommendations for adaptive strategies that dynamically adjust to varying computational demands, potentially improving overall system responsiveness.

RQ7: What role do DNNs play in improving computation offloading for UAVs and vehicle management systems, considering variations such as DNNs, RNNs, LSTM or CNNs? Answering this question helps in evaluating of DNN applications in UAVs and vehicle management, highlighting opportunities to enhance the efficiency and responsiveness of mobile applications in dynamic environments, potentially influencing autonomous systems.

RQ8: How do different DNN variations (DNN, CNN, RNN, LSTM) and learning approaches (supervised, unsupervised, RL) impact the overall effectiveness of computation offloading? Answering this question helps in comparative analysis of the performance of various DNN architectures and learning approaches, providing insights into their strengths and limitations in different application scenarios.

RQ9: How does the choice of offloading platform, such as edge servers, UAVs, and RSUs, impact the efficiency of DNN-driven computation offloading in MEC? Answering this question helps in the investigation of the influence of different offloading platforms on the performance of DNN-based computation offloading. The outcome will provide valuable insights into the optimal configurations and scenarios for each offloading platform, aiding in the informed selection and deployment of platforms in various use cases.

RQ10: What are the different applications and open issues in DNN based computation offloading? Providing answers to this question assists researchers in selecting a more suitable application, and crucial issues for their upcoming research.

RQ11: What are the different specific challenges linked to the utilization of DNNs for computation offloading? The question helps in highlighting the specific challenges associated with the utilization of DNNs for computation offloading. Table 2 provides the motivation, category and mapping sections of each RQ.

Search Criteria: This survey employed major scientific databases like Wiley Interscience, Springer, ACM Digital Library, ScienceDirect (or Elsevier), IEEE Xplore, and MDPI. The search terms encompassed phrases such as "Deep Neural Networks," "Artificial Neural Networks," "computation offloading," etc. The search strings were formulated by combining these keywords using Boolean "AND" and "OR" operators or "in" and "with" preposition and finally searched as follows:

```
("Deep neural networks" or DNNs in "computation offloading")
OR ("Artificial neural networks" or ANN in "computation
offloading") OR ("Convolutional neural networks" or CNNs in
"computation offloading") OR ("Recurrent neural networks" or
RNNs in "computation offloading") OR ("Long short term memory"
or LSTM in "computation offloading") OR ("Neural networks"
or NNs in "Channel Estimation") OR ("Neural networks" or
NNs in "Resource allocation in MEC") OR ("Neural networks"
or NNs in "AR and VR applications") OR ("Neural networks"
or NNs in "Computation offloading with UAVs and Vehicular
networks") OR ("Neural networks" or NNs in "caching with MEC")
OR ("Computation offloading" with "Supervised learning in MEC")
OR ("Computation offloading" with "Unsupervised learning in
```

Table 2: Motivation, category and mapping sections of each RQ

RQ	Motivation	Category	Mapping Section
RQ1	Understand how DNNs improve computation offloading compared to traditional methods.	DNNs impact on computation offloading	Section 2
RQ2	Identify optimal DNN architectures for precise channel estimation in offloading.	Channel Estimation	Section 4.1
RQ3	Evaluate effective DNN-driven caching strategies to minimize data retrieval latency.	Caching	Section 4.1
RQ4	Examine how DNNs enhance efficiency in AR and VR applications in offloading.	AR/VR applications	Section 4.1
RQ5	Investigate DNN-driven strategies for efficient resource allocation in offloading.	Resource Allocation	Section 4.1
RQ6	Explore DNN algorithms for adaptive mode selection in offloading scenarios.	Mode selection	Section 4.1
RQ7	Assess the role of DNNs in improving offloading for UAVs and vehicle management.	UAVs and Vehicle Management	Section 4.1
RQ8	Analyze the impact of DNN variations and learning approaches in computation offloading.	DNN variatios and learning approaches	Section 4.2, 4.3, 5 and 6
RQ9	Investigate how different offloading platforms impact the efficiency of DNN-driven offloading.	Offloading platforms	Section 4.4
RQ10	Explore diverse applications and identify open issues in DNN-based computation offloading.	Applications and open issues	Section 7
RQ11	Address the precise challenges inherent in deploying DNNs for computation offloading, facilitating targeted solutions and advancements in MEC	Challenges in utilizing DNNs	Section 8

`MEC") OR ("Computation offloading" with "Reinforcement learning in MEC").`

Quality Assessment and Criteria for Inclusion and Exclusion: Our review of the literature indicates a significant increase in research on computation offloading from 2018. Therefore, our main attention spans from 2018 to 2023. It's essential to note the exclusion of non-English papers and those not concurrently addressing DNNs and computation offloading together. Additionally, during the quality evaluation, conference papers lacking indexing or demonstrating low relevance were omitted. We also deleted the papers that were published without a rigorous peer review.

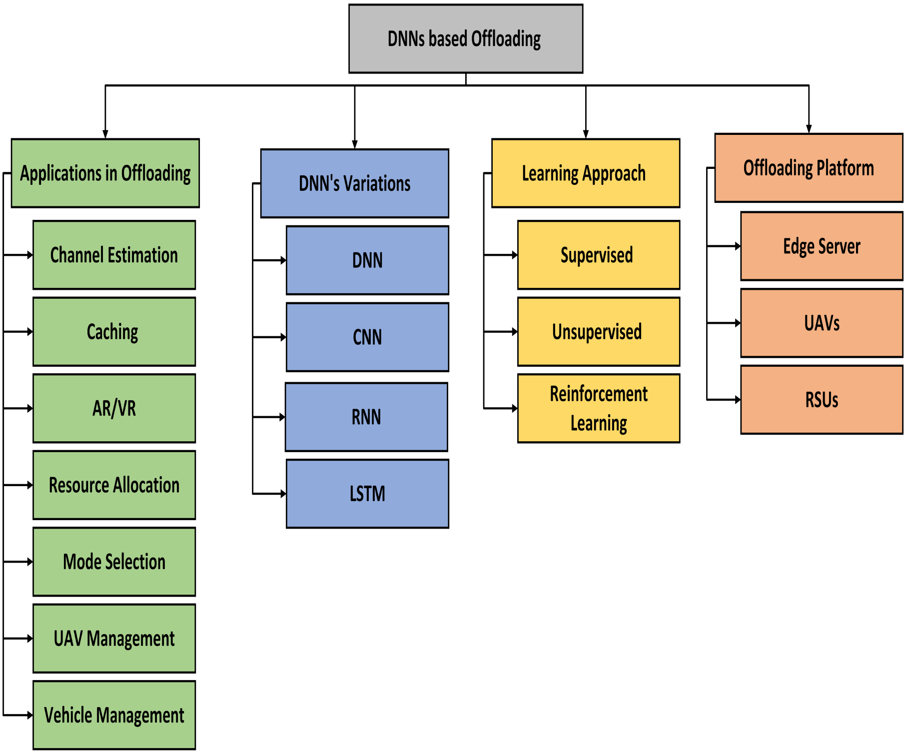


Fig. 4: Taxonomy of DNNs based computation offloading

Paper Distribution: According to our analysis about paper distribution, IEEE holds the majority at 68%, surpassing other publishers, with Springer and Elsevier following at 11%. According to the years, there has been a notable increase in DNN-based research after 2018. Our comprehensive literature review identified only journal and conference papers. Each selected article went through full-text reading and analysis to address the research questions.

4 Taxonomy

We present a generic taxonomy that covers the applications of DNNs with their variations in computation offloading. These applications include a) Channel estimation b) Caching c) Augmented Reality (AR) and Virtual Reality (VR) d) Resource Allocation e) Mode Selection f) UAVs and g) Vehicle Management. DNNs variations include as a) DNN b) CNN c) RNN d) Long Short term Memory (LSTM). We also discuss the learning approaches including a) supervised learning b) Unsupervised learning, and c) Reinforcement learning (RL) for computation offloading with related domain references. Finally, we provide the offloading platforms for compute-intensive tasks including a) Edge servers, b) UAVs and c) Roadside units (RSUs). Figure 4 shows the complete taxonomy.

4.1 DNN's applications in offloading

DNN serves as a key technology to make the computation offloading more intelligent. These DNN-based applications for offloading are discussed below.

4.1.1 Channel Estimation

Channel estimation provides information regarding the distortion in the transmitted signal throughout the propagation process. Prior channel estimation using some pilot values such as modulation symbols through various conventional schemes improves the performance of MEC. For example, one traditional scheme to estimate a channel is a training sequence (i.e. data-aided scenario). However, these traditional methods involve the estimation of interfering channel strength and then optimizing the scheduling based on different models [17]. These methods are resource and compute-intensive. Prior knowledge of wireless fading channels obtained with DNNs specifically CNNs through some known pilot values improves the overall procedure of computation offloading [18]. This intelligent offloading scheme has almost complete knowledge of channel statistics which is practical in a dynamic environment. Moreover, these approaches work better when the receiver in the network is equipped with chains of radio frequency [19]. Moreover, the learning process becomes more efficient by combining a fully convolutional network with learned approximate message-passing networks in massive MIMO systems [20]. The flexible de-noising convolutional network performs better without sacrificing means square error performance [21].

4.1.2 Caching

Edge devices provide popular content to the IoT devices in closer proximity to ensure minimum computational and transmission latency. However, due to the finite storage, and rapid growth in big data, it is hard to predict and provide popular content. Therefore, smart decisions are needed to serve IoT devices with popular content. DNNs-based content popularity effectively reduces the delays in accessing content [22]. RNNs and LSTM-based content popularity decisions are effective to decide which contents should be stored at the network edge [23]. Furthermore, distributed DNNs allow the network to exchange the particulars which reduces content demand errors without revealing privacy concerns [24]. These intelligent DNNs-based content prediction not only utilizes the bandwidth and minimizes latency, but also reduced the quantity of data transmitted [27].

4.1.3 Augmented Reality (AR)/ Virtual Reality (VR)

In the era of MEC, AR, and VR applications have been widely used because of the good experience for users. Besides, their ultralow latency demands and immense resource consumption brings a huge challenge. Machine learning tools provide powerful solutions for latency and resource demands [28]. However, these tools may be compute-intensive. To subsist the aforementioned challenges, DNNs are functional which makes it possible to decide offloading decisions, uplink and downlink transmissions, and resource allocations. These decisions improve the user experiences in terms of latency, resource efficiency, and energy consumption [29].

4.1.4 Resource Allocation

As the time, frequency, and other resources such as computing cycles are limited in the MEC, it is challenging to satisfy user demands in substantial network and mutual interference scenarios. These resource problems are often approached as linear and stochastic programming techniques. However, concrete allocation issues and MEC may not be expressed in the form of linear programming [30]. DNNs, specifically DNN and CNN-based resource allocation and management serve IoT devices with promising performance with many low computations [31]. In addition, video surveillance and face recognition require image recognition algorithms along with communication and computing resources, which can be effectively managed with DNNs with high accuracy of recognition in MEC [33].

4.1.5 Mode Selection

Optimal offloading mode selection among local, remote, and partial modes is challenging particularly in MEC and critically depends upon system parameters, such as the number of computing bits and distance from the access points. Furthermore, it is necessary to predict the effective mode selection for fast-fading channels using the values of channel gains. Conventional schemes such as the coordinate descent method [34] require several iterations, resulting in immense execution latency which may not be possible for dynamic and fading channels DNNs enable to learn the optimal offloading mode among local and remote computation [35], or partial computation [36] via experience with much low execution latency. Through DNNs it is also possible to predict the optimal offloading mode along with subcarrier allocation such as for non-orthogonal multiple access (NOMA) [37]. Different neural network tools may be used such as DNN [38] with reward scheme, CNN [39], and RNN along with LSTM [40] may be used for computing mode selection. This timely decision obtained with DNNs prediction is practical for fast-fading channels. Figure 5 shows the procedure used in [39] for computing mode selection applicable in dynamic networks and Figure 6 shows the procedure involved in [35].

4.1.6 Unmanned aerial vehicles (UAVs)

UAVs assisting edge computing are becoming popular for multiple applications including transmission control, offloading management, cooperation among end devices or edge servers, or autonomous operations supporting users on

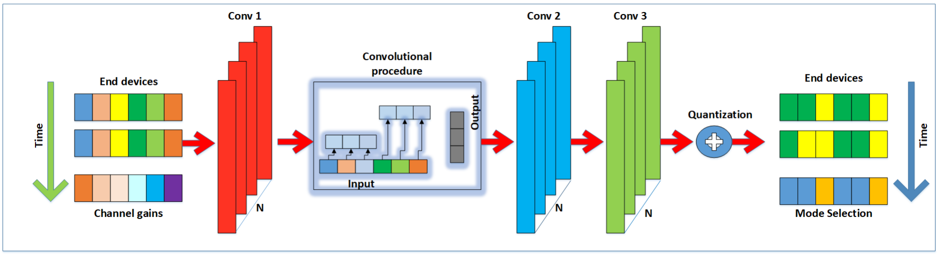


Fig. 5: Computing mode selection procedure with CNN

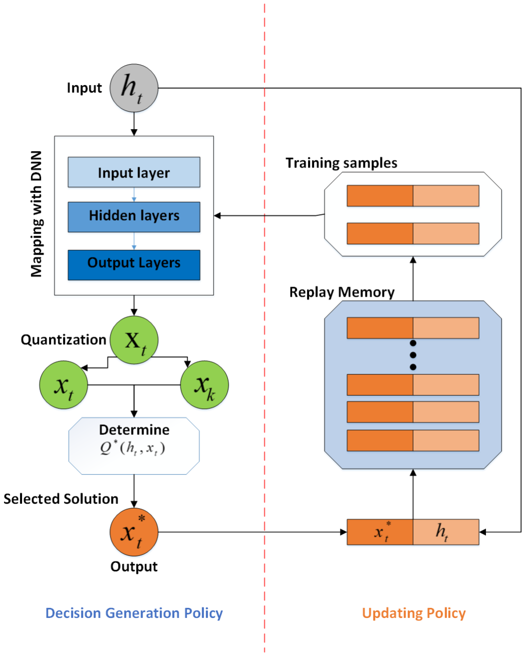


Fig. 6: Computing mode selection procedure with DNN and reward scheme

the ground in a dynamic environment [41]. These UAVs equipped with DNNs tools may have the ability to dynamically detect target sites and environments [42]. To support the user's devices through offloading at these UAVs, DNNs control the trajectories of these UAVs while ensuring user fairness and load [43]. With efficient trajectory management obtained with DNNs, these UAVs may act as edge servers while providing storage and computing resources to IoT devices [44]. Furthermore, DNNs-enabled UAVs specifically CNNs may effectively work for surveillance and monitoring systems [45].

4.1.7 Vehicle Management

The traditional techniques use mathematical programming for solving resource management and other problems for computation offloading in MEC. Nonetheless, edge networks become more complex with traditional programming while solving the above problems to a certain level of optimality such as high mobility in vehicular networks [46]. DNNs-based vehicular applications including resource management and vehicle mobility prediction are becoming popular because of the difficulty relies on traditional frameworks. Moreover, DNNs enable vehicle type prediction by capturing images and videos, offloading these inputs at the edge server intelligently, and extracting features and prediction, especially with CNN at the edge server [32]. DNNs also contribute to other offloading application such as load balancing among edge servers through intelligent offloading [47], environmental identification [48], scheduling and malicious nodes detection [49], etc.

4.2 DNN's Variations

Various DNNs have been used to make the offloading more intelligent and functional. These DNNs are a) Deep neural networks (DNNs), Convolutional neural networks (CNNs), Recurrent neural networks (RNNs), and Long short-term memory (LSTM). We discuss the application and usage of each of these for offloading scenarios in the following subsections. It is important to note that other neural network structures like spiking neural networks (SNNs) exist. However, based on the author's knowledge, their applications and usage in computation offloading are not yet significant.

4.2.1 Deep Neural Networks (DNNs)

These are widely used neural networks with multiple hidden layers between input and output layers. DNNs always consist of the same type of components such as weights, biases, and functions [50]. A DNN provides a high level of abstraction through multiple nonlinear transformations to learn multiple levels of representations. However, due to a large number of parameters, DNN suffers from severe overfitting. To cope with this issue, extensive studies proposed regularization approaches such as dataset augmentation, and weight decay [51].

4.2.2 Convolution Neural Networks (CNNs)

CNNs are effective tools for image understanding. These networks outperform human experts in many image understanding tasks [52]. Moreover, they are very effective in representing spatial patterns. It typically consists of convolution layers, max-pooling layers, and fully connected layers. CNNs are playing a major role in diverse functions such as computer visioning, image processing, and segmentation. An offloading perspective, vehicle management, channel estimation, smart parking, and mode selections are also improved by CNNs.

Table 3: Summary of the use of Learning Approaches for existing papers in specific DNNs applications in offloading

DNNs Applications in Offloading	Existing Papers	Learning Approach			
	Problem	Reference	Supervised	Unsupervised	RL
Channel Estimation	· Uplink channels estimation	· [93]	✓		
	· Quality of channels estimation	· [94]	✓		
	· Channel estimation over selective fading channels	· [95]			✓
	· Doppler spread and channel correlation	· [96]	✓		
	· Beam-space channels estimation	· [97]		✓	
	· Pilotless channel estimation	· [98]		✓	
Caching	· Edge caching in big data	· [99]	✓		
	· Mobility input-based caching	· [100]	✓		
	· Wireless coded caching	· [101]	✓		
	· Network representation learning based caching	· [102]		✓	
	· Proactive caching in 5g	· [103]		✓	
	· File popularity based content caching	· [104]			✓
	· Multi agent cooperative content caching	· [105]			✓
	· Game theory based content caching	· [106]			✓
AR/VR	· Aerial co-robot streaming	· [107]		✓	
	· Depth and motion estimation	· [109]		✓	
	· Binary offloading scheme for AR edge computing	· [89]			✓
	· RL based AR games management	· [110]			✓
Resource Allocation	· Fairness based resource allocation in wireless networks	· [111]	✓		
	· Graph neural network based D2D resource allocation	· [112]	✓		
	· Resource allocation in NOMA	· [113]			✓
	· Wireless power duration allocation with OFDMA	· [13]			✓
	· Wireless power duration allocation with FDMA	· [114]			✓
Mode Selection	· Computing mode selection and resource allocation	· [32]			✓
	· Computing mode selection and subcarrier allocation	· [34]			✓
	· Partial computing mode selection	· [33]			✓
	· Binary offloading mode selection	· [65]	✓		
	· Binary offloading mode selection	· [70]		✓	
UAVs	· Intelligent maximization of UAVs performance	· [115]	✓		
	· Autonomous 3D UAV localization	· [116]	✓		✓
	· Association of base station with UAVs	· [117]		✓	
	· Surveillance planning with UAVs	· [118]		✓	
	· Intrusion detection in UAVs	· [119]			✓
Vehicle Management	· Task offloading for vehicles	· [71]			✓
	· Energy efficient vehicle management	· [120]			✓
	· Spectrum access for cognitive vehicles	· [121]			✓
	· In vehicle alcohol detection	· [122]	✓		
	· Road anomaly segmentation for vehicles	· [123]		✓	

4.2.3 Recurrent Neural Networks (RNNs)

RNNs attracted considerable attention on sequential tasks such as speech recognition, handwriting recognition, and image-to-text. Compared to the general feed-forward networks, RNNs have recurrent loops called feedbacks. This feedback forms a backpropagation mechanism. These setups make RNN more powerful for sequential and time series data. However, this backpropagation causes gradient vanishing problems for RNNs [53]. CNNs usually work with convolutional and max pooling layers, whether RNN feeds the result back into the network. In offloading context, these networks may be used for handling task dependencies and learning offloading policies [54].

4.2.4 Long Short Term Memory (LSTM)

LSTM is the enhanced form of RNN. A typical RNN is limited to looking back in time for approximately ten timestamps. The reason behind this is the feedback signal is exploding are vanishing, which is commonly known as the vanishing gradient problem. To overcome this limitation, LSTM may be applied. LSTM can learn more than 10,000 timestamps [55]. These networks extend biologically plausible [56]. LSTMs perform better than RNNs because LSTM units include memory cells. These memory cells can maintain information for a long period. These networks are composed of forward layers, backward layers, and activation layers. Towards MEC, LSTM advances the offloading process such as in [57]-[58]-[59]-[60]-[61].

4.3 Learning Approach

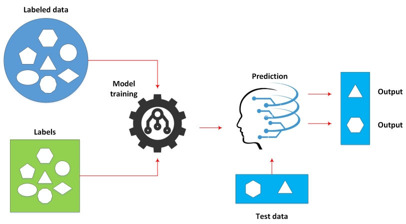
While learning of DNNs, three mechanisms are available, named as a) Supervised learning and b) Unsupervised learning and c) Reinforcement learning (RL). We discuss each of these in the following subsections.

4.3.1 Supervised Learning

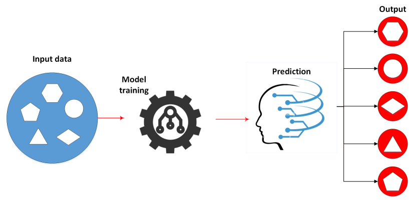
Supervised learning is a technique that relies on labeled data for training. It derives the implicit relationship between input data and predicted data. The training data contains labeled input such as feature vectors and desired output such as supervisory signals [62]. It takes the label data with trained labels before training and predicts the desired output. Supervised learning maps the labeled input data to the classified output according to the training dataset and supervisory signals [63]-[64]. This is the kind of learning where the output of the framework is known before the learning begins. This learning approach is widely used in computation offloading schemes [65]-[66]-[67]-[68].

4.3.2 Unsupervised Learning

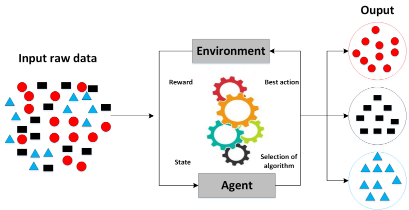
This is a learning approach where the aim is to detect hidden structures from unlabeled data. Here the desired output is known lately [69]. This approach has found much consideration in the fields of data compression, classification, outlier detection, dimensionality reduction algorithms, human learning, and so on [70]. The general technique in this learning is training from probabilistic data models. Moreover, structured patterns are learned in the data by rejecting pure unstructured noise. Classification of unsupervised learning into multiple subsets are Analytical hierarchy process, k-means, Hidden Markov model, clustering, Game-AI, and so on [71]. Computation offloading in MEC is significantly improved with unsupervised learning such as binary offloading [72], offloading in the internet of vehicles [73], and offloading in edge cloud



(a) Supervised Learning



(b) Unsupervised Learning



(c) Reinforcement Learning

Fig. 7: The process of (a) Supervised learning, (b) Unsupervised Learning and (c) Reinforcement Learning

platforms [74]. Besides, we also provide a summary of existing papers using these learning approaches with specific applications in Table 3.

4.3.3 Reinforcement Learning

Figure 7 provides the visual representation of supervised, unsupervised and RL. RL provides a mathematical formalism to control the learning mechanism. Through utilizing RL, one can automatically acquire near optimal behaviors [75]. In simple terms, RL is a learning approach based on rewarding an agent with desired behaviors and punishing undesired ones. The reward function describes what an agent should do, and RL defines how to do it. An RL agent can grasp and interpret its environment, take actions, and learn by trial and error using DNNs as presented in figure 8. It is recently witnessed major advances in solving decision-making problems in multiple domains such as vehicular networks [77], MEC [78], and wireless networks [79].

4.4 Offloading Platforms

A task may be offloaded to one of the following platforms in MEC. a) Edge server, b) UAVs, and c) RSUs. Each of these is discussed in the following subsections.

4.4.1 Edge Server

These are the offloading platforms that work as computing servers located in closer proximity to the end users such as those integrated with eNBs. Edge

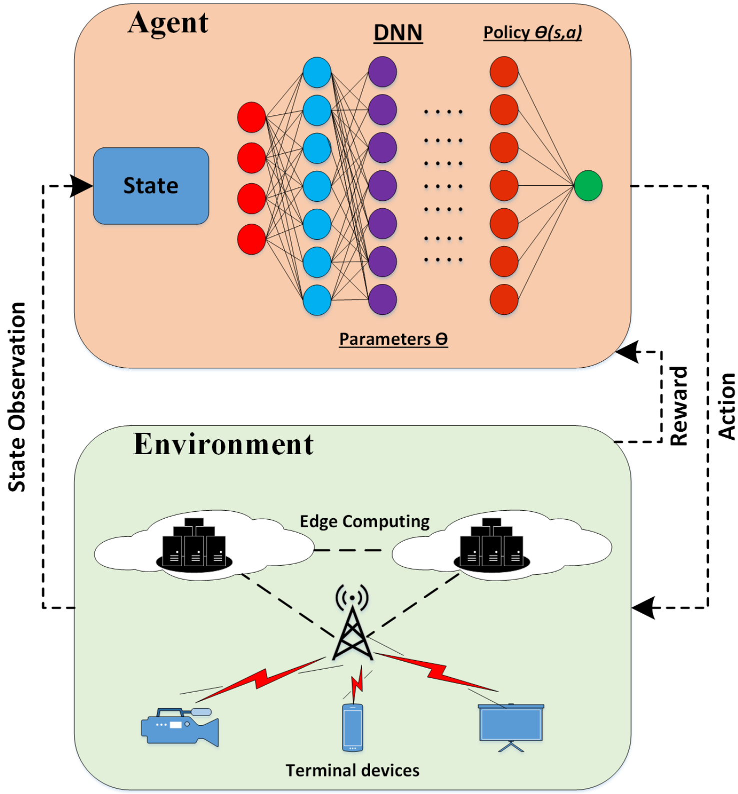


Fig. 8: Reinforcement Learning with DNNs

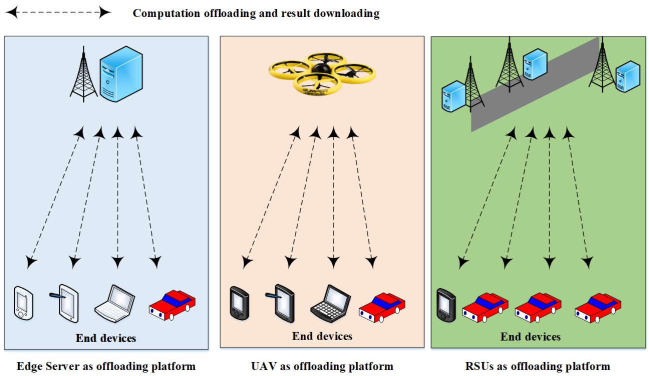


Fig. 9: Offloading Platforms

servers are resource-rich devices that provide the cloud services of storage and remote computation to the end devices. However, compared to cloud servers, edge server has limited resources to serve all the devices in the corresponding network. Hence, choosing a suitable edge server among multiple is a critical task [80]. Traditional methods require future knowledge such as server workload and user mobility which is not known a priori. To this end, DNNs play a crucial role. These networks work based on the observed information of past server selection. The LSTM may be exploited to encode the historical knowledge of dynamic factors such as server workloads and user mobility [81]. One can use a DNN to select the optimal server with resource allocation decisions [82]. Moreover, these DNNs enable the vehicles to act as servers for nearby end devices irrespective of their mobile nature [83].

4.4.2 UAVs

UAVs equipped with vision techniques could be leveraged as the offloading platform that provides pervasive computing capabilities to the end devices [84]. These end devices can offload computing tasks to the UAVs for computation. However, the dynamic nature of UAVs may degrade the offloading performance. Furthermore, the assumption of the availability of UAV dynamics such as mobility and aerodynamic friction is not always practical. To this end, DNNs can be exploited to boost the offloading process by learning nonlinear information such as aerodynamic friction and blade flipping [85]. Similarly, UAVs acting as servers can provide monitoring services and ubiquitous surveillance for agriculture management, military operations, and urban administration. An example can be the use of CNN for object detection at UAVs when end devices offload the sensed data [42]. Through this scheme, smart surveillance and traffic monitoring could be greatly improved. Similarly, a CNN may be exploited at the UAVs to identify vehicles on the ground when sensors and cameras offload input at these UAVs. Furthermore, this scheme also leverages the tracking of vehicles while using UAVs as edge servers [86].

4.4.3 Road Side Units (RSUs)

Road site units (RSUs) deployed in the internet of vehicle settings act as an offloading platform for the offloaded tasks. Vehicles offload compute-intensive tasks to these RSUs for computation. However, urban settings and the high mobility of vehicles have a great impact on computation offloading. More elaborately, migration of task offloading and their results to the users is challenging. For example, when RSU decides to transfer the results to the vehicle, and in the meantime, the vehicle moves to another location. Moreover, when a vehicle moves from the range of one RSU to another, collaboration among these two RSUs is also required for the optimal offloading process. These challenges require prior knowledge such as vehicle mobility and channel statistics etc., hence limiting the traditional schemes. To this end, DNNs can be exploited to predict the suitable RSUs for computation offloading in the case of dynamic topologies [87]. Similarly, a CNN may be deployed to determine the task offloading, computation, and results from migration policy for vehicles [88]. Figure 9 depicts the use of offloading platforms in MEC.

Table 4: Overview of existing papers based on the applications of DNNs in computation offloading

Offloading Applications	**Problem**	**Reference**	**DNN**	**CNN**	**RNN LSTM**
Channel Estimation	. Multi cell interference limited channel estimation	· [144]	✓		
	. Massive MIMO channel estimation	· [145]	✓		
	. Compressive channel estimation	· [146]		✓	
	. Channel estimation for cell free beam-space mmWave massive MIMO	· [147]	✓		
	. Channel estimation for cell free mmWave massive MIMO	· [21]		✓	
Caching	· Proactive caching with data-driven techniques	· [148]		✓	
	· Content caching in D2D networks	· [149]	✓		✓
	· Content popularity prediction in B5G networks	· [150]	✓		
	· Maximization of cache hit ratio	· [98]	✓		
	· Online proactive caching	· [151]			✓
AR/VR	· User-centric task assistance	· [152]		✓	
	· Interaction between front-end devices and backed helpers to improve video quality	· [153]		✓	
	· Real-time object detection in AR	· [154]		✓	
Resource Allocation	· Optimal topology for backhaul network	· [155]	✓		
	· Computational resource allocation	· [156]	✓		
	· Joint optimization of offloading and resource allocation	· [157]	✓		
Mode Selection	· Partial computing mode selection	· [13]	✓		
	· FDMA enabled partial computing mode selection	· [112]	✓		
	· OFDMA enabled partial computing mode selection	· [33]	✓		
	· Binary offloading mode selection	· [158]	✓		
	· OFDMA enabled binary computing mode selection	· [108]		✓	
UAVs	· Prediction of energy consumption for UAVs based MEC	· [159]			✓
	· Maximization of average secrecy rate for UAVs	· [160]	✓		
	· Task offloading strategy in UAVs	· [126]	✓		
	· Trajectory prediction of UAVs	· [161]			✓
	· Synthetic trajectory prediction of UAVs	· [162]			✓
Vehicle Management	· Data-driven intrusion detection system	· [163]	✓		
	· Channel prediction for connected vehicles	· [164]			✓
	· Content caching for vehicles	· [165]	✓		
	· Cooperative sensing among adjacent vehicles	· [166]			✓

5 DNNs based Offloading Mechanism in MEC

This section overviews the DNN-based offloading mechanisms. We provide the details of existing works that used DNNs applications to improve the offloading process. We also provide an overview of existing papers based on DNNs application in Table 4.

5.1 DNN based Offloading Mechanism

DNNs are the most common neural network structures used by researchers to utilize computation offloading due to their versatility and simplicity. In [37], Nduwayezu et al. investigated non-orthogonal multiple access (NOMA) with multi-carrier MEC system. The authors addressed the problem of sub-carrier allocation and decision among remote and local computation using DNNs. The objective of the study was to maximize the weighted sum computation rate under the binary offloading policy by assigning optimal subcarriers to the user equipment in case of offloading the task. The authors investigated that NOMA-based machine learning techniques do not require a complete definition of channel environment and labeled data for training purposes. The proposed algorithm also avoided the complexity of existing optimization algorithms for channel allocation.

Huang et al.,[35] provided a deep reinforcement learning-based online offloading (DROO) for wireless-powered MEC networks with binary computation offloading that maximizes the weighted sum computation rate. The algorithm uses reinforcement learning to learn from previous offloading experiences to improve the DNN-generated offloading action. To achieve quick algorithm convergence, the author developed an order-preserving quantization and an adaptive parameter setting technique. Unlike other optimization methods, the proposed DROO algorithm eliminates the requirement to solve complex mixed integer programming problems entirely. Simulation results demonstrate that DROO delivers near-optimal performance while decreasing CPU execution delay, making real-time system optimization practical for wireless-powered MEC networks in fading environments. Although the allocation of resources sub-problem is handled in the context of a specific wireless-powered network, the suggested DROO framework can be used to offload computation in universal MEC networks.

Tilahun et al., [89] investigated a cell-free multiple inputs multiple outputs (MIMO) mobile edge network to meet the demanding specifications of the recently introduced multimedia services. The authors provided a distributed deep reinforcement learning-based joint communication and resource distribution scheme. This scheme uses two DNNs including actor DNN and critic DNN. Each user is implemented as an autonomous agent such that joint resource allocation depends only on local observation. The simulation results show that the agents develop robust strategies that reduce consumption while fulfilling the extremely low latency requirements of advanced services.

The authors investigated MEC networks for intelligent IoTs, in which many users receive assistance from various computational access points for computational activities [90]. The performance of the system may be increased by lowering latency and energy expenditure. Two critical metrics of relevance in the MEC networks are considered, by offloading some of the operations. The authors designed the system by providing an offloading technique using RL technique. Deep Q-network was employed in this approach to automatically learn the offloading action to enhance system efficiency, and hence a DNN is trained for anticipating the offloading activities, with training data obtained from the environmental system. Furthermore, authors employ network bandwidth to enhance the wireless spectrum for links between users. Finally, simulation data are shown to demonstrate the usefulness of the suggested reinforcement learning offloading technique. The suggested DRL approach with DNN can greatly minimize the system cost of delay and power expenditure.

AR and IoT are emerging technologies that improve efficiency in many fields. To solve the problem of reliability and latency requirements of AR applications, the author studied a binary offloading scheme for AR edge computing which increases the computing capability of AR devices [91]. DNN-based AR offloading decision reduces the computational complexity, hence avoiding the need to solve combinatorial optimization. By using the Markov decision process which is solved by deep reinforcement learning, this scheme shows better performance compared with existing optimization methods.

Due to the rapid development of intelligently connected automobiles, more computing resource optimization strategies are required for network implementations. When practically all tasks are to be performed on MEC servers, many vehicle resources are unused, hence creating a burden on the servers [92]. To allocate efficient resources in distributed computation offloading, difficult work can be broken down into smaller subtasks. Deep Q Learning Network using DNN, which is a distributed compute offloading approach, is the primary way to reduce the amount of time that a complex job takes to complete.

Since vehicular networks are being used more and more, it is necessary to adopt a flexible design to improve the quality of service. Utilizing multi-access MEC is the optimal way to save time in such a network, even though MEC has limited resources and is unable to handle high mobility [124]. A resource allocation problem should be investigated with a wide range of applications to address the response time issue. DNN is a suitable model that rapidly provides a solution for resource allocation while learning the dynamics of network state.

In [125], the authors provided the offloading scheme to determine the QoS on the edges for the internet of vehicles. Consideration is given to the important factors of cache areas, processing power, and channel conditions to progress the fulfillment of quality of experience constrained by energy consumption. The author developed an updated DRL method using two DNNs to search for an optimum offloading mode due to the high complexity of the predicted offloading. The suggested framework can increase instability and speed up training by exchanging stochastic gradient descent with the experience replay buffer and employing prioritized experience replay and stochastic weight averaging. Finally, the results of the experiment show that the proposed scheme performs better than other DRL methods.

End users have limited resources to handle computationally intensive tasks, although MEC has plenty of processing capacity. To this end, a framework with few users was designed to handle compute-intensive by the adjacent processor [126]. MEC servers connected to base stations have sufficient computational power and communication channels. To manage the burden of a user, the authors developed an architecture with several static and car-assisted MEC servers. A partial computation offloading strategy was proposed based on DNNs. The suggested technique can evaluate the best offloading action based on stochastic task arrivals and wireless channel dynamics.

In this research [127], the authors used supervised deep learning to study the partial offloading method in MEC. The suggested method called the comprehensive and energy-efficient deep learning-based offloading technique, effectively chooses the partial offloading strategy and the size of each task's component to minimize service latency and device energy usage. The authors utilized deep learning to concurrently determine the optimal offloading strategy and the appropriate task partitioning. Kumar et al. [128] investigated the task offloading method in UAV-enabled MEC systems, in which end users offload compute-intensive tasks to the UAV to reduce overall costs in terms of weighting delay and energy usage. End users can either perform the task with local computation or offload to a UAV, which operates like a computational server. Nevertheless, according to the computational bottleneck as well as restricted channel capacity among the UAV and the end users, offloading to the UAV becomes challenging. To identify the best offloading decision, the authors deployed a distributed DNN [129]. The simulation findings demonstrate that the trained DNN offloading decision may obtain near-optimal efficiency with a variety of system parameters. The authors provided optimal resource allocation and offloading decisions for a wirelessly powered MEC system. Researchers suggested a modular technique for resolving time fractions and the number of task components, while also addressing the partial offloading scheme using a deep learning approach. With the aid of a trained DNN, the minimization of end devices cost, and energy usage is investigated using a double antenna hybrid access point. Simulation results demonstrate the effectiveness of the proposed study over benchmarks.

5.2 CNN-based Offloading Mechanism

Yang et al., [130] provided a scheme for offloading a portion of the end device's CNN inference computation to the edge servers. Batching activities on GPUs can significantly decrease average computation time on GPUs, according to their findings. Depending on such crucial findings, the authors presented a technique that incorporates all end-device activities and the associated batching effect at the edge servers, in contrast to previous work on cooperative inference which allows every device to make offloading decisions separately.

In addition, an online approach is presented to deal with the CNN inference jobs arrived in random nature. It also substantially decreases the average convergence speed without knowing the future task arrivals. Chai, Song [131] presented a CNN-based cooperative computing system. A joint task management architecture was designed to achieve efficient information interaction and task management. The authors developed an overall task latency reduction problem based on the architecture and solved it to achieve the combined job offloading, CNN layer scheduling, and resource allocation method. Numerous simulations were used to determine the effectiveness of the proposed strategy over baseline methods.

Zeng el al., [132] investigated a multi-layered vehicular edge computing architecture where vehicles can offload different tasks through one of the three given offloading mechanisms. These techniques are chosen based on the network state and server load. The proposed method is to anticipate offloading effectiveness based on previously collected data, such that requested vehicles may select the most successful offloading technique. Moreover, the authors established a deep learning-based approach for task offloading outcomes including success and failure and service-delay estimation. An automatic feature generation model was also proposed to acquire the interconnections of attributes to produce new attributes, attempting to avoid the efficiency chaos caused by manually designed features. The effective task offloading and the shortest service delay are chosen as the final selection based on performance projection.

Deep learning empowered visual target navigation on UAVs is considered in this study. Because small UAVs have restricted computational power as well as limited energy expenditure, a new implementation of a trained CNN model for target detection is used, in which the lower layers of the CNN were being distributed on the UAV and the corresponding higher layers were being distributed at the MEC server [86]. The above configuration meets the requirement for quick and efficient video image processing while considering realistic limitations. The explanatory results help in establishing important insights into wireless edge networks driven by UAV tracking.

The authors in [**?**] investigated the problem of optimal computation offloading in vehicular edge computing. The aim was to reduce the overall cost in terms of the tradeoff between task delays and energy optimization. A Markov decision process was used to model the problem and DRL to deal with enormous state space. To extract the features and approximate the policy and function, the authors considered CNN. The proposed solution was evaluated with six different baseline techniques to show the performance. A joint CNN and LSTM-based offloading decision optimization framework was proposed in [133]. The authors investigated the load at the end devices using joint DNNs to optimize the offloading strategy. The CPU utilization of end devices was also predicted in the proposed framework.

The author formulated the problem of optimal offloading decisions among local and remote computation using DRL [134]. The dynamics of wireless channel gains were taken as input to obtain the optimal offloading decision using CNN. The authors aimed to enhance the computation rate for all end devices while allocating optimal resources. The proposed system learns from experience using DRL technique. A memory was designed to store the optimal action for future offloading actions. Compared with existing algorithms, the proposed solution minimizes the execution latency while achieving optimal computation rates.

5.3 RNN and LSTM based Offloading Mechanism

In [114], the authors provided a multiuser multitasking hybrid computational offloading model based on DRL for offloading a set of tasks generated by several users to an edge server and neighboring devices. The suggested framework, instead of trying to make individual decisions for numerous different computation-intensive tasks, makes global computing offloading decisions considering the impact of users' offloading decisions on the system's competitiveness. The study's main objective is to minimize long-term total system latency. The model increases the convergence stability and speed of the DRL mode by extracting task and network state feature information using RNN.

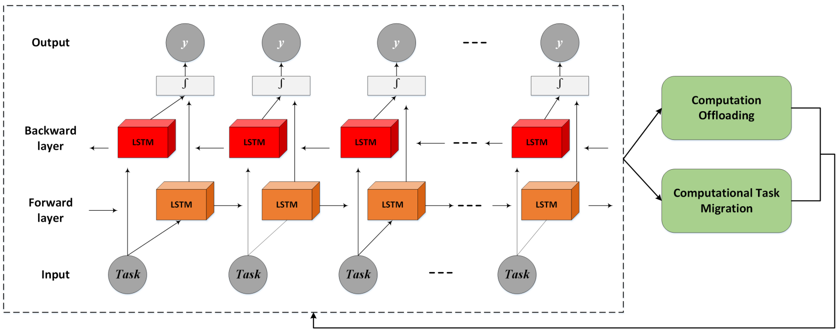


Fig. 10: Efficient Computation Offloading and task migration using LSTM

In close collaboration with artificial intelligence technologies, a novel MEC architecture was proposed that provides intelligent computation offloading. A computation offloading, as well as task migration methodology based on task prediction, is provided [135]. To optimize the edge computing offloading model, task migration is scheduled at the edge and cloud server, where computation task prediction is modeled with LSTM. Results demonstrated that the proposed architecture and its algorithm can successfully reduce the overall task latency even when data and sub-tasks are larger.

In [136], researchers investigated the offloading tasks in cognitive vehicular networks and considered both the onboard computing resource as well as the computing resource of a remote cloud server. The proposed two stages of overall strategic management are resource exploration and computational offloading. An LSTM-based resource discovery technique is suggested to address the dynamics and uncertainty of the availability of resources at vehicle cloud computing while increasing sub-tasks and data. The authors in [57] provided a user prediction-based computation offloading scheme based on LSTM. The authors investigated that the nature of mobility data is nonlinear and time series. The proposed scheme uses mobility features for prediction including location, velocity, and direction. The authors also investigated the priority weights to choose the optimal edge to serve for task offloading. The simulation results demonstrate the effectiveness of the proposed work compared to baseline techniques.

The authors in [58] introduced an innovative approach for secure and energy-efficient computational offloading, harnessing the power of LSTM. The essence of this approach lies in predicting computational tasks through LSTM algorithms, which subsequently guide the strategy for offloading computations on mobile devices. Additionally, the scheme incorporated task migration as part of edge cloud scheduling, ultimately optimizing the entire edge computing offloading model as presented in figure 10. Experimental results substantiate the efficacy of our proposed architecture, combining LSTM-based offloading techniques and the LSTMOTR (LSTM Offloading and Routing) algorithm. This architecture not only significantly reduces total task delay as data and subtasks grow but also minimizes energy consumption. Furthermore, it enhances device security by leveraging the inherent firewall capabilities of LSTM algorithms.

6 Comparison of Existing Schemes

Based on the above section, we compare the state-of-the-art studies. We categorize this comparison into three branches. i) Comparisons of DNN-based schemes, ii) Comparison of CNN-based schemes, and iii) Comparison of RNN-based schemes. All the studies that are discussed in the previous section are compared in tabular form. This comparison is based on the following key parameters that were discussed in the taxonomy. (i) Objective: In this parameter, we provide the major objective of the study. (ii) DNNs variation: This parameter provides the use of specific DNN to achieve the major objective. (iii) Learning approach: In this parameter, we show the learning scheme used by state-of-the-art studies (iv) Offloading platform: In this parameter, we compare the studies and separate them based on the specific offloading platform. Moreover, we also provide the strengths and limitations of state-of-the-art studies that are crucial to unlocking the doors for novel research. Table 5 presents the comparison of state-of-the-art techniques based on various parameters.

Table 5: Comparison of state-of-the-art techniques

Study	Objective	DNNs variation	Learning approach	Offloading platform	Strength(s)	Limitation(s)
[34]	Subcarrier allocation and decision among local and remote computation	DNN	DRL	Edge server	Less execution latency	Dense networks
[32]	Optimal offloading actions among local and remote decision	DNN	DRL	Edge server	Less execution latency	Interference among end devices
[87]	Investigation of MIMO systems for multimedia services	N\A	DRL	Edge server	Less energy consumption	Poor performance analysis
[88]	Allocation of optimal network bandwidth to enhance wireless spectrum	DNN	DRL	Edge server	Optimization of spectrum	Training of deep Q andDNN
[89]	Binary offloading scheme for AR in MEC	DNN	N\A	Edge server	Computation rate	Q value prediction
[122]	Investigation of resource allocation for vehicles	DNN	DRL	Edge server	Low response time	Location prediction
[123]	Investigation of QoS for computation offloading for vehicles	N\A	DRL	RSU	Stability and convergence	Complexity
[124]	Investigation of computation offloading under stochastic task arrivals	DNN	DRL	Edge server	Optimal offloading policy	Response time
[125]	Minimization of service latency and end device energy usage	DNN	Supervised	Edge server	Less service delay	Computational complexity
[126]	Investigation of task offloading scheme in UAV enabled MEC systems	DNN	N\A	UAV	Less average delay	Dynamic networks
[127]	Investigation of time fractions to minimize cost and energy	DNN	N\A	Edge server	Less energy consumption	Resource limitations
[128]	Offloading a portion of CNN inference at the edge servers to minimize inference time	CNN	N\A	Edge server	Minimum inference time	Resource constrained
[129]	Execution of tasks at edge and cloud servers deployed with CNN model	CNN	N\A	Edge and cloud server	Less task latency	Complexity of joint model deployment
[130]	Optimal offloading scheme for vehicles based on network state and server load	CNN	Supervised	RSU	Less failure rate	Energy consumption and huge data
[84]	Joint offloading of CNN model at UAVs and edge servers	CNN	N\A	UAVs and edge server	Less weighted sum cost	Complexity of joint model deployment
[131]	Investigation of computation offloading scheme for vehicular edge computing	CNN	DRL	Edge server	Less latency and energy consumption	Resource limitation
[132]	Investigation of load on end devices to optimize offloading	CNN, LSTM	N\A	Edge server	Less delay and energy consumption	Complexity of joint model deployment
[?]	Investigation of optimal computing mode selection	CNN	DRL	Edge server	Less execution latency	Manually labeling data
[37]	The simultaneous offloading decisions for all the tasks	RNN, LSTM	DRL	Edge server	Less system latency	Single objective optimization
[56]	Task prediction and scheduling to improve security	LSTM	DRL	Edge server	Less delay and energy consumption	Complexity of joint task prediction and migration
[133]	Intelligent task prediction to optimize offloading strategy	LSTM	N\A	Edge server	Less tasks delay	Mobility and security
[134]	Investigation of on-board resources of vehicles	LSTM	DRL	RSUs and cloud	Utilization of on-board resources	Robustness
[55]	Investigation of user direction based computation offloading	LSTM	N\A	Edge server	Less latency and energy consumption	Dynamic environment

6.1 Comparison of DNN-based Schemes

Article [34] investigated the problem of sub-carrier allocation and decision among local and remote computation using a DNN. The authors considered NOMA as a channel access mechanism. Article [32] provided an online algorithm that learns the optimal offloading actions among local and remote decisions using DRL. In [87], the authors used the MIMO system to meet the demanding specifications of the recently introduced multimedia services in MEC. The authors in [88] authors provided optimal network bandwidth to enhance the wireless spectrum for links between users and particular computational activities. An intelligent system was proposed which chooses the optimal offloading techniques using DRL. The author investigated a binary offloading scheme in [89] for AR edge computing which optimizes the computation of AR devices. The proposed DNN-based AR offloading decision effectively reduces the computational complexity, hence neglecting combinational optimization. In [90], the authors distributed the offloading scheme by dividing the tasks into sub-tasks to reduce the cost of offloading. A resource allocation problem was investigated in [122] to tackle vehicle computations. In [123], the quality of service for computation offloading was investigated. The proposed framework decreases the training time by exchanging stochastic gradient descent with the experience replay buffer. An intelligent offloading scheme based on DNN was studied in [124] that tackles the dynamic channels and stochastic task arrivals. The authors effectively choose the partial offloading strategy and the size of each task to minimize service latency and user energy usage [125]. Kumar et al., [126] investigated the task offloading scheme in UAV-enabled MEC systems, where end users offload tasks to the UAVs. The aim was to reduce overall costs in terms of weighting delay and energy usage. Using deep learning, the authors provided a scheme to minimize the cost and energy of end devices [127]. An effective technique was proposed to solve time fractions and the number of tasks while addressing partial offloading.

6.2 Comparison of CNN-based Schemes

In [128], authors offloaded a portion of CNN inference of end devices at the edge servers. The authors also investigated that batching tasks on GPU can reduce the average inference time. A joint task management system was designed in [129] to reduce the overall latency. The authors also formulated a resource allocation method. In [130], the authors investigated an optimal offloading scheme for vehicles based on network state and server load. A deep learning-based scheme was proposed to predict the optimal offloading scheme. The authors in [84] investigated the joint offloading of the CNN model, where the lower layers of the CNN are distributed on the UAV and the corresponding higher layers are distributed at the MEC server. In [131], the authors investigated the computation offloading problem for vehicle edge computing, where a vehicle tends to schedule its task. The aim was to minimize the queuing time and total cost in terms of the trade-off between task latency and energy consumption. In the article [132], a joint CNN and LSTM-based approach was proposed to investigate the load at the end devices and to optimize the offloading strategy. The aim was to minimize the delay and energy consumption for end devices. An optimal offloading decision problem was investigated in [**?**], where the authors considered the DRL approach that learns from past experiences. For effective computation offloading decisions, the binary computation offloading strategy was considered.

6.3 Comparison of RNN and LSTM based Schemes

In [37], the authors provided a hybrid task offloading scheme that utilizes simultaneous decisions instead of one by one. The aim was to minimize the overall system latency using RNN and LSTM-based DRL. Another LSTM-based computation offloading strategy was investigated in [56]. The authors provided an effective approach to predict the tasks for computation offloading and provided a scheme to schedule the tasks at the edge server. The system provides enough security for end devices with the firewall nature of LSTM. The authors in [133] provided an intelligent scheme that predicts the task to optimize the computation offloading strategy. The tasks are scheduled and migrated on the edge cloud server to reduce the task execution delay. The experimental results validate the performance of task delay while increasing data and sub-tasks. A prediction of onboard resources for vehicles was investigated in [134]. Researchers also provided a DRL-based collaborative computation offloading at vehicle cloud servers and remote cloud servers. In [55], the authors investigated a computation offloading strategy based on user direction prediction that was neglected by previous studies. Moreover, the authors calculated the priority weights to select the optimal edge server.

7 Features and Open Issues

In this section, we provide features of DNNs and arise different questions in DNN-based applications for computation offloading in MEC. We present the open issues regarding the applications of DNNs in computation offloading. We also suggest possible solutions to cope with these open issues. Figure 11 shows the significant features of DNN and open issues with their possible solutions.

7.1 Features and Open Issues in DNN based Channel Estimation

How to enhance the performance of DNN-based channel estimation while taking channel selection, mobility, load estimation, and load balancing into consideration? DNNs are an attractive solution for tackling various challenges in channel estimation. These DNNs enable the use of smart channel estimation where an access point can learn when to transmit on each type of channel. Furthermore, DNNs may also enable multi-mode access points to direct their traffic between mm-Wave, and microwave [137]. However, there are still various open issues.

DNNs applications in offloading	Features of DNN	Challenges	Possible Solutions
Channel Estimation	• Channel selection • Load distribution • Load estimation	• Channel selection • MmWave channel modeling • Proactive behavior of channels • Mobility prediction	• CNN based framework • SNN based framework • SNN based framework • RNN based framework
Caching	• Desired Content at network edge • Predicting user behavior	• Content correlation • Content correlation for users • • Data filtering • Finite resources of DNNs to store contents	• CNN based framework • CNN based framework for clustering users • CNN based framework • RNN based framework for computational resource allocation
AR/VR	• Predicting user environment • Predicting user movement • Resource management	• Limited resources • Limited time for training DNNs • Erroneous data	• DNN based framework for managing resources • CNN based framework for fast training and convergence • CNN based framework to predict erroneous data
Resource Allocation	• Intelligent data analytics • Extraction of patterns and relationships from data • Intelligent strategies such as subcarrier allocation	• Data compression • Data classification • Tradeoff between computational needs and accuracy of DNNs	• CNN based framework • CNN based framework • RNN for computational needs management with LSTM for accuracy
Mode Selection	• Optimal offloading action • Allocation of resources such as wireless power transfer duration	• Decision among online or offline training • Separate channels for uplink and downlink • Neglecting assumption of fading channels in single time frame	• CNN based framework for fast convergence in online training • DNN based framework for decision • SNN based framework to deal with dynamic channels in each time frame.
UAVs	• Dynamic location adjustment • Resource allocation • Data analytics to predict user behavior.	• Limited resources • Dynamic air-ground channel issues • UAVs trajectory and erroneous data	• DNN based framework to manage resources for training and services. • SNNs based channel modeling • RNN based trajectory planning to reduce chances or errors
Vehicle Management	• Data analytics from big data • Pattern recognition to detect pedestrians and obstacles • Collection, processing and releasing of traffic information	• Vehicle trajectory • Limited resources • Lack of datasets and Data classification • Uniform decisions	• RNN based framework • DNN based resource management • CNN based data classification to cope with limited datasets

Fig. 11: Features, challenges, and possible solutions of DNNs-based applications for computation offloading

Previous studies consider reactive approaches. In these approaches, the data request is first initiated. After this initiation, resources are allocated based on corresponding delay tolerance. Moreover, previous studies do not consider the proactive and predictable behavior of traffic in channels. These studies do not take into consideration the future of peak times while data traffic is distributed among different channels. In addition, load balancing among channels, load estimation, and mobility aware estimation are also significant directions for advanced research. *For example, CNN-based models may be deployed for channel selections. SNN-based algorithms may enhance the performance of load estimation etc.*

7.2 Features and Open Issues in DNN-based Caching

How to ensure the desired content in the cache using DNNs based caching while ensuring data filtering, content classification, and limited storage? DNNs enables is witnessed as a promising solution to place the desired content at the network edge. Since the problem of cache placement and cache update depends on the user's behaviors. This cache update depends on the frequency of a certain request, and cache placement depends on the user's behavior. DNNs are well recognized as a better solution to predict user behavior. However, there are still some open issues that need to be considered. First, how to clean data and extract useful content from huge data sources? It is necessary to read and extract the data depending on user type. For example, the data may be in hundreds of GB depending on the user's type, which requires pre-processing and filtering of data. This data processing and filtering may take more time than learning [138]. Moreover, the memory of DNNs is limited. Hence, DNNs may contain a limited number of user requests. *The possible solutions may be CNN-based algorithms for content correlation and clustering users for classification. SNN may be used for demand predictions etc.*

7.3 Features and Open Issues in DNN-based AR/VR

How to ensure the desired service of DNNs based AR/VR, while taking limited resources, limited time for training DNNs, and errors in collected data into consideration? DNNs are widely considered a promising solution for AR/ VR application management [139]. Compared to the other offloading applications, AR/VR applications are more concerned with user behavior and environment. DNNs are efficient mechanisms to predict the user's movement and behaviors. These predictions enable the access points to improve the generation of AR/VR images. DNNs also enable effective resource management for these applications. However, it faces many challenges. First, due to the large data sizes, each 360 degrees image, DNNs require a large number of computational resources. In MEC and offloading, it may not possible. Second, data collected from users contains many errors. In this case, access points may need to use inaccurate data for training, which may significantly reduce the prediction accuracy of DNNs. Here, efficient resource management for processing AR/VR images and for training DNNs is also s significant challenge. *CNNs may be used for the prediction of erroneous data in VR offloading applications. RNNs may be used for the prediction of AR/VR users' movement.*

7.4 Features and Open Issues in DNN-based Resource Allocation

How to ensure effective resource allocation while taking limited energy, data classification, and a tradeoff between computational needs and accuracy of DNNs into consideration? DNNs are widely recognized as key models to leverage intelligent data analytics to extract the relationship and patterns from data offloaded by IoT devices [9]-[140]. These can also be used for intelligent data compression and data recovery. Moreover, DNNs may be deployed to adapt intelligent strategies such as channel subcarrier allocations. They can dynamically select the most appropriate frequency bands based on the network state. However, it faces many challenges. First, the offloaded data for storage and computational resources may be in an erroneous form. It is necessary to classify and prevent the data with DNNs. In other words, DNNs should be able to tolerate this erroneous data. Second, it is important to consider the trade-off between the accuracy required for DNNs computational needs for training DNNs. Last, the MEC may generate a thousand types of data offloaded by IoT devices. DNNs should be able to choose accurate data for offloading purposes and resource allocations. *One can use CNNs for data compression, and RNN for computational resource management. In addition, DNN can be used for subcarrier allocation.*

7.5 Features and Open issues in DNNs based Mode Selection

How to ensure the best offloading mode selection while taking channel access mechanisms, online or offline training, and separate channels for uplink and downlink into consideration? Moreover, how to ensure the best offloading action while neglecting the assumption of fading channels in a single time frame? DNNs are commonly used models for offloading mode selection. These are promising solutions for choosing the best offloading actions among local or remote and partial offloading. Moreover, DNNs also contribute to allocating resources such as wireless power transfer duration, transmission time, etc [13]-[141]. However, DNN-based mode selection faces many challenges. First, how to efficiently choose optimal offloading action using different channel access mechanisms such as OFDMA. Second, online training of DNNs requires huge computational resources. In addition, offline training of DNNs may not meet the constraints of fading channels. Third, it is much more difficult to train a DNN while taking dynamic channels for each time frame. Last, in the case of partial offloading, previous studies do not consider which part of the task should be computed locally and which part should be offloaded. *DNNs may be deployed to choose which part should be computed locally and which part should be offloaded. SNNs may be used to deal with dynamic channels in each time frame. CNN may be used for online training to ensure fast convergence.*

7.6 Features and Open Issues in DNN-based UAVs

How to train and improve UAV performance while taking limited power to train DNNs, limited time, and erroneous data due to air-ground channels into consideration? UAVs play a significant role in tracking user environments. DNNs enable UAVs to dynamically adjust their locations, resource allocation decisions, flying direction, and path planning to serve users on the ground. In addition, DNNs may be used for data analytics to predict the user's behaviors on the ground. However, DNNs based UAVs face many challenges. First, UAVs have limited computational power to train DNNs. Second, these have limited time to collect data due to their flying nature. Third, UAVs should be able to tackle the trade-off between energy to train DNN and energy to serve the users on the ground. Lastly, collected data from the ground due to environmental effects may have an error which may result in poor accuracy. *SNNs may be applied to deal with air-ground channel constraints. DNNs may be used to manage the limited resources of UAVs. Moreover, RNN may be implemented to predict the trajectory of UAVs which can reduce the errors in the collection of data.*

7.7 Features and Open Issues in DNN-based Vehicle Management

How to manage vehicles while taking the lack of datasets, trajectory, limited resources, uniform decisions, and computations of big data into consideration? DNNs play a crucial role in data analytics. They extract knowledge from the big data generated by vehicles. Moreover, DNNs help in pattern recognition to detect pedestrians and obstacles on road [142]. These can also be used to deal with raw data. Collection, processing, and releasing of traffic information may also be improved with DNNs. However, it faces many challenges. Even though DNNs have been applied to solve various edge network problems, very few studies have considered them for vehicle management. Second, there is a lack of datasets regarding vehicles and roads to simulate the DNNs. Third, uniform decisions and compatibility are a challenge. Processing big data with limited resources of vehicles and edge servers is also challenging. Finally, mobility and trajectory prediction not works everywhere [143]. *RNNs may be applied to predict a vehicle's mobility trajectory. DNNs may be used for optimal uniform decisions. In addition, CNNs may be used for data classifications.*

8 Challenges linked to the utilization of DNNs for computation offloading

8.1 Dynamic Environment Adaptation

One of the primary challenges in DNN-based computation offloading is adapting to the dynamic nature of edge computing environments. Edge networks can experience rapid changes in network conditions, device availability, and server loads. For example, mobile devices' movement makes the collaborative edge site difficult to determine and the limited battery makes the device unable to work continuously [173]. Ensuring that DNN models can make effective offloading decisions in real-time despite these fluctuations is crucial. Researchers need to develop algorithms and training strategies that enable DNNs to continuously learn and adapt to changing circumstances. Techniques like reinforcement learning and online learning may play a pivotal role in addressing this challenge.

8.2 Data Privacy and Security

Data privacy and security are paramount concerns when deploying DNNs for offloading. Edge devices often process sensitive data, and offloading decisions involve sharing data with remote servers. Balancing the need for data privacy with the utility of offloading is a complex problem. For example, a common problem identified from previous research work is the leakage of privacy at the edge layer and data accessed by unauthorized people [172]. Researchers must explore privacy-preserving techniques like federated learning, homomorphic encryption, and differential privacy to ensure that DNN models can make informed offloading decisions without compromising data confidentiality. Additionally, developing robust security mechanisms to protect against potential attacks on DNN-based offloading systems is essential.

8.3 Energy Efficiency

Edge devices, particularly IoT sensors and mobile devices, are often battery-powered and have limited energy resources. DNN inference can be computationally intensive and energy-consuming. Balancing the benefits of offloading with energy efficiency is a significant research challenge. Researchers need to investigate techniques for energy-efficient DNN model design, lightweight model architectures, and dynamic power management strategies. This includes optimizing the trade-off between local computation and offloading to minimize energy consumption while meeting application performance requirements.

8.4 Real-time Decision Making:

Real-time decision making is critical in many edge computing applications, such as autonomous vehicles and augmented reality. DNN-based offloading decisions must be made quickly to ensure timely responses. However, the inference process itself introduces latency. Researchers need to develop strategies that optimize the decision-making process, possibly through model quantization, edge caching, or parallel processing, to minimize the overall latency while still leveraging the advantages of DNN-based offloading.

8.5 Model Generalization

Model generalization is the ability of a trained DNN to perform well on data it hasn't seen during training. In the context of computation offloading in MEC, DNN models need to generalize effectively across various edge scenarios, device types, and network conditions. The challenge arises because MEC environments are diverse and dynamic. Edge scenarios can range from smart factories to smart cities, each with unique characteristics and requirements. Ensuring that a DNN model trained for one scenario can perform adequately in another is complex. Researchers are exploring techniques to enhance model generalization. This includes data augmentation, where the training dataset is artificially diversified to simulate different edge scenarios. Transfer learning is another approach, where a pre-trained DNN is fine-tuned for specific edge environments, leveraging knowledge from a broader context. Reinforcement learning can be employed to enable models to adapt to changing edge conditions in real-time. Achieving robustness and adaptability in DNN models is essential to make computation offloading practical and reliable across the diverse landscapes of MEC. This challenge highlights the need for versatile and scalable DNN architectures and training strategies.

9 Conclusion

In this article, we offered an overview of how applications based on DNNs addressed the challenges associated with computation offloading. These applications encompassed tasks such as channel estimation, caching, AR and VR applications, resource allocation, mode selection, UAVs, and vehicle management. We constructed a comprehensive taxonomy encompassing the diverse applications of DNNs, their variations, learning approaches, and the platforms used for offloading. We summarized existing research works that concentrated on different learning approaches within offloading scenarios. These studies were systematically compared based on various parameters, highlighting their strengths and limitations. Additionally, we provided open issues based on each application and specific challenges linked to the utilization of DNNs in computation offloading. In conclusion, we found that applications based on DNNs have proven to be effective solutions for addressing various challenges in the context of MEC offloading. In the future, we aim to explore the novel architectures of DNNs specifically optimized for resource-constrained edge devices. We also aim to propose a robust framework for ensuring the security and privacy of offloaded tasks.

Declaration of competing interest

The authors declare no competing interests.

Declaration of generative AI in scientific writing

The authors declare that they used generative AI solely to enhance the readability and language of the Introduction section.

Authors Contributions

Conceptualization EM and JS; Data curation EM, FR, and MM; Funding acquisition MM and MKK; Investigation EM, MB and FR; Methodology EM and AR; Project administration JS; Resources JS and MB; Software EM; Supervision FR and MKK; Validation AR; Visualization EM; Writing - original draft EM; Writing - review and editing JS and MKK.

References

[1] Feng, C., Han, P., Zhang, X., Yang, B., Liu, Y. & Guo, L. Computation offloading in mobile edge computing networks: A survey. *Journal Of Network And Computer Applications*. pp. 103366 (2022)

[2] Huda, S. & Moh, S. Survey on computation offloading in UAV-Enabled mobile edge computing. *Journal Of Network And Computer Applications*. pp. 103341 (2022)

[3] Sadatdiynov, K., Cui, L., Zhang, L., Huang, J., Salloum, S. & Mahmud, M. A review of optimization methods for computation offloading in edge computing networks. *Digital Communications And Networks*. (2022)

[4] Maray, M., Mustafa, E., Shuja, J. & Bilal, M. Dependent task offloading with deadline-aware scheduling in mobile edge networks. *Internet Of Things*. **23** pp. 100868 (2023)

[5] Mustafa, E., Shuja, J., Jehangiri, A., Din, S., Rehman, F., Mustafa, S., Maqsood, T., Khan, A. & Others Joint wireless power transfer and task offloading in mobile edge computing: a survey. *Cluster Computing*. **25**, 2429-2448 (2022)

[6] Jehangiri, A., Maqsood, T., Ahmad, Z., Umar, A., Shuja, J., Alanazi, E., Alasmary, W. & Others Mobility-aware computational offloading in mobile edge networks: a survey. *Cluster Computing*. **24**, 2735-2756 (2021)

[7] Adhikari, M. & Hazra, A. 6G-Enabled Ultra-Reliable Low-Latency Communication in Edge Networks. *IEEE Communications Standards Magazine*. **6**, 67-74 (2022)

[8] Zaman, S., Maqsood, T., Ramzan, A., Rehman, F., Mustafa, S. & Shuja, J. Deadline-aware heuristics for reliability optimization in ubiquitous mobile edge computing. *International Journal Of Data Science And Analytics*. pp. 1-13 (2023)

[9] Chen, M. & Debbah A. Artificial neural networks-based machine learning for wireless networks: A tutorial. *IEEE Communications Surveys and Tutorials*. **21**, 3039-3071 (2019)

[10] Saleh, H., Saber, W. & Rizk, R. Mobile Computation Offloading in Mobile Edge Computing Based on Artificial Intelligence Approach: A Review and Future Directions. *International Conference On Advanced Machine Learning Technologies And Applications*. pp. 593-603 (2022)

[11] Mezair, T., Djenouri, Y., Belhadi, A., Srivastava, G. & Lin*, J. Towards an Advanced Deep Learning for the Internet of Behaviors: Application

to Connected Vehicle. *ACM Transactions On Sensor Networks (TOSN).* (2022)

[12] Emani, S., Soman, K., Sajith Variyar, V. & Adarsh, S. Obstacle detection and distance estimation for autonomous electric vehicle using stereo vision and DNN. *Soft Computing And Signal Processing.* pp. 639-648 (2019)

[13] Zhang, S., Gu, H., Chi, K., Huang, L., Yu, K. & Mumtaz, S. Drl-based partial offloading for maximizing sum computation rate of wireless powered mobile edge computing network. *IEEE Transactions On Wireless Communications.* (2022)

[14] Bhattacharya, S. & Gupta, A. Deep Learning for THz Channel Estimation and Beamforming Prediction via Sub-6GHz Channel. *2022 IEEE International Conference On Signal Processing And Communications (SPCOM).* pp. 1-5 (2022)

[15] Shakarami, A., Ghobaei-Arani, M. & Shahidinejad, A. A survey on the computation offloading approaches in mobile edge computing: A machine learning-based perspective. *Computer Networks.* **182** pp. 107496 (2020)

[16] Shakarami, A., Ghobaei-Arani, M., Masdari, M. & Hosseinzadeh, M. A survey on the computation offloading approaches in mobile edge/cloud computing environment: a stochastic-based perspective. *Journal Of Grid Computing.* **18**, 639-671 (2020)

[17] Cui, W., Shen, K. & Yu, W. Spatial deep learning for wireless scheduling. *Ieee Journal On Selected Areas In Communications.* **37**, 1248-1261 (2019)

[18] Soltani, M., Pourahmadi, V., Mirzaei, A. & Sheikhzadeh, H. Deep learning-based channel estimation. *IEEE Communications Letters.* **23**, 652-655 (2019)

[19] He, H., Wen, C., Jin, S. & Li, G. Deep learning-based channel estimation for beamspace mmWave massive MIMO systems. *IEEE Wireless Communications Letters.* **7**, 852-855 (2018)

[20] Zhang, Y., Mu, Y., Liu, Y., Zhang, T. & Qian, Y. Deep learning-based beamspace channel estimation in mmWave massive MIMO systems. *IEEE Wireless Communications Letters.* **9**, 2212-2215 (2020)

[21] Jin, Y., Zhang, J., Jin, S. & Ai, B. Channel estimation for cell-free mmWave massive MIMO through deep learning. *IEEE Transactions On Vehicular Technology.* **68**, 10325-10329 (2019)

[22] Wang, Y. & Friderikos, V. A survey of deep learning for data caching in edge network. *Informatics.* **7**, 43 (2020)

[23] Rahman, S., Alam, M. & Rahman, M. Deep learning-based predictive caching in the edge of a network. *2020 International Conference On Information Networking (ICOIN).* pp. 797-801 (2020)

[24] Saputra, Y., Hoang, D., Nguyen, D., Dutkiewicz, E., Niyato, D. & Kim, D. Distributed deep learning at the edge: A novel proactive and cooperative caching framework for mobile edge networks. *IEEE Wireless Communications Letters.* **8**, 1220-1223 (2019)

[25] Akhlaqi, M. & Hanapi, Z. Task offloading paradigm in mobile edge computing-current issues, adopted approaches, and future directions. *Journal Of Network And Computer Applications.* **212** pp. 103568 (2023)

[26] Li, T., He, X., Jiang, S. & Liu, J. A survey of privacy-preserving offloading methods in mobile-edge computing. *Journal Of Network And Computer Applications.* **203** pp. 103395 (2022)

[27] Qian, Y., Wang, R., Wu, J., Tan, B. & Ren, H. Reinforcement learning-based optimal computing and caching in mobile edge network. *IEEE Journal On Selected Areas In Communications.* **38**, 2343-2355 (2020)

[28] Chen, P., Chen, Y., Huang, W., Huang, C. & Tirkkonen, O. DDPG-based radio resource management for user interactive mobile edge networks. *2020 2nd 6G Wireless Summit (6G SUMMIT).* pp. 1-5 (2020)

[29] Chen, X. & Liu, G. Energy-efficient task offloading and resource allocation via deep reinforcement learning for augmented reality in mobile edge networks. *IEEE Internet Of Things Journal.* **8**, 10843-10856 (2021)

[30] Wang, T. & Melchior, P. Graph neural network-based resource allocation strategies for multi-object spectroscopy. *Machine Learning: Science And Technology.* **3**, 015023 (2022)

[31] Lei, L., Yuan, Y., Vu, T., Chatzinotas, S. & Ottersten, B. Learning-based resource allocation: Efficient content delivery enabled by convolutional neural network. *2019 IEEE 20th International Workshop On Signal Processing Advances In Wireless Communications (SPAWC).* pp. 1-5 (2019)

[32] Li, Y., Song, B., Kang, X., Du, X. & Guizani, M. Vehicle-type detection based on compressed sensing and deep learning in vehicular networks. *Sensors.* **18**, 4500 (2018)

[33] Hu, H., Shan, H., Wang, C., Sun, T., Zhen, X., Yang, K., Yu, L., Zhang, Z. & Quek, T. Video surveillance on mobile edge networks—a reinforcement-learning-based approach. *IEEE Internet Of Things Journal.* **7**, 4746-4760 (2020)

[34] Bi, S. & Zhang, Y. Computation rate maximization for wireless powered mobile-edge computing with binary computation offloading. *IEEE Transactions On Wireless Communications.* **17**, 4177-4190 (2018)

[35] Huang, L., Bi, S. & Zhang, Y. Deep reinforcement learning for online computation offloading in wireless powered mobile-edge computing networks. *IEEE Transactions On Mobile Computing.* **19**, 2581-2593 (2019)

[36] Chen, W., Zhu, B., Chi, K. & Zhang, S. DRL based offloading of industrial IoT applications in wireless powered mobile edge computing. *IET Communications.* **16**, 951-962 (2022)

[37] Nduwayezu, M., Pham, Q. & Hwang, W. Online computation offloading in NOMA-based multi-access edge computing: A deep reinforcement learning approach. *IEEE Access.* **8** pp. 99098-99109 (2020)

[38] Chen, X., Zhang, H., Wu, C., Mao, S., Ji, Y. & Bennis, M. Optimized computation offloading performance in virtual edge computing systems via deep reinforcement learning. *IEEE Internet Of Things Journal.* **6**, 4005-4018 (2018)

[39] Wang, Y., Li, M., Ji, R., Wang, M., Zhang, Y. & Zheng, L. A convolutional operation-based online computation offloading approach in wireless powered multi-access edge computing networks. *Computers And Electronics In Agriculture.* **197** pp. 106967 (2022)

[40] Cai, J., Fu, H. & Liu, Y. Deep reinforcement learning-based multitask hybrid computing offloading for multiaccess edge computing. *International Journal Of Intelligent Systems.* (2022)

[41] Yazid, Y., Ez-Zazi, I., Guerrero-González, A., El Oualkadi, A. & Arioua, M. UAV-enabled mobile edge-computing for IoT based on AI: A comprehensive review. *Drones.* **5**, 148 (2021)

[42] Albanese, A., Nardello, M. & Brunelli, D. Low-power deep learning edge computing platform for resource constrained lightweight compact UAVs. *Sustainable Computing: Informatics And Systems.* **34** pp. 100725 (2022)

[43] Wang, L., Wang, K., Pan, C., Xu, W., Aslam, N. & Hanzo, L. Multi-agent deep reinforcement learning-based trajectory planning for multi-UAV assisted mobile edge computing. *IEEE Transactions On Cognitive Communications And Networking.* **7**, 73-84 (2020)

[44] Wang, L., Wang, K., Pan, C., Xu, W., Aslam, N. & Nallanathan, A. Deep reinforcement learning based dynamic trajectory control for UAV-assisted mobile edge computing. *IEEE Transactions On Mobile Computing.* (2021)

[45] Wang, X., Cheng, P., Liu, X. & Uzochukwu, B. Fast and accurate, convolutional neural network based approach for object detection from UAV. *IECON 2018-44th Annual Conference Of The IEEE Industrial Electronics Society*. pp. 3171-3175 (2018)

[46] Liang, L., Ye, H., Yu, G. & Li, G. Deep-learning-based wireless resource allocation with application to vehicular networks. *Proceedings Of The IEEE*. **108**, 341-356 (2019)

[47] Li, J., Luo, G., Cheng, N., Yuan, Q., Wu, Z., Gao, S. & Liu, Z. An end-to-end load balancer based on deep learning for vehicular network traffic control. *IEEE Internet Of Things Journal*. **6**, 953-966 (2018)

[48] Elwekeil, M., Wang, T. & Zhang, S. Deep learning for environment identification in vehicular networks. *IEEE Wireless Communications Letters*. **9**, 576-580 (2019)

[49] Eziama, E., Tepe, K., Balador, A., Nwizege, K. & Jaimes, L. Malicious node detection in vehicular ad-hoc network using machine learning and deep learning. *2018 IEEE Globecom Workshops (GC Wkshps)*. pp. 1-6 (2018)

[50] Song, H., Kim, M., Park, D., Shin, Y. & Lee, J. Learning from noisy labels with deep neural networks: A survey. *IEEE Transactions On Neural Networks And Learning Systems*. (2022)

[51] Goodfellow, I., Bengio, Y. & Courville, A. Deep Learning, chapter Chapter 7: Regularization for Deep Learning. (MIT Press,2016)

[52] Sarvamangala, D. & Kulkarni, R. Convolutional neural networks in medical image understanding: a survey. *Evolutionary Intelligence*. **15**, 1-22 (2022)

[53] Chen, G. A gentle tutorial of recurrent neural network with error backpropagation. *ArXiv Preprint ArXiv:1610.02583*. (2016)

[54] Wang, J., Hu, J., Min, G., Zhan, W., Ni, Q. & Georgalas, N. Computation offloading in multi-access edge computing using a deep sequential model based on reinforcement learning. *IEEE Communications Magazine*. **57**, 64-69 (2019)

[55] Lindemann, B., Maschler, B., Sahlab, N. & Weyrich, M. A survey on anomaly detection for technical systems using LSTM networks. *Computers In Industry*. **131** pp. 103498 (2021)

[56] Yu, Y., Si, X., Hu, C. & Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*. **31**, 1235-1270

(2019)

[57] Zaman, S., Jehangiri, A., Maqsood, T., Umar, A., Khan, M., Jhanjhi, N., Shorfuzzaman, M. & Masud, M. COME-UP: Computation Offloading in Mobile Edge Computing with LSTM Based User Direction Prediction. *Applied Sciences.* **12**, 3312 (2022)

[58] Arif, M., Ajesh, F., Shamsudheen, S. & Shahzad, M. Secure and Energy-Efficient Computational Offloading Using LSTM in Mobile Edge Computing. *Security And Communication Networks.* **2022** (2022)

[59] Tu, Y., Chen, H., Yan, L. & Zhou, X. Task offloading based on LSTM prediction and deep reinforcement learning for efficient edge computing in IoT. *Future Internet.* **14**, 30 (2022)

[60] DeWolf, T. Spiking neural networks take control. *Science Robotics.* **6**, eabk3268 (2021)

[61] Liu, Q. & Zhang, Z. Ultra-low Power Always-on Intelligent and Connected SNN-based System for Multimedia IoT-enabled Applications. *IEEE Internet Of Things Journal.* (2022)

[62] Yang, X., Song, Z., King, I. & Xu, Z. A survey on deep semi-supervised learning. *IEEE Transactions On Knowledge And Data Engineering.* (2022)

[63] Tian, Y., Chen, X. & Ganguli, S. Understanding self-supervised learning dynamics without contrastive pairs. *International Conference On Machine Learning.* pp. 10268-10278 (2021)

[64] Liu, Y., Jin, M., Pan, S., Zhou, C., Zheng, Y., Xia, F. & Yu, P. Graph self-supervised learning: A survey. *IEEE Transactions On Knowledge And Data Engineering.* (2022)

[65] Dey, S., Pal, S. & Saravanan, V. Computational Offloading with Deep Supervised Learning for Edge enabled IoT. *Technological Advances In Science, Medicine And Engineering Conference 2021.* (2021)

[66] Alam, T., Ullah, A. & Benaida, M. Deep reinforcement learning approach for computation offloading in blockchain-enabled communications systems. *Journal Of Ambient Intelligence And Humanized Computing.* pp. 1-14 (2022)

[67] Yang, S., Lee, G. & Huang, L. Deep Learning-Based Dynamic Computation Task Offloading for Mobile Edge Computing Networks. *Sensors.* **22**, 4088 (2022)

[68] Qu, C., Calyam, P., Yu, J., Vandanapu, A., Opeoluwa, O., Gao, K., Wang, S., Chastain, R. & Palaniappan, K. DroneCOCoNet: Learning-based edge computation offloading and control networking for drone video analytics. *Future Generation Computer Systems*. **125** pp. 247-262 (2021)

[69] Yang, J., Chen, X., Wang, D., Zou, H., Lu, C., Sun, S. & Xie, L. Deep learning and its applications to wifi human sensing: A benchmark and a tutorial. *ArXiv Preprint ArXiv:2207.07859*. (2022)

[70] Glielmo, A., Husic, B., Rodriguez, A., Clementi, C., Noé, F. & Laio, A. Unsupervised learning methods for molecular simulation data. *Chemical Reviews*. **121**, 9722-9758 (2021)

[71] Baadel, S. Big Data Analytics: A Tutorial of Some Clustering Techniques. *Int. J. Management And Data Analytics*. **1**, 38-46 (2021)

[72] Chen, X., Xu, H., Zhang, G., Chen, Y. & Li, R. Unsupervised Deep Learning for Binary Offloading in Mobile Edge Computation Network. *Wireless Personal Communications*. **124**, 1841-1860 (2022)

[73] Zhang, D., Cao, L., Zhu, H., Zhang, T., Du, J. & Jiang, K. Task offloading method of edge computing in internet of vehicles based on deep reinforcement learning. *Cluster Computing*. **25**, 1175-1187 (2022)

[74] Alelaiwi, A. An efficient method of computation offloading in an edge cloud platform. *Journal Of Parallel And Distributed Computing*. **127** pp. 58-64 (2019)

[75] Levine, S., Kumar, A., Tucker, G. & Fu, J. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. *ArXiv Preprint ArXiv:2005.01643*. (2020)

[76] Feriani, A. & Hossain E. Single and Multi-Agent Deep Reinforcement Learning for AI-Enabled Wireless Networks: A Tutorial *IEEE Communications Surveys and Tutorials*. **23**, 1226-1252 (2021)

[77] Mekrache, A., Bradai, A., Moulay, E. & Dawaliby, S. Deep reinforcement learning techniques for vehicular networks: recent advances and future trends towards 6G. *Vehicular Communications*. pp. 100398 (2021)

[78] Domeke, A., Cimoli, B. & Monroy, I. Integration of Network Slicing and Machine Learning into Edge Networks for Low-Latency Services in 5G and beyond Systems. *Applied Sciences*. **12**, 6617 (2022)

[79] Alwarafy, A., Abdallah, M., Ciftler, B., Al-Fuqaha, A. & Hamdi, M. Deep reinforcement learning for radio resource allocation and management in next generation heterogeneous wireless networks: A survey. *ArXiv Preprint*

ArXiv:2106.00574. (2021)

[80] Wang, S., Xu, J., Zhang, N. & Liu, Y. A survey on service migration in mobile edge computing. *IEEE Access*. **6** pp. 23511-23528 (2018)

[81] Liu, H. & Cao, G. Deep Reinforcement Learning-Based Server Selection for Mobile Edge Computing. *IEEE Transactions On Vehicular Technology*. **70**, 13351-13363 (2021)

[82] Chen, J., Chen, S., Wang, Q., Cao, B., Feng, G. & Hu, J. iRAF: A deep reinforcement learning approach for collaborative mobile edge computing IoT networks. *IEEE Internet Of Things Journal*. **6**, 7011-7024 (2019)

[83] Liu, Y., Yu, H., Xie, S. & Zhang, Y. Deep reinforcement learning for offloading and resource allocation in vehicle edge computing and networks. *IEEE Transactions On Vehicular Technology*. **68**, 11158-11168 (2019)

[84] Abualigah, L., Diabat, A., Sumari, P. & Gandomi, A. Applications, deployments, and integration of internet of drones (iod): a review. *IEEE Sensors Journal*. (2021)

[85] Dierks, T. & Jagannathan, S. Output feedback control of a quadrotor UAV using neural networks. *IEEE Transactions On Neural Networks*. **21**, 50-66 (2009)

[86] Yang, B., Cao, X., Yuen, C. & Qian, L. Offloading optimization in edge computing for deep-learning-enabled target tracking by internet of UAVs. *IEEE Internet Of Things Journal*. **8**, 9878-9893 (2020)

[87] Aljeri, N. & Boukerche, A. A probabilistic neural network-based road-side unit prediction scheme for autonomous driving. *ICC 2019-2019 IEEE International Conference On Communications (ICC)*. pp. 1-6 (2019)

[88] Li, M., Gao, J., Zhao, L. & Shen, X. Deep reinforcement learning for collaborative edge computing in vehicular networks. *IEEE Transactions On Cognitive Communications And Networking*. **6**, 1122-1135 (2020)

[89] Tilahun, F., Abebe, A. & Kang, C. Joint Communication and Computing Resource Allocation over Cell-Free Massive MIMO-enabled Mobile Edge Network: A Deep Reinforcement Learning-based Approach. *2021 International Conference On Artificial Intelligence In Information And Communication (ICAIIC)*. pp. 344-346 (2021)

[90] Zhao, R., Wang, X., Xia, J. & Fan, L. Deep reinforcement learning based mobile edge computing for intelligent Internet of Things. *Physical Communication*. **43** pp. 101184 (2020)

[91] Chen, M., Liu, W., Wang, T., Liu, A. & Zeng, Z. Edge intelligence computing for mobile augmented reality with deep reinforcement learning approach. *Computer Networks.* **195** pp. 108186 (2021)

[92] Chen, C., Zhang, Y., Wang, Z., Wan, S. & Pei, Q. Distributed computation offloading method based on deep reinforcement learning in ICV. *Applied Soft Computing.* **103** pp. 107108 (2021)

[93] Dampahalage, D., Manosha, K., Rajatheva, N. & Latva-Aho, M. Supervised Learning Based Sparse Channel Estimation For RIS Aided Communications. *ICASSP 2022-2022 IEEE International Conference On Acoustics, Speech And Signal Processing (ICASSP).* pp. 8827-8831 (2022)

[94] Kwon, H. Bundle-based channel estimation for 5g nr with supervised learning. *GLOBECOM 2020-2020 IEEE Global Communications Conference.* pp. 1-6 (2020)

[95] Bai, Q., Wang, J., Zhang, Y. & Song, J. Deep learning-based channel estimation algorithm over time selective fading channels. *IEEE Transactions On Cognitive Communications And Networking.* **6**, 125-134 (2019)

[96] Ding, Y. & Kwon, H. Doppler spread estimation for 5G NR with supervised learning. *GLOBECOM 2020-2020 IEEE Global Communications Conference.* pp. 1-7 (2020)

[97] He, H., Wang, R., Jin, W., Jin, S., Wen, C. & Li, G. Beamspace channel estimation for wideband millimeter-wave MIMO: A model-driven unsupervised learning approach. *IEEE Transactions On Wireless Communications.* (2022)

[98] Jung, K. & Wang, H. Pilotless channel estimation scheme using clustering-based unsupervised learning. *2018 15th International Symposium On Wireless Communication Systems (ISWCS).* pp. 1-5 (2018)

[99] Chang, Z., Lei, L., Zhou, Z., Mao, S. & Ristaniemi, T. Learn to cache: Machine learning for network edge caching in the big data era. *IEEE Wireless Communications.* **25**, 28-35 (2018)

[100] Mohammed, L., Anpalagan, A., Khwaja, A. & Jaseemuddin, M. Performance of cache placement using supervised learning techniques in mobile edge networks. *IET Networks.* **10**, 304-321 (2021)

[101] Zhang, Z. & Tao, M. Deep learning for wireless coded caching with unknown and time-variant content popularity. *IEEE Transactions On Wireless Communications.* **20**, 1152-1163 (2020)

[102] Chen, J., Gong, Z., Wang, W., Wang, C., Xu, Z., Lv, J., Li, X., Wu, K. & Liu, W. Adversarial caching training: Unsupervised inductive network representation learning on large-scale graphs. *IEEE Transactions On Neural Networks And Learning Systems.* (2021)

[103] Lei, F., Dai, Q., Cai, J., Zhao, H., Liu, X. & Liu, Y. A proactive caching strategy based on deep Learning in EPC of 5G. *International Conference On Brain Inspired Cognitive Systems.* pp. 738-747 (2018)

[104] Zhong, C., Gursoy, M. & Velipasalar, S. Deep reinforcement learning-based edge caching in wireless networks. *IEEE Transactions On Cognitive Communications And Networking.* **6**, 48-61 (2020)

[105] Jiang, W., Feng, G., Qin, S. & Liu, Y. Multi-agent reinforcement learning based cooperative content caching for mobile edge networks. *IEEE Access.* **7** pp. 61856-61867 (2019)

[106] Xu, Q., Su, Z. & Lu, R. Game theory and reinforcement learning based secure edge caching in mobile social networks. *IEEE Transactions On Information Forensics And Security.* **15** pp. 3415-3429 (2020)

[107] Bentz, W., Dhanjal, S. & Panagou, D. Unsupervised learning of assistive camera views by an aerial co-robot in augmented reality multitasking environments. *2019 International Conference On Robotics And Automation (ICRA).* pp. 3003-3009 (2019)

[108] Maray, M., Mustafa, E. & Shuja, J. Wireless Power Assisted Computation Offloading in Mobile Edge Computing: A Deep Reinforcement Learning Approach. *HUMAN-CENTRIC COMPUTING AND INFORMATION SCIENCES.* **14** (2024)

[109] Li, L., Li, X., Yang, S., Ding, S., Jolfaei, A. & Zheng, X. Unsupervised-learning-based continuous depth and motion estimation with monocular endoscopy for virtual reality minimally invasive surgery. *IEEE Transactions On Industrial Informatics.* **17**, 3920-3928 (2020)

[110] Magdy, R., Rashad, S., Hany, S., Tarek, M., Hassan, M. & Mohammed, A. Deep reinforcement learning approach for augmented reality games. *2021 International Mobile, Intelligent, And Ubiquitous Computing Conference (MIUCC).* pp. 330-336 (2021)

[111] Jang, J., Park, J. & Yang, H. Supervised-Learning-Based Resource Allocation in Wireless Networks. *2020 International Conference On Information And Communication Technology Convergence (ICTC).* pp. 1022-1024 (2020)

[112] Chen, T., Zhang, X., You, M., Zheng, G. & Lambotharan, S. A GNN-Based Supervised Learning Framework for Resource Allocation in Wireless IoT Networks. *IEEE Internet Of Things Journal.* **9**, 1712-1724 (2021)

[113] Niu, J., Zhang, S., Chi, K., Shen, G. & Gao, W. Deep learning for online computation offloading and resource allocation in NOMA. *Computer Networks.* **216** pp. 109238 (2022)

[114] Chen, W., Shen, G., Chi, K., Zhang, S. & Chen, X. DRL based partial offloading for maximizing sum computation rate of FDMA-based wireless powered mobile edge computing. *Computer Networks.* **214** pp. 109158 (2022)

[115] Galkin, B., Amer, R., Fonseca, E. & DaSilva, L. Intelligent base station association for UAV cellular users: A supervised learning approach. *2020 IEEE 3rd 5G World Forum (5GWF).* pp. 383-388 (2020)

[116] Afifi, G. & Gadallah, Y. Autonomous 3-D UAV localization using cellular networks: deep supervised learning versus reinforcement learning approaches. *IEEE Access.* **9** pp. 155234-155248 (2021)

[117] Shehzad, M., Hassan, S., Mahmood, A. & Gidlund, M. On the association of small cell base stations with UAVs using unsupervised learning. *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring).* pp. 1-5 (2019)

[118] Faigl, J., Váňa, P., Pěnička, R. & Saska, M. Unsupervised learning-based flexible framework for surveillance planning with aerial vehicles. *Journal Of Field Robotics.* **36**, 270-301 (2019)

[119] Praveena, V. & Vijayaraj. Optimal deep reinforcement learning for intrusion detection in UAVs. *Cmc Computers.* **70**, 2639-2653 (2022)

[120] Kong, X., Duan, G., Hou, M., Shen, G., Wang, H., Yan, X. & Collotta, M. Deep Reinforcement Learning based Energy Efficient Edge Computing for Internet of Vehicles. *IEEE Transactions On Industrial Informatics.* (2022)

[121] Liu, X., Sun, C., Yau, K. & Wu, C. Joint Collaborative Big Spectrum Data Sensing and Reinforcement Learning Based Dynamic Spectrum Access for Cognitive Internet of Vehicles. *IEEE Transactions On Intelligent Transportation Systems.* (2022)

[122] Abu Al-Haija, Q. & Krichen, M. A Lightweight In-Vehicle Alcohol Detection Using Smart Sensing and Supervised Learning. *Computers.* **11**, 121 (2022)

[123] Ganguly, B., Dey, D. & Munshi, S. An Unsupervised Learning Approach for Road Anomaly Segmentation Using RGB-D Sensor for Advanced Driver Assistance System. *IEEE Transactions On Intelligent Transportation Systems.* (2022)

[124] Karimi, E., Chen, Y. & Akbari, B. Task offloading in vehicular edge computing networks via deep reinforcement learning. *Computer Communications.* **189** pp. 193-204 (2022)

[125] He, X., Lu, H., Du, M., Mao, Y. & Wang, K. QoE-based task offloading with deep reinforcement learning in edge-enabled Internet of Vehicles. *IEEE Transactions On Intelligent Transportation Systems.* **22**, 2252-2261 (2020)

[126] Wang, J., Ke, H., Liu, X. & Wang, H. Optimization for computational offloading in multi-access edge computing: A deep reinforcement learning scheme. *Computer Networks.* **204** pp. 108690 (2022)

[127] Abbas, Z., Ali, Z., Abbas, G., Jiao, L., Bilal, M., Suh, D. & Piran, M. Computational offloading in mobile edge with comprehensive and energy efficient cost function: a deep learning approach. *Sensors.* **21**, 3523 (2021)

[128] Mukherjee, M., Kumar, V., Lat, A., Guo, M., Matam, R. & Lv, Y. Distributed deep learning-based task offloading for UAV-enabled mobile edge computing. *IEEE INFOCOM 2020-IEEE Conference On Computer Communications Workshops (INFOCOM WKSHPS).* pp. 1208-1212 (2020)

[129] Irshad, A., Abbas, Z., Ali, Z., Abbas, G., Baker, T. & Al-Jumeily, D. Wireless powered mobile edge computing systems: Simultaneous time allocation and offloading policies. *Electronics.* **10**, 965 (2021)

[130] Yang, Q., Luo, X., Li, P., Miyazaki, T. & Wang, X. Computation offloading for fast CNN inference in edge computing. *Proceedings Of The Conference On Research In Adaptive And Convergent Systems.* pp. 101-106 (2019)

[131] Chai, R., Song, X. & Chen, Q. Joint task offloading, CNN layer scheduling, and resource allocation in cooperative computing system. *IEEE Systems Journal.* **14**, 5350-5361 (2020)

[132] Zeng, F., Tang, J., Liu, C., Deng, X. & Li, W. Task-Offloading Strategy Based on Performance Prediction in Vehicular Edge Computing. *Mathematics.* **10**, 1010 (2022)

[133] Zhao, M., Li, Y., Asif, S., Zhu, Y. & Tang, F. C-LSTM: CNN and LSTM Based Offloading Prediction Model in Mobile Edge Computing (MEC). *2022 IEEE 23rd International Conference On High Performance Switching*

And Routing (HPSR). pp. 245-251 (2022)

[134] Mustafa, E., Shuja, J., Bilal, K., Mustafa, S., Maqsood, T., Rehman, F. & Khan, A. Reinforcement learning for intelligent online computation offloading in wireless powered edge networks. *Cluster Computing*. **26**, 1053-1062 (2023)

[135] Miao, Y., Wu, G., Li, M., Ghoneim, A., Al-Rakhami, M. & Hossain, M. Intelligent task prediction and computation offloading based on mobile-edge cloud computing. *Future Generation Computer Systems*. **102** pp. 925-931 (2020)

[136] Xu, S. & Guo, C. Computation offloading in a cognitive vehicular networks with vehicular cloud computing and remote cloud computing. *Sensors*. **20**, 6820 (2020)

[137] Ma, K., Wang, Z., Tian, W., Chen, S. & Hanzo, L. Deep Learning for mmWave Beam-Management: State-of-the-Art, Opportunities and Challenges. *IEEE Wireless Communications*. (2022)

[138] Łaskawiec, S., Choraś, M., Kozik, R. & Varadarajan, V. Intelligent operator: Machine learning based decision support and explainer for human operators and service providers in the fog, cloud and edge networks. *Journal Of Information Security And Applications*. **56** pp. 102685 (2021)

[139] Miller, R., Banerjee, N. & Banerjee, S. Combining Real-World Constraints on User Behavior with Deep Neural Networks for Virtual Reality (VR) Biometrics. *2022 IEEE Conference On Virtual Reality And 3D User Interfaces (VR)*. pp. 409-418 (2022)

[140] Shen, Y., Shi, Y., Zhang, J. & Letaief, K. Graph neural networks for scalable radio resource management: Architecture design and theoretical analysis. *IEEE Journal On Selected Areas In Communications*. **39**, 101-115 (2020)

[141] Xiang, T. & Joe, I. A Reinforcement learning-based for Multi-user Task Offloading and Resource Allocation in MEC. **29**, 45-47 (2022)

[142] Chiroma, H., Abdulhamid, S., Hashem, I., Adewole, K., Ezugwu, A., Abubakar, S. & Shuib, L. Deep learning-based big data analytics for internet of vehicles: taxonomy, challenges, and research directions. *Mathematical Problems In Engineering*. **2021** (2021)

[143] Bahari, M., Saadatnejad, S., Rahimi, A., Shaverdikondori, M., Shahidzadeh, A., Moosavi-Dezfooli, S. & Alahi, A. Vehicle trajectory prediction works, but not everywhere. *Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition*. pp. 17123-17133

(2022)

[144] Balevi, E., Doshi, A. & Andrews, J. Massive MIMO channel estimation with an untrained deep neural network. *IEEE Transactions On Wireless Communications*. **19**, 2079-2090 (2020)

[145] Yang, Y., Gao, F., Ma, X. & Zhang, S. Deep learning-based channel estimation for doubly selective fading channels. *IEEE Access*. **7** pp. 36579-36589 (2019)

[146] Liu, S., Gao, Z., Zhang, J., Di Renzo, M. & Alouini, M. Deep denoising neural network assisted compressive channel estimation for mmWave intelligent reflecting surfaces. *IEEE Transactions On Vehicular Technology*. **69**, 9223-9228 (2020)

[147] Huang, H., Yang, J., Huang, H., Song, Y. & Gui, G. Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system. *IEEE Transactions On Vehicular Technology*. **67**, 8549-8560 (2018)

[148] Wang, Y., Hu, Y., Yang, Z., Saad, W., Wong, K. & Friderikos, V. Learning from Images: Proactive Caching with Parallel Convolutional Neural Networks. *IEEE Transactions On Mobile Computing*. (2022)

[149] Bajpai, R., Chakraborty, S. & Gupta, N. Adapting Deep Learning for Content Caching Frameworks in Device-to-Device Environments. *IEEE Open Journal Of The Communications Society*. (2022)

[150] Lai, S., Zhao, R., Wang, Y., Zhu, F. & Xia, J. Content popularity prediction for cache-enabled wireless B5G networks. *EURASIP Journal On Advances In Signal Processing*. **2021**, 1-16 (2021)

[151] Ale, L., Zhang, N., Wu, H., Chen, D. & Han, T. Online proactive caching in mobile edge computing using bidirectional deep recurrent neural network. *IEEE Internet Of Things Journal*. **6**, 5520-5530 (2019)

[152] Park, K., Kim, M., Choi, S. & Lee, J. Deep learning-based smart task assistance in wearable augmented reality. *Robotics And Computer-Integrated Manufacturing*. **63** pp. 101887 (2020)

[153] Ran, X., Chen, H., Zhu, X., Liu, Z. & Chen, J. Deepdecision: A mobile deep learning framework for edge video analytics. *IEEE INFOCOM 2018-IEEE Conference On Computer Communications*. pp. 1421-1429 (2018)

[154] Liu, L., Li, H. & Gruteser, M. Edge assisted real-time object detection for mobile augmented reality. *The 25th Annual International Conference On Mobile Computing And Networking*. pp. 1-16 (2019)

[155] Mirzaei, A. & Najafi Souha, A. Towards optimal configuration in MEC Neural networks: deep learning-based optimal resource allocation. *Wireless Personal Communications.* **121**, 221-243 (2021)

[156] Li, J. & Lv, T. Deep neural network based computational resource allocation for mobile edge computing. *2018 IEEE Globecom Workshops (GC Wkshps).* pp. 1-6 (2018)

[157] Jiang, F., Wang, K., Dong, L., Pan, C., Xu, W. & Yang, K. Deep-learning-based joint resource scheduling algorithms for hybrid MEC networks. *IEEE Internet Of Things Journal.* **7**, 6252-6265 (2019)

[158] Li, J., Gao, H., Lv, T. & Lu, Y. Deep reinforcement learning based computation offloading and resource allocation for MEC. *2018 IEEE Wireless Communications And Networking Conference (WCNC).* pp. 1-6 (2018)

[159] Alzaghir, A., Abdellah, A. & Koucheryav, A. Predicting Energy Consumption for UAV-Enabled MEC Using Machine Learning Algorithm. *Internet Of Things, Smart Spaces, And Next Generation Networks And Systems.* pp. 297-309 (2021)

[160] Jing, L., Jia, X., Lv, Y. & Wan, N. Maximizing the average secrecy rate for UAV-assisted MEC: A DRL method. *2021 IEEE 5th Advanced Information Technology, Electronic And Automation Control Conference (IAEAC).* **5** pp. 2514-2518 (2021)

[161] Xiao, K., Zhao, J., He, Y. & Yu, S. Trajectory prediction of UAV in smart city using recurrent neural networks. *ICC 2019-2019 IEEE International Conference On Communications (ICC).* pp. 1-6 (2019)

[162] Becker, S., Hug, R., Hübner, W., Arens, M. & Morris, B. Generating synthetic training data for deep learning-based uav trajectory prediction. *ArXiv Preprint ArXiv:2107.00422.* (2021)

[163] Nie, L., Ning, Z., Wang, X., Hu, X., Cheng, J. & Li, Y. Data-driven intrusion detection for intelligent Internet of vehicles: A deep convolutional neural network-based method. *IEEE Transactions On Network Science And Engineering.* **7**, 2219-2230 (2020)

[164] Liu, G., Xu, Y., He, Z., Rao, Y., Xia, J. & Fan, L. Deep learning-based channel prediction for edge computing networks toward intelligent connected vehicles. *IEEE Access.* **7** pp. 114487-114495 (2019)

[165] Qiao, G., Leng, S., Maharjan, S., Zhang, Y. & Ansari, N. Deep reinforcement learning for cooperative content caching in vehicular edge computing and networks. *IEEE Internet Of Things Journal.* **7**, 247-257 (2019)

[166] Du, H., Leng, S., Wu, F., Chen, X. & Mao, S. A new vehicular fog computing architecture for cooperative sensing of autonomous driving. *IEEE Access.* **8** pp. 10997-11006 (2020)

[167] Sadatdiynov, K., Cui, L., Zhang, L., Huang, J., Salloum, S. & Mahmud, M. A review of optimization methods for computation offloading in edge computing networks. *Digital Communications And Networks.* (2022)

[168] Zabihi, Z., Eftekhari Moghadam, A. & Rezvani, M. Reinforcement Learning Methods for Computation Offloading: A Systematic Review. *ACM Computing Surveys.* **56**, 1-41 (2023)

[169] Hortelano, D., Miguel, I., Barroso, R., Aguado, J., Merayo, N., Ruiz, L., Asensio, A., Masip-Bruin, X., Fernández, P., Lorenzo, R. & Others A comprehensive survey on reinforcement-learning-based computation offloading techniques in Edge Computing Systems. *Journal Of Network And Computer Applications.* **216** pp. 103669 (2023)

[170] Abdulazeez, D. & Askar, S. Offloading mechanisms based on reinforcement learning and deep learning algorithms in the fog computing environment: A comprehensive review. *IEEE Access.* (2023)

[171] Ma, H., Ji, B., Wu, H. & Xing, L. Video data offloading techniques in Mobile Edge Computing: A survey. *Physical Communication.* pp. 102261 (2023)

[172] Sharma, J., Kim, D., Lee, A. & Seo, D. On Differential Privacy-Based Framework for Enhancing User Data Privacy in Mobile Edge Computing Environment. *IEEE Access.* **9** pp. 38107-38118 (2021)

[173] Li, M., Yang, L. & Huang, J. A Deep Reinforcement Learning Approach for DNN Collaborative Inference in Mobile Edge Computing. *2023 10th International Conference On Electrical And Electronics Engineering (ICEEE).* pp. 248-254 (2023)