When 5G Meets Deep Learning: A Systematic Review

Guto Leoni Santos [1], Patricia Takako Endo [2],[*], Djamel Sadok [1] and Judith Kelner [1]

[1]   Centro de Informática, Universidade Federal de Pernambuco, Recife 50670-901, Brazil;
      gls4@cin.ufpe.br (G.L.S.); jamel@gprt.ufpe.br (D.S.); jk@gprt.ufpe.br (J.K.)
[2]   Programa de Pós-Graduação em Engenharia da Computação, Universidade de Pernambuco,
      Recife 50100-010, Brazil
[*]   Correspondence: patricia.endo@upe.br

Abstract: This last decade, the amount of data exchanged on the Internet increased by over a staggering factor of 100, and is expected to exceed well over the 500 exabytes by 2020. This phenomenon is mainly due to the evolution of high-speed broadband Internet and, more specifically, the popularization and wide spread use of smartphones and associated accessible data plans. Although 4G with its long-term evolution (LTE) technology is seen as a mature technology, there is continual improvement to its radio technology and architecture such as in the scope of the LTE Advanced standard, a major enhancement of LTE. However, for the long run, the next generation of telecommunication (5G) is considered and is gaining considerable momentum from both industry and researchers. In addition, with the deployment of the Internet of Things (IoT) applications, smart cities, vehicular networks, e-health systems, and Industry 4.0, a new plethora of 5G services has emerged with very diverging and technologically challenging design requirements. These include high mobile data volume per area, high number of devices connected per area, high data rates, longer battery life for low-power devices, and reduced end-to-end latency. Several technologies are being developed to meet these new requirements, and each of these technologies brings its own design issues and challenges. In this context, deep learning models could be seen as one of the main tools that can be used to process monitoring data and automate decisions. As these models are able to extract relevant features from raw data (images, texts, and other types of unstructured data), the integration between 5G and DL looks promising and one that requires exploring. As main contribution, this paper presents a systematic review about how DL is being applied to solve some 5G issues. Differently from the current literature, we examine data from the last decade and the works that address diverse 5G specific problems, such as physical medium state estimation, network traffic prediction, user device location prediction, self network management, among others. We also discuss the main research challenges when using deep learning models in 5G scenarios and identify several issues that deserve further consideration.

Keywords: the next generation of telecommunication (5G); deep learning; reinforcement learning; systematic review; cellular networks

1. Introduction

According to Cisco, the global Internet traffic will reach around 30 GB per capita by 2021, where more than 63% of this traffic is generated by wireless and mobile devices [1]. The new generation of mobile communication system (5G) will deal with a massive number of connected devices at base stations, a massive growth in the traffic volume, and a large range of applications with different features and requirements. The heterogeneity of devices and applications makes infrastructure management even more complex. For example, IoT devices require low-power connectivity,

trains moving at 300 KM/h need a high-speed mobile connection, users at their home need fiber-like broadband connectivity [2] whereas Industry 4.0 devices require ultra reliable low delay services. Several underlying technologies have been put forward in order to support the above. Examples of these include multiple-input multiple-output (MIMO), antenna beamforming [3], virtualized network functions (VNFs) [4], and the use of tailored and well provisioned network slices [5].

Some data based solutions can be used to manage 5G infrastructures. For instance, analysis of dynamic mobile traffic can be used to predict the user location, which benefits handover mechanisms [6]. Another example is the evaluation of historical physical channel data to predict the channel state information, which is a complex problem to address analytically [7]. Another example is the network slices allocation according to the user requirements, considering network status and the resources available [2]. All these examples are based on data analysis. Some examples are based on historical data analysis, used to predict some behavior, and others are based on the current state of the environment, used to help during decision making process. These type of problems can be addressed through machine learning techniques.

However, the conventional machine learning approaches are limited to process natural data in their raw form [8]. For many decades, constructing a machine learning system or a pattern-recognition system required a considerable expert domain knowledge and careful engineering to design a feature extractor. After this step, the raw data could be converted into a suitable representation to be used as input to the learning system [9].

In order to avoid the effort for creating a feature extractor or suffering possible mistakes in the development process, techniques that automatically discover representations from the raw data were developed. Over recent years, deep learning has outperformed conventional machine learning techniques in several domains such as computer vision, natural language processing, and genomics [10]. According to [9], deep learning methods *"are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transforms the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level"*. Therefore, several complex functions can be learned automatically through sufficient and successive transformations from raw data.

Similarly to many application domains, deep learning models can be used to address problems of infrastructure management in 5G networks, such as radio and compute resource allocation, channel state prediction, handover prediction, and so on. This paper presents a systematic review of the literature in order to identify how deep learning has been used to solve problems in 5G environments.

In [11], Ahmed et al. presented some works that applied deep learning and reinforcement learning to address the problem of resource allocation in wireless networks. Many problems and limitations related to resource allocation, such as throughput maximization, interference minimization, and energy efficiency were examined. While the survey presented in [11] focused on the resource allocation problem, in this paper, we offer a more general systematic review spanning the used different deep learning models applied to 5G networks. We also cover other problems present in 5G networks, that demand the use of different deep learning models.

Recently, in [12], Zhang et al. presented an extensive survey about the usage of deep learning in mobile wireless networks. Authors focused on how deep learning was used in mobile networks and potential applications, while identifying the crossover between these areas. Although it is very related to our work, Zhang et al. had a more general focus, addressing problems related to generic wireless networks such as mobility analysis, wireless sensor networks (WSN) localization, WSN data analysis, among others. Our systematic review is focused on 5G networks and their scenarios, applications, and problems. The deep learning models proposed in the analyzed works deal with specific cellular network problems such as channel state information, handover management, spectrum allocation. The scenarios addressed in the works that we select are also related with 5G networks and influence the deep learning-based solution proposed.

Differently, the existing work in the literature, our research identifies some of the main 5G problems addressed by deep learning, highlights the specific types of suitable deep learning models adopted in this context, and delineates the major open challenges when 5G networks meet deep learning solutions.

This paper is structured as follows: Section 2 an overview of the methodology adopted to guide this literature review. The results of the review including descriptive and thematic analysis are presented in Section 3. The paper concludes with a summary of the key findings and contributions of the paper in Section 4.

2. Systematic Review

In this paper, we based our systematic review on the protocol established in [13] with the purpose of finding the works that addressed the usage of deep learning models in the 5G context. We describe the methodology steps in the following subsections.

2.1. Activity 1: Identify the Need for the Review

As discussed previously, both 5G and deep learning are technologies that have received considerable and increasing attention in recent years. Deep learning has become a reality nowadays due to the availability of powerful off-the-shelf hardware and the emergence of new processing processing units such as GPUs. The research community has taken this opportunity to create several public repositories of big data to use in the training and testing of the proposed intelligent models. 5G on the other hand, has a high market appeal as it promises to offer new advanced services that, up until now, no other networking technology was able to offer. 5G importance is boosted by the popularity and ubiquity of mobile, wearable, and IoT devices.

2.2. Activity 2: Define Research Questions

The main goal of this work is to answer the following research questions:

- RQ. 1: What are the main problems deep learning is being used to solve?
- RQ. 2: What are the main learning types used to solve 5G problems (supervised, unsupervised, and reinforcement)?
- RQ. 3: What are the main deep learning techniques used in 5G scenarios?
- RQ. 4: How the data used to train the deep learning models is being gathered or generated?
- RQ. 5: What are the main research outstanding challenges in 5G and deep learning field?

2.3. Activity 3: Define Search String

The search string used to identify relevant literature was: (5G and "deep learning"). It is important to limit the number of strings in order to keep the problem tractable and avoid cognitive overwhelming.

2.4. Activity 4: Define Sources of Research

We considered the following databases as the main sources for our research: IEEE Xplore (http://ieeexplore.ieee.org/Xplore/home.jsp), Science Direct (http://www.sciencedirect.com/), ACM Digital Library (http://dl.acm.org/), and Springer Library (https://link.springer.com/).

2.5. Activity 5: Define Criteria for Inclusion and Exclusion

With the purpose of limiting our scope to our main goal, we considered only papers published in conferences and journals between 2009 and 2019. A selected paper must discuss the use of deep learning in dealing with a 5G technological problem. Note that solutions based on traditional machine learning (shallow learning) approaches were discarded.

2.6. Activity 6: Identify Primary Studies

The search returned 3, 192, 161, and 116 papers (472 in total) from ACM Digital Library, Science Direct, Springer Library, and IEEE Xplore, respectively. We performed this search in early November 2019. After reading all the 472 abstracts and using the cited criteria for inclusion or exclusion, 60 papers were selected for the ultimate evaluation. However, after reading the 60 papers, two papers were discarded because they were considered as being out of scope of this research. Next, two others were eliminated. The first paper was discarded because it was incomplete, and the second one was removed due presenting several inconsistencies in its results. Therefore, a total of 56 papers were selected for the for ultimate data extraction and evaluation (see Table A1 in Appendix A).

2.7. Activity 7: Extract Relevant Information

After reading the 56 papers identified in Activity 6, the relevant information was extracted as it attempted to answer some of the research questions presented in the Activity 2.

2.8. Activity 8: Present an Overview of the Studies

An overview of all works will be presented in this activity (see Section 3), in order to classify and clarify the conducted works according to our research questions presented in Activity 2.

2.9. Activity 9: Present the Results of the Research Questions

Finally, an overview of the studies in deep learning as it is applied to 5G is produced. It will discuss our findings and address our research questions stated in Activity 2 (see Section 3).

3. Results

In this section, we present our answers for the research question formulated previously.

3.1. What are the Main Problems Deep Learning Is Being Used to Solve?

In order to answer RQ. 1, this subsection presents an overview of the papers found in the systematic review. We separated the papers according to the problem addressed as shown in Figure 1. The identified problems can be categorized in three main layers: physical medium, network, and application.

At the physical level of the OSI reference model, we detected papers that addressed problems related to channel state information (CSI) estimation, coding/decoding scheme representation, fault detection, device prediction location, self interference, beamforming definition, radio frequency characterization, multi user detection, and radio parameter definition. At the network level, the works addressed traffic prediction through deep learning models and anomaly detection. Research on resource allocation can be related to the physical or network level. Finally, at the application level, existing works proposed deep learning-based solutions for application characterization.

In the following subsections, we will describe the problems solved by deep learning models; further details about the learning and the deep learning types used in the models will be presented in Sections 3.2 and 3.3, respectively.
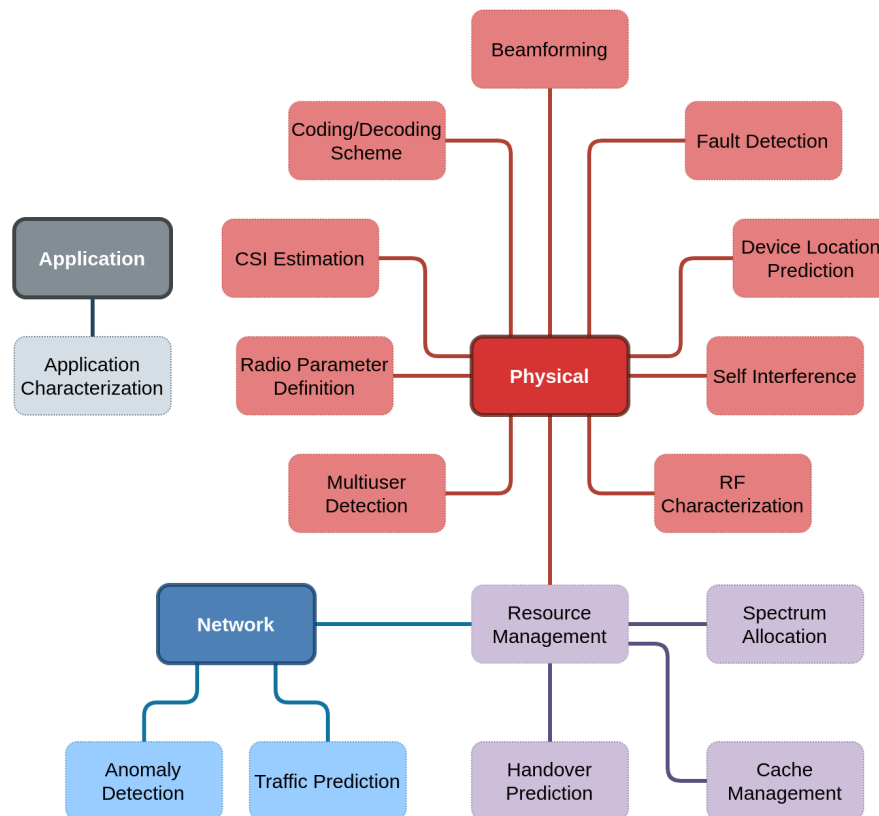
Figure 1. The problems related to 5G addressed in the works examined.

3.1.1. Channel State Information Estimation

CSI estimation is a common problem in wireless communication systems. It refers to the channel properties of a communication link [7]. In a simplified way, these properties describe how the signal will propagate from the transmitter to the receiver. Based on the CSI, the transmission can be adapted according to the current channel conditions, in order to improve the whole communication. CSI is an important factor in determining radio resource allocation, the type of modulation and coding schemes to use, etc.

Traditional CSI estimation techniques usually require high computation capability [14]. In addition, these techniques may not be suitable for 5G scenarios due to the complexity of the new scenarios and the presence of different technologies (e.g., massive MIMO, orthogonal frequency division multiplexing (OFDM), and millimeter-Wave), that impact the physical medium conditions [7]. Therefore, several authors have used deep learning models for CSI estimation. In our systematic review, we came across five papers related to CSI estimation with deep learning.

Three works proposed a deep learning-based solution focused on MIMO systems [15–17]. In MIMO systems both transmitter and receiver are equipped with an array of antennas. This is a very important technology for 5G, offering multiple orders of spectral and energy efficiency gains in comparison to LTE technologies [18]. Note that LTE uses MIMO but 5G takes this technology a notch further as it adopts massive antenna configurations in what is known as massive MIMO.

In [15], the authors adopted deep learning for decision-directed for channel estimation (DD-CE) in MIMO systems, to avoid the Doppler rate estimation. Authors considered vehicular channels, where the Doppler rate varies from one packet to another, making the CSI estimation difficult. Therefore, the deep learning model was used to learn and estimate the MIMO fading channels over different Doppler rates.

In [16], the authors proposed a combination of deep learning and superimposed code (SC) techniques for channel state CSI feedback. The main goal is to estimate downlink CSI and detect user data in the base stations.

In [17], Jiang et al. presented some evaluations for CSI estimation using deep learning models in three use cases. The first one focused on MIMO with multi users where the angular power spectrum (APS) information is estimated using deep learning models; and the two other scenarios were (a) static CSI estimation framework based on deep learning; and (b) a variant of the first scheme, but considering time variation, i.e., a deep learning model is proposed to estimate the CSI through time.

In [7], Luo et al. proposed an online CSI prediction taking into account relevant features that affect the CSI of a radio link, such as frequency band, user location, time, temperature, humidity, and weather.

In [19], a residual network was proposed for CSI estimation in filter bank multicarrier (FBMC) systems. The traditional CSI estimation and equalization and demapping module are replaced by deep learning model.

3.1.2. Coding/Decoding Scheme Representation

The generation of the information at the source and the reconstruction of such information at the receiver makes up the coding and decoding processes, respectively. However, due to the unstable nature of the channels, some disturbances and noise in the signal can cause data corruption [20]. Considering the 5G networks, where new technologies, such as MIMO, non-orthogonal multiple access (NOMA), mmWave will be deployed, the coding/decoding schemes must be adapted to work properly. These schemes need to characterize several phenomena that can impact the data transmission, such as signal diffraction, fading, path loss, and scattering.

We identified a total of seven works that addressed the coding/decoding schemes using deep learning models.

Three of these considered NOMA technology using deep learning models. In [21], the authors proposed a deep learning-based solution to parameterize the bit-to-symbol mapping and multi-user detection. Recall that as we are using non orthogonal modulation, multi-user detection becomes a cumbersome issue. In [22], the authors proposed a deep learning model to learn the coding/decoding process of MIMO-NOMA system in order to minimize the total mean square error of the users signals. In [23], the authors proposed a deep learning model to be used in sparse code multiple access (SCMA) system, which is a promising code-based NOMA technique, with the goal to minimize the bit error rate.

The authors in [24] considered multiuser single-input multiple-output (MU-SIMO) systems. A simple deep learning model was considered for joint multi user waveform design at the transmitter side, and non coherent signal detection at the receiver side. The main goal was to reduce the difference between the transmitted and received signals.

In [25], Kim et al. proposed a novel peak-to-average power ratio (PAPR) reduction scheme using deep learning of OFDM systems. The presence of large PAPR values is harmful to battery life as high peaks tend to draw high levels of energy from sometimes energy limited devices. The model proposed map and demap symbols on each subcarrier adaptively and both bit error rate (BER) and the PAPR of the OFDM system could be jointly minimized.

In [26], a deep learning based unified polar-low-density parity-check (LDPC) is proposed. The deep learning model was created to receive the observed symbols and an additional information introduced by the authors called "indicator section", and to output the signal decoded.

In [27], a coding mechanism under low latency constraints based on deep learning was proposed. The idea was to create a robust and adaptable mechanism for generic codes for future communications.

3.1.3. Fault Detection

Fault detection systems are very important to achieving ultra-reliable low latency communication (URLLC). For example, mission-critical industrial automation applications is a type of application that demands stringent timing and reliability guarantees for data collection, transmission, and processing [28]. Identifying faults is crucial to ensure low latency (since damaged equipment may increase the time transmission) and reliable communication (since point of failure may reduce the overall network performance). However, due to the device heterogeneity of 5G networks, identifying faults is a complex task that requires sophisticated techniques in order to automate such task.

In this systematic review, we found two papers that addressed fault detection in 5G scenarios using deep learning models.

In [29], a deep-learning-based schema was proposed to detect and locate antenna faults in mmWave systems. Firstly, the scheme detects the faults (using a simple neural network with a low cost), and then it locates where the fault occurred. Since the second step is a more complex task due to the high number of antennas present in a mmWave system, a more complex neural network was proposed.

In [30], Yu et al. covered fronthaul network faults. The model was designed to locate single-link faults in 5G optical fronthaul networks. The proposed model was able to identify faults and false alarms among alarm information considering single-link connections.

3.1.4. Device Location Prediction

Unlike traditional networks, in telecommunication networks, the nodes are characterized by a high mobility; and determining or estimating their mobility behavior is a complex task. Device location prediction has many applications, such as location-based services, mobile access control, mobile multimedia quality of service (QoS) provision, as well as the resource management for mobile computation and storage [31].

Considering urban scenarios, it is known that movement of people has a high degree of repetition, because they visit regular places in the city such as their own homes and places of work. These patterns can help to build services for specific places in order to increase user experience [32]. In addition, more detailed information about human mobility across the city can be collected using smartphones [33]. This information (combined with other data sources) can be used as input for models to estimate the device and consequently user location with high accuracy.

In this systematic review, three articles presented models to deal with device location prediction. Two works focused on device location prediction in mmWave systems [34,35]. In these systems, predicting the device location is a complex task due to the radiation reflected on most visible objects, which creates a rich multi path (interference) environment. In [34], a deep learning model was used to predict user location based on the radiation sent by the obstacles encountered. These carry latent information regarding their relative positions; while in [35], fingerprint historical data was used to estimate the device location over beamformed fingerprints.

In [36], the authors proposed a deep learning model to predict the device location in ultra-dense networks. Predicting the device location in this scenario is important because the deployment of small cells inevitably leads to more frequent handovers, making the mobility process more challenging. The model was used to predict user mobility and anticipate the handover preparation. The model was designed to estimate the future position of an user based on her/his historical data. If a handover is estimated as being eminent, deep learning model was able to determine the best base station to receive the user.

3.1.5. Anomaly Detection

Future 5G networks will lead with different types of devices over heterogeneous wireless networks with higher data rates, lower latency and lower power consumption. Autonomous management mechanisms will be needed to reduce the control and monitoring of these complex networks [37].

Anomaly detection systems are important to identify malicious network flows that may impact users and the network performance. However, developing these systems remains a considerable challenge due to the large data volume generated in 5G systems [38,39].

Four articles addressing the anomaly detection problem using deep learning in 5G were identified in this systematic review. In [38,40], the authors deal with cyber security defense systems in 5G networks, proposing the use of deep learning models that are capable of extracting features from network flows and the quick identification of cyber threats.

In [10,41], the authors proposed a deep learning-based solution to detect anomalies in the network traffic, considering two types of behavior as network anomalies: sleeping cells and soared traffic. Sleeping cells can happen due to failures in the antenna hardware or random access channel (RACH) failures due to RACH misconfiguration, while soared traffic can result in network congestion, where traffic increases but with relatively smaller throughput to satisfy the users' demand. Recall that RACH is the channel responsible for giving users radio resources so when RACH is not working properly we effectively have a sleepy cell with no transmission activity taking place.

3.1.6. Traffic Prediction

It is expected that Internet traffic will grow tenfold by 2027. This acts as a crucial anchor to create the new generation of cellular network architecture [42]. Predicting traffic for the next day, hour, or even the next minute can be used to optimize the available system resources, for example by reducing the energy consumption, applying opportunistic scheduling, or preventing problems in the infrastructure [42].

In this systematic review, we found eight works that addressed traffic prediction using deep learning.

The works presented in [43,44] proposed a deep learning-based solution to predict traffic for network slicing mechanisms. Note that 5G relies on the use of network slicing in order to accommodate different services and tenants while virtually isolating them. In [43], a proactive network slice mechanism was proposed and a deep learning model was used to predict the traffic with high accuracy. In [44], a mechanism named DeepCog was proposed with a similar purpose. DeepCog can forecast the capacity needed to allocate future traffic demands in network slices while minimizing service request violations and resource overprovisioning.

Three works considered both temporal and spatial dependence of cell traffic. In [6], the authors proposed a deep learning model to predict citywide traffic. The proposed model was able to capture the spatial dependency and two temporal dependencies: closeness and period. In [45], the authors proposed different deep learning models for mobile Internet traffic prediction. The authors used the different models to consider spatial and temporal aspects of the traffic. The maximum, average, and minimum traffic were predicted for the proposed models. In [46], the authors proposed a deep learning-based solution to allocate remote radio heads (RRHs) into baseband unit (BBU) pools in a cloud radio access network (C-RAN) architecture. The deep learning model was used to predict traffic demand of the RRHs considering the spatial and temporal aspects. The prediction was used to create RRH clusters and map them to BBU pools in order to maximize the average BBU capacity utility and minimize the overall deployment cost.

In [47], the authors considered traffic prediction in ultra-dense networks, which is a complicated scenario due to the presence of beamforming and massive MIMO technologies. A deep learning model was used to predict the traffic in order to detect if a congestion will take place and then take decisions to avoid/alleviate such congestion.

In [48], the authors addressed the benefits of cache offloading in small base stations considering the mobile edge computing (MEC). The offloading decision is based on the users' data rate, where the users with low data rates are offloaded first. Consequently, the authors proposed a deep learning model to predict the traffic data rate of the users in order to have a guide for the scheduling offloading mechanism.

3.1.7. Handover Prediction

The handover process ensures continuous data transfer when users are on the move between call towers. For that, the mobile management entity (MME) must update the base stations where the users are connected. This procedure is known as location update. The handover delay is one of the main problems in wireless networks [49]. Conventionally, a handover is carried out based on a predefined threshold of the Reference Signal Receiver Power (RSRP), the Reference Signal Receiver Quality (RSRQ), among other signal strength parameters [50]. Predicting the handover based on the nearby stations' parameters can be a fruitful strategy to avoid handover errors, temporary disconnections and improve user experience [49].

In this systematic review, we located two papers that addressed handover prediction. In [51], Khunteta et al. proposed a deep learning model to avoid handover failures. For that, the deep learning model was trained to detect if the handover will fail or be successful based on the historical signal condition data.

In [52], the handover prediction was tested to provide uninterrupted access to wireless services without compromising the expected QoS. The authors proposed both analytical and deep learning-based approaches to predict handover events in order to reduce the holistic cost.

3.1.8. Cache Optimization

In the last decade, multimedia data became dominant in mobile data traffic. This raised additional challenges in transporting the big volume of data from the content providers to the end users with high-rates and low latency. The main bottleneck point is the severe traffic congestion observed in the backhaul links, specially in 5G scenarios, where several small base stations will be scattered [53]. To mitigate this issue, the most popular content can be stored (cached) at the edge of the network (e.g., in the base stations) in order to free backhaul link usage [54]. However, finding the best strategy for the cache placement is a challenge. The best content to cache and the best location for storing this content are both decisions that can impact the cache scheme performance.

Two works addressed the cache placement problem in 5G environments using deep learning models. In [55], authors proposed a collaborative cache mechanism in multiple RRHs to multiple BBUs based on reinforcement learning. This approach was used because rule-based and metaheuristics methods suffer some limitations and fail to consider all environmental factors. Therefore, by using reinforcement learning, the best cache strategy can be selected in order to reduce the transmission latency from the remote cloud and the traffic load of backhaul.

In [56], the authors considered ultra-dense heterogeneous networks where the content cache is performed at small base stations. The goal is to minimize energy consumption and reduce the transmission delay, optimizing the whole cache placement process. Instead of using traditional optimization algorithms, a deep learning model was trained to learn the best cache strategy. This model reduces the computational complexity achieving a real time optimization.

3.1.9. Resource Allocation/Management

As the numbers of users, services, and resources increase, the management and orchestration complexity of resources also increase. The efficient usage of resources can be translated into cost reduction and avoid over/under resource dimensioning. Fortunately, under such a very dynamic and complex network environment, recent achievements in machine learning that interact with surrounding environments can provide effective way to address these problems [57].

Four papers addressed resource allocation in network slices using solutions based on deep learning [5,57–59]. A network slice is a very important technology for 5G since it will allow a network operator to offer a diverse set of tailored and isolated services over a shared physical infrastructure.

A deep learning-based solution was proposed in [58] to allocate slices in 5G networks. The authors proposed a metric called REVA that measures the amount of Physical Resource Blocks (PRBs) available to active bearers for each network slice, and a deep learning model was proposed to predict such metric.

Yan et al. proposed a framework that combined deep learning and reinforcement learning to resource scheduling and allocation [57]. The main goal was to minimize resource consumption at the same time guaranteeing the required performance isolation degree by a network slice. In [5], the authors proposed a framework for resource allocation in network slices and a deep learning model was used to predict the network status based on historical data. In [59], a model was proposed to predict the medium usage for network slices in 5G environments while meeting service level agreement (SLA) requirements.

Three papers proposed deep learning-based solutions to optimize the energy consumption in 5G networks [60–63]. The works proposed by [60,61] focused on NOMA systems. A framework was proposed in [60] to optimize energy consumption. A deep learning model is part of the framework and was used to map the input parameters (channel coefficients, the user demands, user power, and the transmission deadline) into an optimal scheduling scheme. In [61], a similar strategy was used, where a deep learning model was used to find the approximated optimal joint resource allocation strategy to minimize energy consumption. In [62], a deep learning model was used in the MME for user association taking into account the behavior of access points in the offloading scheme. In [63], the authors proposed a deep learning model to allocate carriers in multi-carrier power amplifier (MCPA) dynamically, taking into account the energy efficiency. The main idea was to minimize the total power consumption while finding the optimal carrier to MPCA allocation. To solve this problem, two approaches were used: convex relaxation and deep learning. The deep learning model was used to approximate the power consumption function formulated in the optimization problem, since it is a non-convex and non-continuous function.

In [64], the authors proposed a deep learning-based solution for downlink coordinated multi-point (CoMP) in 5G. The model receives physical layer measurements from the user equipment and "formulates a modified CoMP trigger function to enhance the downlink capacity" [64]. The output of the model is the decision to enable/disable the CoMP mechanism.

In [65], the authors proposed a deep learning model for smart communication systems with high density D2D mmWave environments using beamforming. The model selects the best relay node taking into account multiple reliability metrics in order to maximize the average system throughput. The authors in [11] also proposed a deep learning-based solution to maximize the network throughput considering resource allocation in multi-cell networks. A deep learning model was proposed to predict the resource allocation solution (taking as input the channel quality indicator and user location) without intensive computations.

3.1.10. Application Characterization

In cellular networks, self-organizing networks (SON) is a technology designed to plan, deploy, operate, and optimize mobile radio access networks in a simple, fast, and automated way. SON is a key technology for future cellular networks due to the potential of saving capital expenditure (CAPEX) and operational expenditure (OPEX). However, SON is not only about network performance but also QoS. A better planning of network resources can be translated into a better service quality and increasing revenues.

The authors in [66,67] presented a framework for self-optimization in 5G networks called APP-SON. It was designed to optimize some target network key performance indicators (KPIs) based on the mobile applications characteristics, by identifying similar application features and creating clusters using the Hungarian Algorithm Assisted Clustering (HAAC). The homogeneous application

characteristics of cells in a cluster are identified to prioritize target network KPIs in order to improve user quality of experience (QoE). This is achieved through cell engineering parameters adjustments. The deep learning model was used to establish cause effect between the cell engineering parameters and the network KPIs. For instance, Video application KPIs can be used to detect that this type of traffic occupies more than 90% of the total traffic, and thus adjust the cell engineering parameters to give priority to video traffic.

3.1.11. Other Problems

Some papers addressed problems which are not related to the ones previously listed. Thus, we will describe them separately.

The work presented in [68] applied a deep learning model to a massive MIMO system to solve the pilot contamination problem [69]. The authors highlighted that conventional approaches of pilot assignment are based on heuristics that are difficult to deploy in a real system due to high complexity. The model was used to learn the relationship between the users' location and the near-optimal pilot assignment with low computational complexity, and consequently could be used in real MIMO scenarios.

The self-interference problem was addressed in [70]. A digital cancellation scheme based on deep learning was proposed for full-duplex systems. The proposed model was able to discover the relationship between the signal sent through the channel and the self-interference signal received. The authors evaluated how the joint effects of non-linear distortion and linear multi-path channel impact the performance of digital cancellation using the deep learning model.

The authors in [71] represented the characterization of radio frequency (RF) power amplifiers (PAs) using deep learning. While in previous works they have considered only linear aspects of PA, the authors included non-linear aspects of PA taking into account memory aspects of deep learning models in [71]. They defined the map between the digital base station stimulus and the response of PA as a non-linear function. However, the conventional methods to solve this function require a designer to extract the interest parameters for each input (base station stimulus) manually. As a result, a deep learning model was proposed to represent this function, extracting the parameters automatically from measured base station stimulus and giving as output the PA response.

In [2], reinforcement learning was used to learn the optimal physical-layer control parameters of different scenarios. Authors proposed a self-driving radio, which learns the near-optimal control algorithm while taking int account the high-level design specifications provided by the network designer. A deep learning model was proposed to map the network specifications into physical-layer control instructions. This model was then used in the reinforcement learning algorithm to take decisions according to feedback from the environment.

In [72], the spectrum auction problem was addressed using deep learning. The idea was to allocate spectrum among unlicensed users taking into account the interests of the channel for the auction, and the interference suffered during communication as well as economic capability. A deep learning model was proposed for spectrum auction, and it receives as input three factors: the interference, experience, and economic ability; and gives as output a number between zero and one that determines whether the channel will be allocated for a user or not.

In [73], path scheduling in a multi path scenario was addressed using reinforcement learning. In these systems, the traffic is distributed across the different paths according to policies, packet traffic classes, and the performance of the available paths. Thus, reinforcement learning was used to learn from the network the best approach for scheduling packets across the different paths.

The security aspect of cooperative NOMA systems was considered in [74]. In cooperative NOMA, the user with a better channel condition acts as a relay between the source and a user experiencing poor channel conditions (user receiver). The security may be compromised in the presence of an eavesdropper in the network. Therefore, a deep learning model was proposed to find the optimal power allocation factor of a receiver in a communication system has the presence of an eavesdropper node.

The model input data are the channel realization while the output are the power allocation factor of the user with poor channel conditions.

In [75], authors considered the propagation prediction using deep learning models. Predicting the propagation characteristics accurately is needed for optimum cell design. Thus, the authors proposed a deep learning model to learn propagation loss from the map of a geographical area with high accuracy.

The authors in [76] considered the multiuser detection problem in an SCMA system. A deep learning model was used to mimic the message passing algorithm (MPA), which is the most popular approach to implement multiuser detection with low complexity. The deep learning model was designed to estimate the probability that a user is assigned into a resource block from a pool of resource blocks, taken the signal sent by the users as input.

In [3], an intelligent beamforming technique based on MIMO technology was proposed using reinforcement learning. The proposal builds a self-learning system to determine the phase shift and the amplitude of each antenna. The reinforcement learning algorithm can adapt the signal concentration based on the number of users located in a given area. If there are many users in a given small area, the solution may produce a more targeted signal for users located at that area. However, if users are spread out over a wide area, a signal with wide coverage will be sent to cover the entire area.

In [77], Tsai et al. proposed a reinforcement learning-based solution in order to choose the best configuration of uplink and downlink channels in dynamic time-division duplexing (TDD) systems. The main goal was to optimize the mean opinion score (MOS), which is a QoE metric. This metric has a direct relationship with the system throughput. The optimization problem was formulated as one that maximizes the MOS of the system by allocating uplink and downlik traffic for the time frames. Thus, a set of downlink and uplink configurations was defined by the authors and, for each frame, these configurations are chosen for each base station.

3.2. What Are the Main Types of Learning Techniques Used to Solve 5G Problems?

The works captured in this systematic review used three different learning techniques, as shown in Figure 2. The majority of the these works used supervised learning (fifty articles), followed by reinforcement learning (seven articles), and unsupervised learning (four articles only).
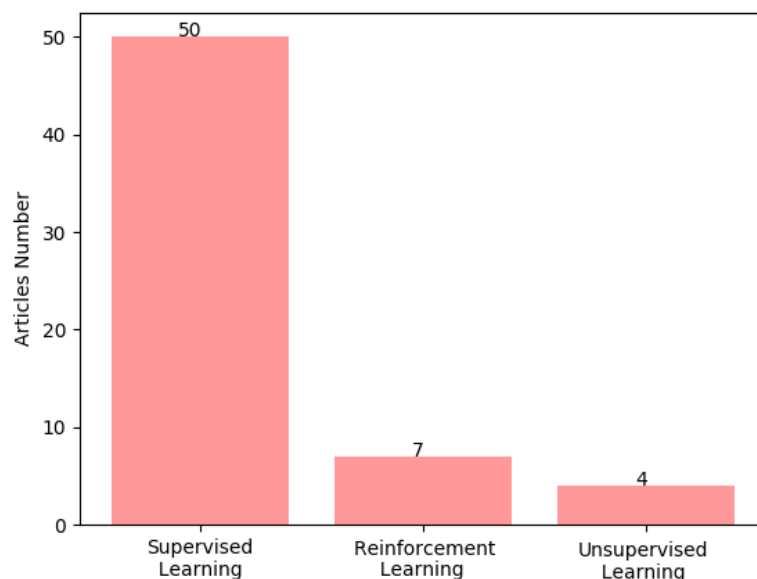

Figure 2. Most common learning type used in the deep learning models for 5G.

3.2.1. Supervised Learning

Although it is hard to find labeled datasets in 5G scenarios, most of the papers used the supervised learning approach. This approach is widely used for classification tasks (such as [78–81]) and regression problems (such as [82–85]), what are the most common problems addressed in the works found in this systematic review.

We classified the 50 articles that used supervised learning between classification and regression problems as shown in Table 1. We can see that 32 articles addressed classification problems in 5G scenarios whereas 19 articles dealt with regression models.

Table 1. Articles that used supervised learning in their deep learning models.

Problem Type	Number of Articles	References
Classification	32	[2,10,11,16,17,19,21–23,26,27,29,30,34,38,40,41,52,56,60–66,68,71,72,74–76]
Regression	19	[5–7,15,17,35,36,43–48,51,57–59,67,70]

3.2.2. Reinforcement Learning

Reinforcement learning has received a lot of attention in the last years. This paradigm is based on trial and error, where software agents learn a behavior that optimizes the reward observing the consequences of their actions [86]. The works we reviewed addressed different problems while taking into account context information and solving optimization problems. For instance, authors in [3] used reinforcement learning to determine phase shift and amplitude of each antenna element with the purpose to optimize the aggregated throughput of the antennas. In [62], authors used reinforcement learning to improve the URLLC energy efficiency and delay tolerant services through resource allocation. In [73], the authors also considered a URLLC service but this time they worked on optimizing packet scheduling of a multipath protocol using reinforcement learning. In [57], the authors adopted reinforcement learning for network slicing in RAN in an attempt to optimize resource utilization. To handle the cache allocation problem in multiple RRHs and multiple BBU pools, the authors in [55] used reinforcement learning to maximize the cache hit rate and maximize the cache capacity. In [77], reinforcement learning was used to configure indoor small cell networks in order to optimize opinion score (MOS) and user QoE. Finally, in [2], reinforcement learning was used to select radio parameters and optimize different metrics according with the scenario addressed.

3.2.3. Unsupervised Learning

We examined four articles that used unsupervised learning to train the models proposed. In [61], the authors proposed a hybrid approach with both supervised and unsupervised learning to train the model with the purpose to determine an approximate solution for optimal joint resource allocation strategy and energy consumption. The authors in [30] also used a hybrid learning approach, combining supervised and unsupervised learning to train the model in order to identify faults and false alarms among alarm information considering single link connections. In [25], the authors trained a deep learning model through unsupervised learning to map constellation mapping and demapping of symbols on each subcarrier in an OFDM system, while minimizing the BER. In [24], an unsupervised deep learning model was proposed to represent a MU-SIMO system. Its main purpose was to reduce the difference between the signal transmitted and the signal received.

3.3. What Are the Main Deep Learning Techniques Used in 5G Scenarios?

Figure 3 shows the common deep learning techniques used to address 5G problems in the literature. Traditional neural networks with fully connected layers is the deep learning technique that most appears in the works (reaching 24 articles), followed by long short-term memory (LSTM) (with 14 articles), and convolutional neural network (CNN) (adopted by only 9 articles).
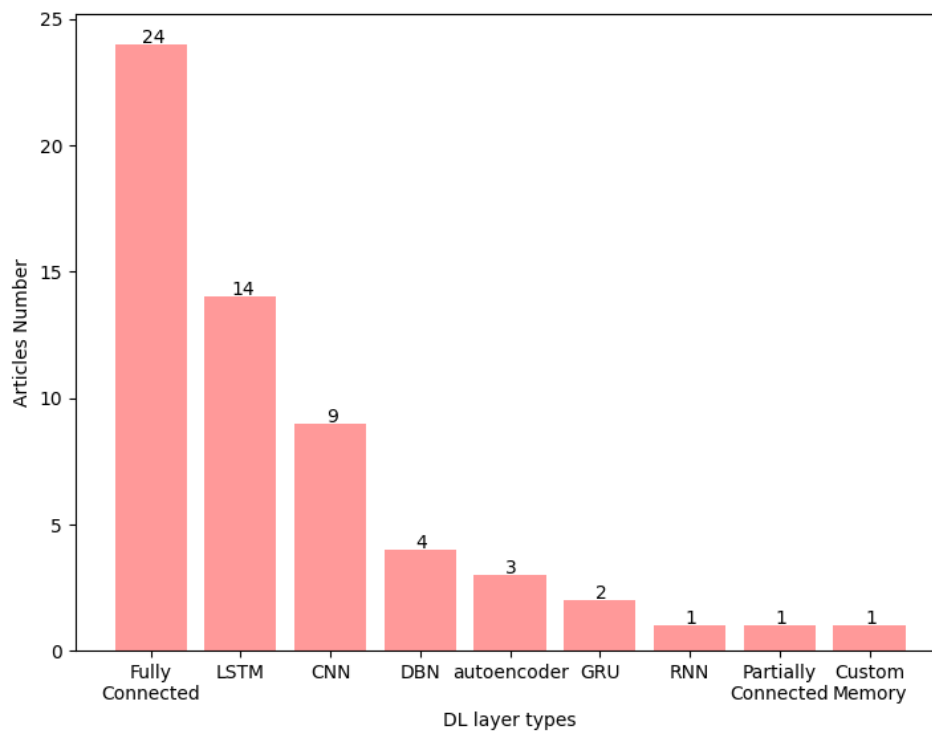
Figure 3. Most common deep learning techniques for 5G.

3.3.1. Fully Connected Models

Most of the works that used fully connected layers addressed problems related to the physical medium in 5G systems [2,11,15–17,21,22,24,26,56,60,62–65,68,72,74,76]. This can be justified because physical information usually can be structured (e.g., CSI, channel quality indicator (CQI), radio condition information, etc.). In addition, these works did not consider more complex data, such as historical information. It is understandable that the 5G physical layer receives such attention. It is the scene of a number of new technologies such as mmWave, MIMO and antenna beamforming. These are very challenging technologies that require real time fine tuning.

However, although fully connected layers were not designed to deal with sequential data, some works found in this systematic review proposed models based on time series. In [10,41], the authors considered real data of cellular networks such as Internet usage, SMS, and calls. Although the dataset has spatio-temporal characteristics, the authors extracted features to compose a new input for the deep learning model. In [52], the authors proposed a fully connected model to deal with user coordinate location data. In this work both fully connected and LSTM models were proposed for comparison and the temporal aspect of dataset was maintained. In [66], the authors adopted a dataset composed of historical data records for urban and rural areas. Unfortunately, the paper did not provide more details about the data used, but a deep learning model composed of fully connected layers was used to process this data.

In [73], a fully connected model was used with a reinforcement learning algorithm. In this work, the open source public Mininet simulator was used to create a network topology (the environment) in order to train the agent. Subsequently, the deep learning model was used to chose the best action according with the environment.

3.3.2. Recurrent Neural Networks

As highlighted in [9], a recurrent neural network (RNN) is able to deal with sequential data, such as time series, speech and language. It is due to its capacity for, given an element in a sequence, storing information of past elements. Therefore, one work used RNN [17] and several others used RNN variations (such as LSTM [87–90]) to deal with sequential data.

In [70], Zhang et al. proposed a digital cancellation scheme to eliminate linear and non-linear interference based on deep learning. The deep learning model receives a signal and the custom loss function represents the residual interference between the real and estimated self-interference signal. This model was based on RNN but with a custom memory unit.

In [17], authors used data from channel estimations using the ray tracing propagation software. The data was processed using traditional RNN layers to capture the time-varying nature of CSI. Similarly, several works adopted deep learning models with LSTM layers. This can be justified as LSTM is widely used in the literature to process sequential data.

The authors in [45,46] used the same dataset to train their models (Telecom Italia, see the Section 3.4). In [46], a multivariate LSTM model was proposed to learn the temporal and spatial correlation among the base station traffic and make an accurate forecast. In [45], an LSTM model was proposed to extract temporal features of mobile Internet traffic and predict Internet flows for cellular networks.

In [52], an LSTM model was suggested to deal with another open dataset in order to predict handover. The dataset is composed of historical location of the users, and the model exploits the long-term dependencies and temporal correlation of data.

In [48], the authors proposed an LSTM model for handling historical data of traffic volume. The model was constructed to predict real time traffic of base stations in order to give relevant information to increase the accuracy of the offloading scheme proposed.

In [47], Zhou et al. also proposed an LSTM model to predict traffic in base stations in order to avoid flow congestion in 5G ultra dense networks. Uplink and downlink flows data were considered as input for the model. With the predicted data, it is possible to allocate more resources for uplink or downlink channels accordingly.

In [7], an LSTM model was proposed to make online CSI prediction. The model explored the temporal dependency of historical data of frequency band, location, time, temperature, humidity, and weather. The dataset was measured through experiments within a testbed.

In [58], a variation of LSTM called X-LSTM was proposed in order to predict a metric called REVA, which measures the amount of PRBs available in a network slice. X-LSTM is based on X-11, which is an interative process that decomposes the data into seasonal patterns. X-LSTM uses different LSTM models to evaluate different time scales of data. "*It filters out higher order temporal patterns and uses the residual to make additional predictions on data with a shorter time scale*" [58]. The input data of the model is the historical data of PRB measured through a testbed, where the REVA metric was calculated.

In [71], the authors represented the memory aspect PA using a biLSTM model. The authors established a bridge between the theoretical formalism of PA behavior and the characteristic of biLSTM models to consider both forward an backward temporal aspect of the input data (baseband measurements using a testbed).

In [35,36,51,59], the authors used LSTM to deal with sequential data generated through simulation. In [59], the LSTM model was used to predict if a new network slice can be allocated given the sequential data of allocated resources and channel conditions. In [51], the LSTM model was used to evaluate historical signal condition in order to classify event in either handover fail or success in advance. In [36], the developed LSTM model was applied to learn the users mobility pattern in order to predict their movement trends in the future based on historical trajectories. In [35], the authors used LSTM to predict position of users based on historical beamformed fingerprint data (considering the presence o buildings in a scenario generated through simulations).

The work presented in [26] proposed an LSTM model to represent the coding/decoding schema considering a hybrid approach to support polar codes. Unfortunately, the authors did not describe the data used to train their model.

In [27,43], gated recurrent unit (GRU) layers are considered to deal with sequential data. In [43], real ISP data is used to train the model. The authors used a testbed to create the dataset composed of GPON (ZTE C320) to demonstrate the fronthaul, while midhaul and backhaul are enabled by the MPLS feature of SDN switches. Details about the dataset used in [27] are not provided.

3.3.3. CNN

CNN models are created to deal with data that come from multiple arrays or multivariate arrays and extract relevant features from them. In other words, the convolution layer is applied to process data with different dimensionality: 1D for signals and sequences, 2D for images or audio spectrograms, and 3D for video or volumetric images [9]. As a result, this layer was typically used to deal with several types of data in the works found in this systematic review.

The works in [29,34,35,75], presented the input data for the CNN models as an image form in order to take advantage of the natural features of the convolutions applied by the CNN layers. Both temporal and geographical aspects were considered in the works presented in [6,44,45]. These are relevant plans since the metrics have different behavior according to the time of the day and the base station location. As a result, these works used CNN to take into consideration temporal and space aspects at the same time and extract relevant joint patterns. The works presented in [7] used CNN models and considered several aspects that affect the CSI as input for the models such as frequency band, location, time, temperature, humidity, and weather. The authors considered 1D and 2D convolutions in order to extract frequency representative vector from CSI information.

A separate work used a special architecture of CNN called ResNet [19]. This architecture was proposed to solve the notorious problem of a vanishing/exploding gradient. The main difference offered by the ResNet architecture is that a shortcut connection is added every two or three layers in order to skip the connections and reuse activation from a previous layer until the adjacent layer learns its weights. This architecture was used to process modulated frequency-domain sequence data for the purpose of channel estimation.

In addition to the LSTM and CNN models, the authors proposed a model named a temporal convolutional network (TCN) in [35]. Unlike the other models, the TCN architecture considers the temporal dependency in a more accurate way. The interested reader may find out more detailed information TCN by consulting [91].

In [26], besides describing a fully connected layers and an LSTM models, the authors also proposed a CNN model for use with LSTM to represent the coding/decoding schema as convolution functions.

3.3.4. DBN

Deep belief networks (DBNs) are attractive for problems with few labeled data and a large amount of unlabeled ones. This is mainly due to the fact that during the training process, unlabeled data are used for training the model and the labeled data are used for fine-tuning the entire network [92]. Therefore, this deep learning technique combines both supervised and unsupervised learning during the training process.

For instance, the works presented in [38,40] used a dataset composed of several network flows of computers infected with botnets. The DBN model was used to detect traffic anomalies.

Similarly, in [61], the authors proposed a DBN model where the dataset used consisted of the channel coefficients and the respective optimal downlink resource allocation solution.

In [30], another DBN model was trained using a hybrid approach (supervised and unsupervised) for fault location on optical fronthalls. The dataset used was taken from a real management system of a network operator, and consists of link faults events.

3.3.5. Autoencoder

Autoencoders networks can be trained to reconstruct their input as the output [8]. Internally, these networks have a hidden layer that describes the internal representation of input. This representation can be used to construct back the input, that is the output of these networks. Therefore, the works used autoencoder architecture to encode and decode signal transmitted through the physical medium. We found three works in this systematic review that used an autoencoder architecture [23,25,27].

3.3.6. Combining Models

Most of the examined works make use of only one deep learning technique, but we have seen that there are eight works that considered more than one technique and provided a combination with other(s).

For instance, the authors in [52] proposed the joint use of an LSTM and a fully connected model. The research in [57] combined LSTM with reinforcement learning, [75] proposed a solution combining a CNN model with a fully connected model, [62,73] combined a fully connected model with reinforcement learning. Finally a combination of LSTM with CNN was proposed in [7,45]. A hybrid model, generative adversarial network (GAN), combining both LSTM and CNN layers was adopted in [5].

Next we discuss how the datasets were used to train these deep learning models.

3.4. How the Data Used to Train the Deep Learning Models Was Gathered/Generated?

Deep learning models (both supervised and unsupervised) require datasets for their training and testing. However, acquiring a good dataset, in some cases, remains a considerable challenge.

The works we reviewed either used different datasets or created their own data using different techniques, as shown in the Table 2.

Table 2. Data source.

Data Source	Number of Articles	References
Generated through simulation	24	[2,15–17,25,34–36,47,51,55–57,59–64,68,72,73,75,77]
Real data (generated using prototypes or public dataset)	18	[5–7,10,30,38,40,41,43–46,48,52,58,66,67,71]
Synthetic (generated randomly)	4	[11,19,74,76]
Not described (the work did not provide information about the dataset used)	10	[3,21–24,26,27,29,65,70]

Most of works (more precisely 24 of them) used simulation to generate their dataset. This is often justified as the authors are unable to a suitable variety of available datasets focused on 5G, since this is a novel technology and is being slowly deployed since 2020. Nonetheless, as many as 18 works used actual datasets to train their models. Some works measured the data through experiments using their own platform, whereas other works used public datasets available across the Internet. Four papers generated synthetic datasets. They contained some parameters of the evaluated 5G environment that were randomly generated. Finally, 10 works did not describe the source of the data used to train the proposed models. This is a point of concern in our view, as it makes the reproducibility and verification of the results of these works very difficult if not impossible altogether.

Unfortunately, none of the the works that created their own datasets (through simulation, measurements, or generated synthetically) made the data available. As a result, future works cannot use them to train new deep learning models or even use their results for comparison. Indeed, the availability of datasets for cellular networks is usually restricted to researchers subject to non-disclosure agreements (NDAs) and contracts with telecommunication operators and other private companies as also confirmed by [93].

Therefore, in this section, we describe some of the few public datasets used in the works that we managed to verify during this systematic review. The idea is to provide a brief description of these datasets that may be used in new works based on deep learning and provide useful pointers to the reader on where to find these. Note that public 5G traces and datasets remain difficult to find and that most of the existing traces are relatively old and related to 4G technology.

3.4.1. Telecom Italia Big Challenge Dataset

The Telecom Italia dataset [93] was used by majority of works [6,10,41,45,46]. It was provided as apart of a *Big Data Challenge* and is composed of various open multi-source aggregations of telecommunications, weather, news, social networks and electricity data. In 2014, the data was collected from two Italian areas: the city of Milan and the Province of Trentino.

With regard to the Call Detail Records (CDRs) present in the dataset, Telecom Italia recorded the following activities: (i) data about SMS, (ii) data about incoming and outgoing calls, and (iii) data about the Internet traffic. A CDR is generated every time a user starts or terminates an Internet session, if a connection takes more than 15 min, or more than five MB is transferred during a user session.

Further, the Telecom Italia dataset also includes the Social Pulse dataset (The Social Pulse dataset is composed of geo-located tweets that were posted by users from Trentino and Milan between 1 November 2013 and 31 December 2013), and other data such as weather, electricity (only for the Tentrino region), and news. For more information about this dataset, please see [93].

It comes at no surprise that the Telecom Italia dataset was used in several papers found in this systematic review. In [45], the dataset was used to train a model for predicting the minimum, maximum, and average traffic (multitask learning) of the next hour based on the traffic of the current hour. In [6], the models were proposed to predict traffic in a city environment taking into account spatial and temporal aspects. The data was sliced using a sliding window scheme generating several samples according with the closeness and the periodicity. In [10,41], the dataset was used to train the model to detect anomalies and data for short messages (SMS), calls, and Internet usage were considered. The authors divided the dataset into samples of three-hour ranges (morning, from 6 to 9 a.m.; afternoon, from 11 to 2 p.m.; and evening, from 5 to 8 p.m.). Another work that used the dataset for traffic prediction was presented in [46]. Here the authors compiled the traffic volume from the covered areas of cells of the dataset, and then normalized to the [0, 1] range for the convenience of carrying the analysis.

3.4.2. CTU-13 Dataset

The CTU-13 dataset [94] was compiled in 2011, in CTU University, in the Czech Republic, and comprises real botnet, normal, and background traffic.

The dataset is composed of 30 captures (corresponding to different scenarios) for several botnets samples. In each scenario, a specific malware with different protocols is used. After the capture, the authors analyzed the flow data in order to create the labels. There are four types of flows in the dataset: background, botnet, command and control channels, and normal. However, the dataset is unbalanced. For example, for a given scenario, there are 114,077 flows, where 37 (0.03%) is botnet traffic, and 112,337 (98.47%) of normal traffic.

Two works found in this systematic review used this dataset [38,40] to train deep learning models for anomaly detection. The authors made two different training/testing data partition. In the first partition, the CTU dataset was divided between training (80%), and testing (20%), both containing samples of every botnet. In the second partition, the botnet flows were divided into training and testing, i.e., the botnet flows that are present in the training set were not present in testing set.

3.4.3. 4G LTE Dataset with Channel from University College Cork (UCC)

This next dataset is provided by UCC [95] and is composed of client-side cellular KPIs. The information was collected from two major Irish mobile operators for different mobility patterns (static, pedestrian, car, train, and tram). There are 135 traces in the dataset, and each trace has an average duration of 50 min and a throughput that varies from 0 to 173 Mbit/s at a granularity of one sample per second. An Android network monitoring application was used to capture several channel related KPIs, downlink and uplink throughput, context-related metrics, and also cell-related information.

In an attempt to supplement the actual measured dataset, another dataset was generated through simulation and is also provided as a 4G LTE Dataset. The popular open source public Ns-3 simulator was used to create this dataset. It includes one hundred users randomly scattered across a seven-cell cluster. The main purpose of this complementary dataset is to provide information about the base stations (not present in the real dataset). In addition, the code and context information are offered to allow other researchers to generate their own synthetic datasets.

Nonetheless, only one work found in this systematic review actually used this dataset [5]. The model proposed was trained considering as input historical network data such as donwlink bitrate, uplink bitrate, and network delay. After the training, the model is then able to predict these network performance parameters for the next 1 min time interval.

3.5. What Are the Most Common Scenarios Used to Evaluate the Integration between 5G and Deep Learning?

Evaluating the works found in this systematic review, we noted that most of the works (40 of them) considered a generic scenario in their evaluations, and that only 16 articles considered specific ones.

The urban environment tops the studies as the most common scenario presented in the works [6,34,41,44–46,48,66,67,72,75]. This is justifiable as urban scenarios are very dynamic, very challenging and heterogeneous, with the presence of different obstacles (persons, vehicles, and buildings). They reflect extreme conditions that could not be easily handled by the previous cellular generations and where 5G requires special solutions such as the use of milli-meter waves, advanced beamforming, NOMA, etc., to deliver its promises. Notably, efficient usage of the frequency spectrum and the high energy consumption are two big challenges present in these scenarios [96].

Two recent works [2,17] considered vehicular networks as use case to evaluate their solutions. These demonstrate the increased research and interest in the domain of autonomous and connected vehicles, where 5G networks play a important role, providing a low latency with high availability [97]. Vehicular networks present unprecedented challenges that are not present in traditional wireless networks, such as fast-varying wireless propagation channels and ever-changing network topology [98]. Therefore, many researchers see the use of deep learning as a promising venue to solve some the stringent 5G problems.

In [2], the authors considered two vehicular network configurations while varying the device battery capacity and the available bandwidth; and also scenarios with a smartphone transmission of high definition (720p) real-time video conferencing signals.

Three different scenarios were evaluated in [5], namely, a video medical consultation (full duplex two direction live stream uplink/downlink), a virtual treatment (propagating a single direction video live stream (downlink)) and a simple Data Submit (Single direction data exchange over the uplink).

Cellular networks can also differ in terms of device location such as when operating indoor or outdoor. Indoor environments (homes and offices) have different characteristics that outdoor ones (road intersections, squares, stadiums, etc.). Evaluations carried out in outdoor scenarios are more common in the works found in this systematic review. Additionally to the works that considered urban cities and vehicular network, the work presented in [35] considered the users location problem within the New York University campus. A hybrid setup was considered in [7]. Here two outdoor and two indoor scenarios were examined: The two outdoor scenarios were parking lots situated outside a building, while the two indoor scenarios were a workroom and an aisle inside a building.

3.6. What Are the Main Research Challenges in 5G and Deep Learning Field?

Unfortunately, the majority of the articles examined in this systematic review did not present challenges or plans for future works (29 articles). It was then difficult to identify opportunities for new researches.

From the works that present next steps for the research, we can highlight the following relevant research issues. Some works plan to evaluate their solution based on real system aspects or real datasets. Authors often cite the lack of real datasets and traces as a major drawback of their current work and resort to the use of datasets generated through simulations. Though not an error in itself, the use of synthetic data may limit the scope of the findings.

Generally, many studies point out that the complexity of the 5G scenario remains a challenge. In fact, mathematical models are more difficult to develop here which makes the use of deep learning techniques more attractive. However, although deep learning models are able to process a big variety of data and receive multi-variate input data, the solutions proposed are often simplified to achieve low computational complexity. In this line of thought, some studies plan to include and add more input parameters for their models. For example, some works plan to consider more realistic parameters about the physical medium in their systems, while others considered to add new parameters. The inclusion of these parameters can considerably increase the complexity of the scenarios to be addressed. In addition, the presence of more parameters have a direct impact on the system performance [11]. This is an issue when dealing with real-time systems as in the case of 5G. Furthermore, it is always important to determine the level of abstraction needed to study a problem. It is not always the case that more parameters and detailing of a model are guaranteed to bring more accurate results and insights. The price may be higher than the benefits.

An alternative technique can be in the use of reinforcement learning. This is known to adapt and react dynamically to the environment at hand. This paradigm does not require data for training the agent, but it needs to describe a reward function and to represent the environment so that the agent learns to take actions that optimize that reward. The problem can be that one cannot afford to let an agent take wrong decisions in an attempt to learn as these can be costly to the operation of the network. We find this kind of problem also present in other critical application domains such as medical applications were one cannot afford to the use of deep learning due to the the risk on human life it may generate.

A further challenge pointed out by [11] is to consider deep learning solutions in scenarios with massive connections. This is indeed seen as a considerable challenge due to the the presence of different mobility patterns and different wireless channel fading conditions for the thousands of users. More robust models are needed. This complex scenario can hardly be subject to the application of traditional models. Instead, deep learning models represent a powerful tool to handle the different mobility patterns (using new recurrent models obtained from historical data) and different wireless channels (for example by considering reinforcement learning for environment-based learning).

The use of deep learning can sometimes be hampered by the processing power and timeliness especially in the presence of massive numbers of devices as in the case Industry 4.0. Understandably, many papers identified as challenges and future works the need to improve the performance of their solutions. After all, the performance of the overall system is slightly dependent on that of the adopted deep learning model. To achieve such improvement, some works intend to make a fine-tuning in the solution proposed, while others plan to trim or compress their networks, and there are those who consider new deep learning models altogether, with appropriated type of layers.

Last but not least, we find it important to highlight the strong integration between IoT and 5G networks. Future IoT applications will require new performance requirements, such as massive connectivity, security, trustworthy, coverage of wireless communication, ultra-low latency, throughput, and ultra-reliable [99]. It is not a coincidence that most part of these requirements are part of the planned 5G services. The authors in [10,41] plan to evaluate their deep learning-based solutions

in IoT scenarios: in [41] the authors plan to consider aspects about security (anomaly detection), while in [10] the authors plan consider energy consumption.

3.7. Discussions

As presented in this systematic review, all the selected papers are very recent as most of them were published in the year 2019 (57.1%). The oldest paper we examined is from the year 2015. This reflects the novelty and hotness of the technologies 5G and deep learning, and of course their integration.

5G is a technology in development and is set to solve several limitations present in the previous generations of cellular communication systems. It offers services, so far limited, such as massive connectivity, security, trust, large coverage, ultra-low latency (in the range of 1 ms over the air interface), throughput, and ultra-reliability, (99.999% of availability). On the other side of the spectrum, deep learning has received a lot of attention in the last few years as it has surpassed several state-of-the-art solutions in several fields, such as computer vision, text recognition, robotics, etc. The many reviewed recent publications attest the benefits that 5G technology would enjoy by making use of deep learning advances.

For the purpose of illustration only, as commented in [11], resource allocation in real cellular wireless networks can be formulated and solved using tools from optimization theory. However, the solutions often used have a high computational complexity. deep learning models may be a surrogate solution keeping the same performance but with reduced computational complexity.

We also noted that many works (a total of 25 to be precise) were published in conferences with few pages (around six pages). We believe that they represent works in progress, as they only show initial results. It reinforces the general view that the the integration between 5G and deep learning is still an evolving area of interest with many advances and contributions expected soon.

By observing the different scenarios considered in the examined articles, they generally do not focus on a real application (30 out of 57 articles found). However, a project called Mobile and wireless communications Enablers for the Twenty-twenty Information Society (METIS) published a document that explained several 5G scenarios and their specific requirements [100]. Nine use cases are presented: gaming, marathon, media on demand, unnamed aerial vehicles, remote tactile interaction, e-health, ultra-low cost 5G network, remote car sensing and control, and forest industry on remote control. Each of these scenarios have different characteristics and different requirements regarding 5G networks. For instance, remote tactile iterations scenarios can be considered a critical application (e.g., remote surgeries) and demand ultra-low latency (not be greater than 2 ms) and high reliability (99.999%). On the other hand, in the marathon use case, the participants commonly use attached tracing devices. This scenario must handle thousands of users simultaneously requiring high signaling efficiency and user capacity. As result, we believe that in order to achieve high impact results, deep learning solutions need to be targeted towards addressing use cases with specific requirements instead of trying to deal with the more general picture. Planning deep learning models for dynamic scenarios can be a complex task, since deep learning models need to capture the patterns present in the dataset. Thus, if the data varies widely between scenarios, it can certainly impact the performance of the models. One approach that can be used to deal with this limitation is the use of reinforcement learning. As presented in Section 3.2, seven works considered this paradigm in their solutions. Indeed, this approach considers training software agents to react to the environment in order to maximize (or minimize) a metric of interest. This paradigm can be a good approach to train software agents to dynamically adapt according to changes in the environment, and thus meet the different requirements of the use cases presented above.

However, reinforcement learning requires an environment where the software agent needs to be inserted during their training. Simulators can be a good approach, due the low cost of implementation. For example, consider an agent trained to control physical medium parameters instead of having to manually set up these, e.g., by fine tuning rules and thresholds. After training, the agent must be placed in a scenario with greater fidelity for validation, for example a prototype that can represent a

real scenario. Finally, the reinforcement learning agent can be deployed in a software-driven solution in the real scenario. These steps are necessary to avoid the drawbacks to deploying a non trained agent within a real operating 5G network. This is a cost, operators cannot afford.

4. Final Considerations

This work presented a systematic review on the use of deep learning models to solve 5G-related problems. 5G stands to benefit from deep learning as reported in this review. Though these models remove some of the traditional modeling complexity, developers need to determine the right balance between performance and abstraction level. More detailed models are not necessarily more powerful and many times the added complexity cannot be justified.

The review has also shown that the used deep learning techniques range across a plethora of possibilities. A developer must carefully opt for the right strategy to a given problem. We also showed that many works developed hybrid approaches in an attempt to cover a whole problem. Deep learning techniques are often also combined in the case of 5G with optimization algorithms such as genetic algorithm among others to produce optimized solutions.

Establishing clear use cases is important to determine the scope of a problem and therefore the deep learning parameters applied to it. 5G is known to offer services that have different and sometimes conflicting requirements. Hence, a solution that works for a given scenario may not work for another one.

Deep learning techniques are known to be data based. The more data, the most testing and development can be done and consequently, the better models we can produce. Unfortunately, due to reasons of business privacy very limited datasets are available. This is in contrast to other research communities that offer several datasets for research as in the case of image processing for example. We therefore believe that the industry and scientific community must make a similar effort to create more recent and representative 5G datasets. The use of simulated, old, and synthetic data has major limitations and may have questionable results.

A major point of concern in the 5G and deep learning integration remains that of performance. As we are dealing with real-time problems, the adopted solutions must not only deliver the expected solution but they must do it at the right time. Two things emerge from this point. The first one is related to the scope of deep-learning applications. In this case, we need to be careful in using it for problems that require agile answers sometimes at the nanosecond level. A second approach would be to develop simpler or compressed models.

Overall, the use of deep learning in 5G has already produced many important contributions and one expects these to evolve even further in the near future despite the many limitations identified in this review.

Author Contributions: Conceptualization, G.L.S. and P.T.E.; methodology, G.L.S. and P.T.E.; validation, G.L.S. and P.T.E.; formal analysis, G.L.S. and P.T.E.; investigation, G.L.S. and P.T.E.; resources, G.L.S. and P.T.E.; writing-original draft preparation, G.L.S. and P.T.E.; writing-review and editing, G.L.S., P.T.E., J.K., and D.S.; visualization, G.L.S. and P.T.E.; supervision, J.K. and D.J. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The Table A1 presents a summary of all works covered in this systematic review. Each row is about a paper and describes briefly the type of DL used in the paper, the learning type, the data source, and the paper objective using DL.

Table A1. Summary of works found in this systematic review.

Article	Layer Type	Learning Type	Data Source	Paper Objective
[56]	fully connected	supervised	simulation	to use a deep learning approach to reduce the network energy consumption and the transmission delay via optimizing the placement of content in heterogeneous networks.
[61]	DBN	supervised	simulation	a deep learning model was used to find the approximated optimal joint resource allocation strategy to minimize the energy consumption
[76]	fully connected	supervised	synthetic	the paper proposed a deep learning model to multiuser detection problem in the scenario of SCMA
[25]	autoencoder	unsupervised	simulation	the paper proposed the use of autoencoders to reduce PAPR in OFDM techniques called PRNet. The model is used to map constellation mapping and demapping of symbols on each subcarrier in an OFDM system, while minimize BER
[19]	residual network	supervised	synthetic	a deep-learning model was proposed for CSI estimation in FBMC systems. The traditional CSI estimation and equalization and demapping module are replaced by deep learning model
[67]	not described	supervised	real data	the paper propose a solution for optimize the self-organization in LTE networks. The solution, called APP-SON, makes the optimization based on the applications characteristics
[70]	a memory with custom memory	supervised and unsupervised	not described	the work proposed a digital cancellation scheme eliminating linear and non-linear interference based on deep learning
[38]	DBN	supervised	real data	the paper proposed a deep learning-based solution for anomaly detection on 5G network flows
[26]	fully connected and LSTM	supervised	not described	the authors proposed a deep learning model for channel decoding. The model is based on polar and LDPC mechanisms for decode signals in the receiver devices
[59]	LSTM	supervised	simulation	the authors proposed a machine learning-based solution to predict the medium usage for network slices in 5G environments meeting some SLA requirements
[34]	CNN	supervised	simulation	the authors proposed a system to to convert the received millimeter wave radiation into the device's position using CNN

Table A1. *Cont.*

Article	Layer Type	Learning Type	Data Source	Paper Objective
[71]	biLSTM	supervised	real data	a BiLSTM model was used to represent the effects of non-linear PAs, which is a promising technology for 5G. The authors defined the map between the digital baseband stimuluses and the response as a non-linear function.
[6]	CNN	supervised	real data	the authors proposed a framework based on CNN models to predict traffic in a city environment taking into account spatial and temporal aspects
[21]	fully connected	supervised	not described	the authors proposed a deep learning scheme to represent a constellation-domain multiplexing at the transmitter. This scheme was used to parameterize the bit-to-symbol mapping as well as the symbol detector
[23]	autoencoder	supervised	not described	the paper proposes a deep learning model to learning automatically the codebook SCMA. The codebook is responsible to code the transmitted bits into multidimensional codewords. Thus, the model proposed maps the bits into a resource (codebook) after the transmission and decode the signal received into bits at the receiver
[51]	LSTM	supervised	simulation	the paper proposed deep learning based scheme to avoid handover failures based on early prediction. This scheme can be used to evaluate the signal condition and make the handover before a failure happen
[7]	CNN and LSTM	supervised	real data	the authors proposed an online framework to estimate CSI based on deep learning models called OCEAN. OCEAN is able to find CSI for a mobile device during a period ate a specific place
[3]	not described	deep learning and reinforcement learning	not described	the authors proposed a beamforming scheme based on deep reinforcement learning. The problem addressed was the beamforming performance in dynamic environments. Depends on the number of users concentrated in a area, the beamforming configuration is produce a more directed signal, on the other hand a signal with wide coverage is sent. The solution proposed is composed of three different models. The first one, is a model that generated synthetic user mobility patterns. The second model tries to response with a more appropriated antenna diagram (beamforming configuration). The third model evaluates the performance of results obtained by the models and returns a reward for the previous models. The authors did not make any experiments about the scheme proposed

Table A1. *Cont.*

Article	Layer Type	Learning Type	Data Source	Paper Objective
[15]	fully connected	supervised	simulation	the authors proposed a deep learning scheme for DD-CE in MIMO systems. The core part of DD-CE is the channel prediction, where the "current channel state is estimated base on the previous estimate and detected symbols". Deep learning can avoid the need of complex mathematical models for doppler rates estimation
[16]	fully connected	supervised	simulation	the authors combined deep learning and superimposed coding techniques for CSI feedback. In a traditional superimposed coding-based CSI feedback system, the main goal of a base station is to recover downlink CSI and detect user data
[63]	fully connected	supervised	simulation	the authors proposed an algorithm to allocate carrier in MCPA dynamically, taking into account the energy efficiency and the implementation complexity. The main idea is to minimize the total power consumption finding the optimal carrier to MPCA allocation. To solve this problem, two approaches were used: convex relaxation and deep learning
[29]	CNN	supervised	not described	the authors presented a deep learning model to fault detection and fault location in wireless communication systems through deep learning, focusing in mmWave systems
[44]	3D CNN	supervised	real data	the authors proposed a deep learning-based solution for allocation resources previously based on data analytic. The solution is called DeepCog, which receives as input measurement data of a specific network slice, make a prediction of network flow and allocate resources in data center to meet the demand
[17]	fully connected and RNN	supervised	simulation	the authors proposed a systematic review about CSI and then presented some evaluations using deep learning models. The solutions presented in the systematic review have a focus on "linear correlations such as sparse spatial steering vectors or frequency response, and Gauss-Markov time correlations"
[36]	LSTM	supervised	simulation	the authors proposed a deep learning-based algorithm for handover mechanism. The model is used to predict the user mobility and anticipate the handover preparation previously. The algorithm will estimate the future position of the an user based on its historical data

Table A1. *Cont.*

Article	Layer Type	Learning Type	Data Source	Paper Objective
[62]	fully connected	deep learning and reinforcement learning	simulation	the authors proposed a solution to improve the energy efficiency of user equipment in MEC environments in 5G. In the work, two different types of applications were considered: URLLC and high data rate delay tolerant applications. The solution uses a "digital twin" of the real network to train the neural network models
[11]	fully connected	supervised	synthetic (through genetic algorithm)	the authors proposed a deep learning model for resource allocation to maximize the network throughput by performing joint resource allocation (i.e., both power and channel). Firstly a review about deep learning techniques applied to wireless resource allocations problem was presented. After, a deep learning model was presented. This model takes as input the CQI and the location indicator (position between the user from the base stations) of users for all base stations and predicts the power and sub-band allocations
[68]	fully connected	supervised	simulation	the work proposed a pilot allocation scheme based on deep learning for massive MIMO systems. The model was used to learn the relationship between the users' location and the near-optimal pilot assignment with low computational complexity
[65]	fully connected	supervised	not described	the authors proposed a deep learning model for smart communication systems for highly density D2D mmWave environments using beamforming. The model can be used to predict the best relay for relaying data taking into account several reliability metrics for select the relay node (e.g., another device or a base station)
[64]	fully connected	supervised	simulation	the authors proposed a deep learning-based solution for downlink CoMP in 5G environments. The model receives as input some physical layer measurements from the connected user equipment and "formulates a modified CoMP trigger function to enhance the downlink capacity". The output of the model is the decision to enable/disable the CoMP mechanism
[22]	fully connected	supervised	not described	the authors proposed a deep learning-based scheme for precoding and SIC decoding for scheme for the MIMO-NOMA system

Table A1. *Cont.*

Article	Layer Type	Learning Type	Data Source	Paper Objective
[57]	LSTM	supervised and reinforcement learning	simulation	the authors proposed a framework to resource scheduling allocation based on deep learning and reinforcement learning. The main goal is to minimize the resource consumption at the same time guaranteeing the required performance isolation degree. A LSTM and reinforcement learning are used in cooperation to do this task. A LSTM model was used to predict the traffic based on the historical data.
[45]	LSTM, 3D CNN, and CNN+LSTM	supervised	real data	the authors proposed a multitask learning based on deep learning for predict data flow in 5G environments. The model is able to predict the minimum, maximum, and average traffic (multitask learn) of the next hour based on the traffic of the current hour.
[30]	DBN	unsupervised and supervised	real data	the authors proposed a DBN model for fault location in optical fronthaul networks. The model proposed identify faults and false alarms in alarm information considering single link connections
[41]	fully connected	supervised	real data	the paper proposed a deep learning model to detect anomalies in the network traffic, considering two types of behavior as network anomalies: sleeping cells and soared traffic.
[47]	LSTM	supervised	simulation	the authors proposed a deep learning model to predict traffic in base stations in order to avoid flow congestion in 5G ultra dense networks
[52]	fully connected and LSTM	supervised	real data	the authors proposed a analytical model for holistic handover cost and a deep learning model to handover prediction. The holistic handover cost model takes into account signaling overhead, latency, call dropping, and radio resource wastage
[48]	LSTM	supervised	real data	a system model that combine mobile edge computing and mobile data offloading was proposed in the paper. In order to improve the system performance, a deep learning model was proposed to predict the traffic and decide if the offloading can be performed on the base station
[55]	-	reinforcement learning	simulation	the authors proposed a network architecture that integrates MEC and C-RAN. In order to reduce the latency, a caching mechanism can be adopted in the MEC. Thus, reinforcement learning was used to maximize the cache hit rate the cache use

Table A1. *Cont.*

Article	Layer Type	Learning Type	Data Source	Paper Objective
[46]	LSTM	supervised	real data	the paper proposed a framework to cluster RRHs and map them into BBU pools using predicted data of mobile traffic. Firstly, the future traffic of the RRHs are estimated using a deep learning model based on the historical traffic data, then these RRHs are grouped according with their complementarity
[40]	DBN	supervised	real data	the paper proposes a deep learning-based approach to analyze network flows and detect network anomalies. This approach executes in a MEC in 5G networks. A system based on NFV and SDN was proposed to detect and react to anomalies in the network
[77]	-	reinforcement learning	simulation	the paper proposed two schemes based on Q-learning to choose the best downlink and uplink configuration in dynamic TDD systems. The main goal is to optimize the MOS, which is a QoE measure that correspond a better experience of users.
[35]	CNN, LSTM, and temporal convolutional network	supervised	simulation	the authors proposed a deep learning-based approach to predict the user position for mmwave systems based on beamformed fingerprint
[2]	LSTM	supervised and reinforcement learning	simulation	the authors deal with physical layer control problem. A reinforcement learning-based solution was used to learn the optimal physical-layer control parameters of different scenarios. The scheme proposed use reinforcement learning to choose the best configuration for the scenario. In the scheme proposed, a radio designer need to specify the network configuration that varies according with the scenario specification
[58]	X-LSTM	supervised	real data	the paper proposed models to predict the mount of PRBs available to allocate network slices in 5G networks
[66]	fully connected	supervised	real data	the authors proposed a algorithm to achieve self-optimization in LTE and 5G networks trough wireless analysis. The deep learning model is used to perform a regression to derive the relationship between the engineering parameters and the performance indicators
[10]	fully connected	supervised	real data	the paper proposed a deep learning-based solution to detect anomalies in 5G networks powered by MEC. The model detects sleeping cells events and soared traffic as anomalies
[60]	fully connected	supervised	simulation	the paper proposed a framework to optimize the energy consumption of NOMA systems in a resource allocation problem.

<div align="center">

Table A1. *Cont.*

</div>

Article	Layer Type	Learning Type	Data Source	Paper Objective
[72]	fully connected	supervised	simulation	the paper proposed an auction mechanism for spectrum sharing using deep learning models in order to improve the channel capacity
[73]	fully connected	supervised and reinforcement learning	simulation	the paper proposed a deep reinforcement learning mechanism for packet scheduler in multi-path networks.
[5]	Generative adversarial networks (GAN) with LSTM and CNN layers	supervised	real data	the paper proposed a deep learning-based framework for address the problem of the network slicing scheme for the mobile network. The deep learning model is used to predict network flow in other to make resource allocation
[27]	Autoencoder with Bi-GRU layers	supervised	not described	the paper proposed a deep learning-based solution for channel coding in low-latency scenarios. The idea was to create a robust and adaptable mechanism for generic codes for future communications
[74]	fully connected	supervised	synthetic	the paper proposed a deep learning model for physical layer security. The model was used to optimize the value of the power allocation factor in a secure communication system
[75]	CNN and fully connected	supervised	simulation	the paper proposed a radio propagation model based on deep learning. The model maps geographical area in the radio propagation (path loss)
[24]	partially and fully connected layers	unsupervised	not described	a deep learning model was proposed to represent a MU-SIMO system. The main purpose is to reduce the difference between the signal transmitted and the signal received
[43]	GRU	supervised	real data	the paper proposed a deep learning-based framework for traffic prediction in order to enable proactive adjustment in network slice

References

1. Cisco. Global—2021 Forecast Highlights. 2016. Available online: https://www.cisco.com/c/dam/m/en_us/solutions/service-provider/vni-forecast-highlights/pdf/Global_2021_Forecast_Highlights.pdf (accessed on 19 August 2020).
2. Joseph, S.; Misra, R.; Katti, S. Towards self-driving radios: Physical-layer control using deep reinforcement learning. In Proceedings of the 20th International Workshop on Mobile Computing Systems and Applications, Santa Cruz, CA, USA, 27–28 February 2019; pp. 69–74.
3. Maksymyuk, T.; Gazda, J.; Yaremko, O.; Nevinskiy, D. Deep Learning Based Massive MIMO Beamforming for 5G Mobile Network. In Proceedings of the 2018 IEEE 4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS), Lviv, Ukraine, 20–21 September 2018; pp. 241–244.

4.　Arteaga, C.H.T.; Anacona, F.B.; Ortega, K.T.T.; Rendon, O.M.C. A Scaling Mechanism for an Evolved Packet Core based on Network Functions Virtualization. *IEEE Trans. Netw. Serv. Manag.* **2019**, *17*, 779–792. [CrossRef]

5.　Gu, R.; Zhang, J. GANSlicing: A GAN-Based Software Defined Mobile Network Slicing Scheme for IoT Applications. In Proceedings of the 2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–7.

6.　Zhang, C.; Zhang, H.; Yuan, D.; Zhang, M. Citywide cellular traffic prediction based on densely connected convolutional neural networks. *IEEE Commun. Lett.* **2018**, *22*, 1656–1659. [CrossRef]

7.　Luo, C.; Ji, J.; Wang, Q.; Chen, X.; Li, P. Channel state information prediction for 5G wireless communications: A deep learning approach. *IEEE Trans. Netw. Sci. Eng.* **2018**, *7*, 227–236. [CrossRef]

8.　Goodfellow, I.; Bengio, Y.; Courville, A.; Bengio, Y. *Deep Learning*; MIT Press: Cambridge, UK, 2016; Volume 1.

9.　LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [CrossRef]

10.　Hussain, B.; Du, Q.; Zhang, S.; Imran, A.; Imran, M.A. Mobile Edge Computing-Based Data-Driven Deep Learning Framework for Anomaly Detection. *IEEE Access* **2019**, *7*, 137656–137667. [CrossRef]

11.　Ahmed, K.I.; Tabassum, H.; Hossain, E. Deep learning for radio resource allocation in multi-cell networks. *IEEE Netw.* **2019**, *33*, 188–195. [CrossRef]

12.　Zhang, C.; Patras, P.; Haddadi, H. Deep learning in mobile and wireless networking: A survey. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2224–2287. [CrossRef]

13.　Coutinho, E.F.; de Carvalho Sousa, F.R.; Rego, P.A.L.; Gomes, D.G.; de Souza, J.N. Elasticity in cloud computing: A survey. *Ann. Telecommun.* **2015**, *70*, 289–309. [CrossRef]

14.　Caire, G.; Jindal, N.; Kobayashi, M.; Ravindran, N. Multiuser MIMO achievable rates with downlink training and channel state feedback. *IEEE Trans. Inf. Theory* **2010**, *56*, 2845–2866. [CrossRef]

15.　Mehrabi, M.; Mohammadkarimi, M.; Ardakani, M.; Jing, Y. Decision Directed Channel Estimation Based on Deep Neural Network k-Step Predictor for MIMO Communications in 5G. *IEEE J. Sel. Areas Commun.* **2019**, *37*, 2443–2456. [CrossRef]

16.　Qing, C.; Cai, B.; Yang, Q.; Wang, J.; Huang, C. Deep learning for CSI feedback based on superimposed coding. *IEEE Access* **2019**, *7*, 93723–93733. [CrossRef]

17.　Jiang, Z.; Chen, S.; Molisch, A.F.; Vannithamby, R.; Zhou, S.; Niu, Z. Exploiting wireless channel state information structures beyond linear correlations: A deep learning approach. *IEEE Commun. Mag.* **2019**, *57*, 28–34. [CrossRef]

18.　Prasad, K.S.V.; Hossain, E.; Bhargava, V.K. Energy efficiency in massive MIMO-based 5G networks: Opportunities and challenges. *IEEE Wirel. Commun.* **2017**, *24*, 86–94. [CrossRef]

19.　Cheng, X.; Liu, D.; Zhu, Z.; Shi, W.; Li, Y. A ResNet-DNN based channel estimation and equalization scheme in FBMC/OQAM systems. In Proceedings of the 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP), Hangzhou, China, 18–20 October 2018; pp. 1–5.

20.　Ez-Zazi, I.; Arioua, M.; El Oualkadi, A.; Lorenz, P. A hybrid adaptive coding and decoding scheme for multi-hop wireless sensor networks. *Wirel. Pers. Commun.* **2017**, *94*, 3017–3033. [CrossRef]

21.　Jiang, L.; Li, X.; Ye, N.; Wang, A. Deep Learning-Aided Constellation Design for Downlink NOMA. In Proceedings of the 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, 24–28 June 2019; pp. 1879–1883.

22.　Kang, J.M.; Kim, I.M.; Chun, C.J. Deep Learning-Based MIMO-NOMA With Imperfect SIC Decoding. *IEEE Syst. J.* **2019**. [CrossRef]

23.　Kim, M.; Kim, N.I.; Lee, W.; Cho, D.H. Deep learning-aided SCMA. *IEEE Commun. Lett.* **2018**, *22*, 720–723. [CrossRef]

24.　Xue, S.; Ma, Y.; Yi, N.; Tafazolli, R. Unsupervised deep learning for MU-SIMO joint transmitter and noncoherent receiver design. *IEEE Wirel. Commun. Lett.* **2018**, *8*, 177–180. [CrossRef]

25.　Kim, M.; Lee, W.; Cho, D.H. A novel PAPR reduction scheme for OFDM system based on deep learning. *IEEE Commun. Lett.* **2017**, *22*, 510–513. [CrossRef]

26.　Wang, Y.; Zhang, Z.; Zhang, S.; Cao, S.; Xu, S. A unified deep learning based polar-LDPC decoder for 5G communication systems. In Proceedings of the 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP), Hangzhou, China, 18–20 October 2018; pp. 1–6.

27. Jiang, Y.; Kim, H.; Asnani, H.; Kannan, S.; Oh, S.; Viswanath, P. Learn codes: Inventing low-latency codes via recurrent neural networks. In Proceedings of the 2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019; pp. 1–7.

28. Hu, P.; Zhang, J. 5G Enabled Fault Detection and Diagnostics: How Do We Achieve Efficiency? *IEEE Internet Things J.* **2020**, *7*, 3267–3281. [CrossRef]

29. Chen, K.; Wang, W.; Chen, X.; Yin, H. Deep Learning Based Antenna Array Fault Detection. In Proceedings of the 2019 IEEE 89th Vehicular Technology Conference (VTC2019), Honolulu, HI, USA, 22–25 September 2019; pp. 1–5.

30. Yu, A.; Yang, H.; Yao, Q.; Li, Y.; Guo, H.; Peng, T.; Li, H.; Zhang, J. Accurate Fault Location Using Deep Belief Network for Optical Fronthaul Networks in 5G and Beyond. *IEEE Access* **2019**, *7*, 77932–77943. [CrossRef]

31. Xiong, H.; Zhang, D.; Zhang, D.; Gauthier, V.; Yang, K.; Becker, M. MPaaS: Mobility prediction as a service in telecom cloud. *Inf. Syst. Front.* **2014**, *16*, 59–75. [CrossRef]

32. Cheng, Y.; Qiao, Y.; Yang, J. An improved Markov method for prediction of user mobility. In Proceedings of the 2016 12th International Conference on Network and Service Management (CNSM), Montreal, QC, Canada, 31 October–4 November 2016; pp. 394–399.

33. Qiao, Y.; Yang, J.; He, H.; Cheng, Y.; Ma, Z. User location prediction with energy efficiency model in the Long Term-Evolution network. *Int. J. Commun. Syst.* **2016**, *29*, 2169–2187. [CrossRef]

34. Gante, J.; Falcão, G.; Sousa, L. Beamformed fingerprint learning for accurate millimeter wave positioning. In Proceedings of the 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 27–30 August 2018; pp. 1–5.

35. Gante, J.; Falcão, G.; Sousa, L. Deep Learning Architectures for Accurate Millimeter Wave Positioning in 5G. *Neural Process. Lett.* **2019**. [CrossRef]

36. Wang, C.; Zhao, Z.; Sun, Q.; Zhang, H. Deep learning-based intelligent dual connectivity for mobility management in dense network. In Proceedings of the 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall), Chicago, IL, USA, 27–30 August 2018; pp. 1–5.

37. Santos, J.; Leroux, P.; Wauters, T.; Volckaert, B.; De Turck, F. Anomaly detection for smart city applications over 5g low power wide area networks. In Proceedings of the 2018 IEEE/IFIP Network Operations and Management Symposium, Taipei, Taiwan, 23–27 April 2018; pp. 1–9.

38. Maimó, L.F.; Gómez, Á.L.P.; Clemente, F.J.G.; Pérez, M.G.; Pérez, G.M. A self-adaptive deep learning-based system for anomaly detection in 5G networks. *IEEE Access* **2018**, *6*, 7700–7712. [CrossRef]

39. Parwez, M.S.; Rawat, D.B.; Garuba, M. Big data analytics for user-activity analysis and user-anomaly detection in mobile wireless network. *IEEE Trans. Ind. Inform.* **2017**, *13*, 2058–2065. [CrossRef]

40. Maimó, L.F.; Celdrán, A.H.; Pérez, M.G.; Clemente, F.J.G.; Pérez, G.M. Dynamic management of a deep learning-based anomaly detection system for 5G networks. *J. Ambient Intell. Humaniz. Comput.* **2019**, *10*, 3083–3097. [CrossRef]

41. Hussain, B.; Du, Q.; Ren, P. Deep learning-based big data-assisted anomaly detection in cellular networks. In Proceedings of the 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, UAE, 9–13 December 2018; pp. 1–6.

42. Li, R.; Zhao, Z.; Zheng, J.; Mei, C.; Cai, Y.; Zhang, H. The learning and prediction of application-level traffic data in cellular networks. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 3899–3912. [CrossRef]

43. Guo, Q.; Gu, R.; Wang, Z.; Zhao, T.; Ji, Y.; Kong, J.; Gour, R.; Jue, J.P. Proactive Dynamic Network Slicing with Deep Learning Based Short-Term Traffic Prediction for 5G Transport Network. In Proceedings of the 2019 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 3–7 March 2019; pp. 1–3.

44. Bega, D.; Gramaglia, M.; Fiore, M.; Banchs, A.; Costa-Perez, X. DeepCog: Cognitive network management in sliced 5G networks with deep learning. In Proceedings of the IEEE INFOCOM 2019—IEEE Conference on Computer Communications, Paris, France, 29 April–2 May 2019; pp. 280–288.

45. Huang, C.W.; Chiang, C.T.; Li, Q. A study of deep learning networks on mobile traffic forecasting. In Proceedings of the 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC, Canada, 8–13 October 2017; pp. 1–6.

46. Chen, L.; Yang, D.; Zhang, D.; Wang, C.; Li, J. Deep mobile traffic forecast and complementary base station clustering for C-RAN optimization. *J. Netw. Comput. Appl.* **2018**, *121*, 59–69. [CrossRef]

47. Zhou, Y.; Fadlullah, Z.M.; Mao, B.; Kato, N. A deep-learning-based radio resource assignment technique for 5G ultra dense networks. *IEEE Netw.* **2018**, *32*, 28–34. [CrossRef]

48. Zhao, X.; Yang, K.; Chen, Q.; Peng, D.; Jiang, H.; Xu, X.; Shuang, X. Deep learning based mobile data offloading in mobile edge computing systems. *Future Gener. Comput. Syst.* **2019**, *99*, 346–355. [CrossRef]

49. Hosny, K.M.; Khashaba, M.M.; Khedr, W.I.; Amer, F.A. New vertical handover prediction schemes for LTE-WLAN heterogeneous networks. *PLoS ONE* **2019**, *14*, e0215334. [CrossRef] [PubMed]

50. Svahn, C.; Sysoev, O.; Cirkic, M.; Gunnarsson, F.; Berglund, J. Inter-frequency radio signal quality prediction for handover, evaluated in 3GPP LTE. In Proceedings of the 2019 IEEE 89th Vehicular Technology Conference (VTC2019), Kuala Lumpur, Malaysia, 28 April–1 May 2019; pp. 1–5.

51. Khunteta, S.; Chavva, A.K.R. Deep learning based link failure mitigation. In Proceedings of the 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, Mexico, 18–21 December 2017; pp. 806–811.

52. Ozturk, M.; Gogate, M.; Onireti, O.; Adeel, A.; Hussain, A.; Imran, M.A. A novel deep learning driven, low-cost mobility prediction approach for 5G cellular networks: The case of the Control/Data Separation Architecture (CDSA). *Neurocomputing* **2019**, *358*, 479–489. [CrossRef]

53. Wen, J.; Huang, K.; Yang, S.; Li, V.O. Cache-enabled heterogeneous cellular networks: Optimal tier-level content placement. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 5939–5952. [CrossRef]

54. Serbetci, B.; Goseling, J. Optimal geographical caching in heterogeneous cellular networks with nonhomogeneous helpers. *arXiv* **2017**, arXiv: 1710.09626.

55. Chien, W.C.; Weng, H.Y.; Lai, C.F. Q-learning based collaborative cache allocation in mobile edge computing. *Future Gener. Comput. Syst.* **2020**, *102*, 603–610. [CrossRef]

56. Lei, L.; You, L.; Dai, G.; Vu, T.X.; Yuan, D.; Chatzinotas, S. A deep learning approach for optimizing content delivering in cache-enabled HetNet. In Proceedings of the 2017 International Symposium on Wireless Communication Systems (ISWCS), Bologna, Italy, 28–31 August 2017; pp. 449–453.

57. Yan, M.; Feng, G.; Zhou, J.; Sun, Y.; Liang, Y.C. Intelligent resource scheduling for 5G radio access network slicing. *IEEE Trans. Veh. Technol.* **2019**, *68*, 7691–7703. [CrossRef]

58. Gutterman, C.; Grinshpun, E.; Sharma, S.; Zussman, G. RAN resource usage prediction for a 5G slice broker. In Proceedings of the 20th ACM International Symposium on Mobile Ad Hoc Networking and Computing, Catania, Italy, 2–5 July 2019; pp. 231–240.

59. Toscano, M.; Grunwald, F.; Richart, M.; Baliosian, J.; Grampín, E.; Castro, A. Machine Learning Aided Network Slicing. In Proceedings of the 2019 21st International Conference on Transparent Optical Networks (ICTON), Angerrs, France, 9–13 July 2019; pp. 1–4.

60. Lei, L.; You, L.; He, Q.; Vu, T.X.; Chatzinotas, S.; Yuan, D.; Ottersten, B. Learning-assisted optimization for energy-efficient scheduling in deadline-aware NOMA systems. *IEEE Trans. Green Commun. Netw.* **2019**, *3*, 615–627. [CrossRef]

61. Luo, J.; Tang, J.; So, D.K.; Chen, G.; Cumanan, K.; Chambers, J.A. A deep learning-based approach to power minimization in multi-carrier NOMA with SWIPT. *IEEE Access* **2019**, *7*, 17450–17460. [CrossRef]

62. Dong, R.; She, C.; Hardjawana, W.; Li, Y.; Vucetic, B. Deep learning for hybrid 5G services in mobile edge computing systems: Learn from a digital twin. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 4692–4707. [CrossRef]

63. Zhang, S.; Xiang, C.; Cao, S.; Xu, S.; Zhu, J. Dynamic Carrier to MCPA Allocation for Energy Efficient Communication: Convex Relaxation Versus Deep Learning. *IEEE Trans. Green Commun. Netw.* **2019**, *3*, 628–640. [CrossRef]

64. Mismar, F.B.; Evans, B.L. Deep Learning in Downlink Coordinated Multipoint in New Radio Heterogeneous Networks. *IEEE Wirel. Commun. Lett.* **2019**, *8*, 1040–1043. [CrossRef]

65. Abdelreheem, A.; Omer, O.A.; Esmaiel, H.; Mohamed, U.S. Deep learning-based relay selection in D2D millimeter wave communications. In Proceedings of the 2019 International Conference on Computer and Information Sciences (ICCIS), Aljouf, Saudi Arabia, 10–11 April 2019; pp. 1–5.

66. Ouyang, Y.; Li, Z.; Su, L.; Lu, W.; Lin, Z. APP-SON: Application characteristics-driven SON to optimize 4G/5G network performance and quality of experience. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 1514–1523.

67. Ouyang, Y.; Li, Z.; Su, L.; Lu, W.; Lin, Z. Application behaviors Driven Self-Organizing Network (SON) for 4G LTE networks. *IEEE Trans. Netw. Sci. Eng.* **2018**, *7*, 3–14. [CrossRef]

68.　Kim, K.; Lee, J.; Choi, J. Deep learning based pilot allocation scheme (DL-PAS) for 5G massive MIMO system. *IEEE Commun. Lett.* **2018**, *22*, 828–831. [CrossRef]

69.　Jose, J.; Ashikhmin, A.; Marzetta, T.L.; Vishwanath, S. Pilot contamination problem in multi-cell TDD systems. In Proceedings of the 2009 IEEE International Symposium on Information Theory, Seoul, Korea, 21–26 June 2009; pp. 2184–2188.

70.　Zhang, W.; Yin, J.; Wu, D.; Guo, G.; Lai, Z. A Self-Interference Cancellation Method Based on Deep Learning for Beyond 5G Full-Duplex System. In Proceedings of the 2018 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Qingdao, China, 14–17 September 2018; pp. 1–5.

71.　Sun, J.; Shi, W.; Yang, Z.; Yang, J.; Gui, G. Behavioral modeling and linearization of wideband RF power amplifiers using BiLSTM networks for 5G wireless systems. *IEEE Trans. Veh. Technol.* **2019**, *68*, 10348–10356. [CrossRef]

72.　Zhao, F.; Zhang, Y.; Wang, Q. Multi-slot spectrum auction in heterogeneous networks based on deep feedforward network. *IEEE Access* **2018**, *6*, 45113–45119. [CrossRef]

73.　Roselló, M.M. Multi-path Scheduling with Deep Reinforcement Learning. In Proceedings of the 2019 European Conference on Networks and Communications (EuCNC), Valencia, Spain, 18–21 June 2019; pp. 400–405.

74.　Jameel, F.; Khan, W.U.; Chang, Z.; Ristaniemi, T.; Liu, J. Secrecy analysis and learning-based optimization of cooperative NOMA SWIPT systems. In Proceedings of the 2019 IEEE International Conference on Communications Workshops (ICC Workshops), Shanghai, China, 20–24 May 2019; pp. 1–6.

75.　Imai, T.; Kitao, K.; Inomata, M. Radio propagation prediction model using convolutional neural networks by deep learning. In Proceedings of the 2019 13th European Conference on Antennas and Propagation (EuCAP), Krakow, Poland, 31 March–5 April 2019; pp. 1–5.

76.　Lu, C.; Xu, W.; Shen, H.; Zhang, H.; You, X. An enhanced SCMA detector enabled by deep neural network. In Proceedings of the 2018 IEEE/CIC International Conference on Communications in China (ICCC), Beijing, China, 16–18 August 2018; pp. 835–839.

77.　Tsai, C.H.; Lin, K.H.; Wei, H.Y.; Yeh, F.M. QoE-aware Q-learning based approach to dynamic TDD uplink-downlink reconfiguration in indoor small cell networks. *Wirel. Netw.* **2019**, *25*, 3467–3479. [CrossRef]

78.　Wang, D.; Khosla, A.; Gargeya, R.; Irshad, H.; Beck, A.H. Deep learning for identifying metastatic breast cancer. *arXiv* **2016**, arXiv:1606.05718.

79.　Kachuee, M.; Fazeli, S.; Sarrafzadeh, M. Ecg heartbeat classification: A deep transferable representation. In Proceedings of the 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY, USA, 4–7 June 2018; pp. 443–444.

80.　Patil, K.; Kulkarni, M.; Sriraman, A.; Karande, S. Deep learning based car damage classification. In Proceedings of the 2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA), Cancun, Mexico, 18–21 December 2017; pp. 50–54.

81.　Song, Q.; Zhao, L.; Luo, X.; Dou, X. Using deep learning for classification of lung nodules on computed tomography images. *J. Healthc. Eng.* **2017**, *2017*, 1–7. [CrossRef]

82.　Kendall, A.; Cipolla, R. Geometric loss functions for camera pose regression with deep learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 5974–5983.

83.　Liu, C.; Wang, Z.; Wu, S.; Wu, S.; Xiao, K. Regression Task on Big Data with Convolutional Neural Network. In Proceedings of the International Conference on Advanced Machine Learning Technologies and Applications, Cairo, Egypt, 28–30 March 2019; pp. 52–58.

84.　Maqueda, A.I.; Loquercio, A.; Gallego, G.; García, N.; Scaramuzza, D. Event-based vision meets deep learning on steering prediction for self-driving cars. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 5419–5427.

85.　Fahrettin Koyuncu, C.; Gunesli, G.N.; Cetin-Atalay, R.; Gunduz-Demir, C. DeepDistance: A Multi-task Deep Regression Model for Cell Detection in Inverted Microscopy Images. *arXiv* **2019**, arXiv:1908.11211.

86.　Arulkumaran, K.; Deisenroth, M.P.; Brundage, M.; Bharath, A.A. A brief survey of deep reinforcement learning. *arXiv* **2017**, arXiv:1708.05866.

87.　Zaheer, M.; Ahmed, A.; Smola, A.J. Latent LSTM allocation joint clustering and non-linear dynamic modeling of sequential data. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; Volume 70, pp. 3967–3976.

88. Niu, D.; Liu, Y.; Cai, T.; Zheng, X.; Liu, T.; Zhou, S. A Novel Distributed Duration-Aware LSTM for Large Scale Sequential Data Analysis. In Proceedings of the CCF Conference on Big Data, Wuhan, China, 26–28 September 2019; pp. 120–134.

89. Yildirim, Ö. A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. *Comput. Biol. Med.* **2018**, *96*, 189–202. [CrossRef] [PubMed]

90. Greff, K.; Srivastava, R.K.; Koutník, J.; Steunebrink, B.R.; Schmidhuber, J. LSTM: A search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *28*, 2222–2232. [CrossRef]

91. Bai, S.; Kolter, J.Z.; Koltun, V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv* **2018**, arXiv:1803.01271.

92. Jiang, M.; Liang, Y.; Feng, X.; Fan, X.; Pei, Z.; Xue, Y.; Guan, R. Text classification based on deep belief network and softmax regression. *Neural Comput. Appl.* **2018**, *29*, 61–70. [CrossRef]

93. Barlacchi, G.; De Nadai, M.; Larcher, R.; Casella, A.; Chitic, C.; Torrisi, G.; Antonelli, F.; Vespignani, A.; Pentland, A.; Lepri, B. A multi-source dataset of urban life in the city of Milan and the Province of Trentino. *Sci. Data* **2015**, *2*, 150055. [CrossRef]

94. Garcia, S.; Grill, M.; Stiborek, J.; Zunino, A. An empirical comparison of botnet detection methods. *Comput. Secur.* **2014**, *45*, 100–123. [CrossRef]

95. Raca, D.; Quinlan, J.J.; Zahran, A.H.; Sreenan, C.J. Beyond throughput: A 4G LTE dataset with channel and context metrics. In Proceedings of the 9th ACM Multimedia Systems Conference, Amsterdam, The Netherlands, 12–15 June 2018; pp. 460–465.

96. Borges, V.C.; Cardoso, K.V.; Cerqueira, E.; Nogueira, M.; Santos, A. Aspirations, challenges, and open issues for software-based 5G networks in extremely dense and heterogeneous scenarios. *EURASIP J. Wirel. Commun. Netw.* **2015**, *2015*, 1–13. [CrossRef]

97. Ge, X.; Li, Z.; Li, S. 5G software defined vehicular networks. *IEEE Commun. Mag.* **2017**, *55*, 87–93. [CrossRef]

98. Ye, H.; Liang, L.; Li, G.Y.; Kim, J.; Lu, L.; Wu, M. Machine learning for vehicular networks: Recent advances and application examples. *IEEE Veh. Technol. Mag.* **2018**, *13*, 94–101. [CrossRef]

99. Li, S.; Da Xu, L.; Zhao, S. 5G Internet of Things: A survey. *J. Ind. Inf. Integr.* **2018**, *10*, 1–9. [CrossRef]

100. Kusume, K.; Fallgren, M.; Queseth, O.; Braun, V.; Gozalvez-Serrano, D.; Korthals, I.; Zimmermann, G.; Schubert, M.; Hossain, M.; Widaa, A.; et al. Updated scenarios, requirements and KPIs for 5G mobile and wireless system with recommendations for future investigations. In *Mobile and Wireless Communications Enablers for the Twenty-Twenty Information Society (METIS) Deliverable, ICT-317669-METIS D*; METIS: Stockholm, Sweden, 2015; Volume 1.