



Article

Deep Neural Networks for Spatial-Temporal Cyber-Physical Systems: A Survey

Abubakar Ahmad Musa ¹, Adamu Hussaini ¹, Weixian Liao ^{1,*}, Fan Liang ² and Wei Yu ^{1,*}

¹ Department of Computer and Information Sciences, Towson University, Towson, MD 21252, USA; aahmadm1@students.towson.edu (A.A.M.); ahussa7@students.towson.edu (A.H.)

² Department of Computer Science, Sam Houston State University, Huntsville, TX 77340, USA; fxl027@shsu.edu

* Correspondence: wliao@towson.edu (W.L.); wyu@towson.edu (W.Y.)

Abstract: Cyber-physical systems (CPS) refer to systems that integrate communication, control, and computational elements into physical processes to facilitate the control of physical systems and effective monitoring. The systems are designed to interact with the physical world, monitor and control the physical processes while in operation, and generate data. Deep Neural Networks (DNN) comprise multiple layers of interconnected neurons that process input data to produce predictions. Spatial-temporal data represents the physical world and its evolution over time and space. The generated spatial-temporal data is used to make decisions and control the behavior of CPS. This paper systematically reviews the applications of DNNs, namely convolutional, recurrent, and graphs, in handling spatial-temporal data in CPS. An extensive literature survey is conducted to determine the areas in which DNNs have successfully captured spatial-temporal data in CPS and the emerging areas that require attention. The research proposes a three-dimensional framework that considers: CPS (transportation, manufacturing, and others), Target (spatial-temporal data processing, anomaly detection, predictive maintenance, resource allocation, real-time decisions, and multi-modal data fusion), and DNN schemes (CNNs, RNNs, and GNNs). Finally, research areas that need further investigation are identified, such as performance and security. Addressing data quality, strict performance assurance, reliability, safety, and security resilience challenges are the areas that are required for further research.

Keywords: deep neural networks; spatial-temporal data; cyber-physical systems



Citation: Musa, A.A.; Hussaini, A.; Liao, W.; Liang, F.; Yu, W. Deep Neural Networks for Spatial-Temporal Cyber-Physical Systems: A Survey. *Future Internet* **2023**, *15*, 199. <https://doi.org/10.3390/fi15060199>

Academic Editor: Paolo Bellavista

Received: 17 April 2023

Revised: 22 May 2023

Accepted: 22 May 2023

Published: 30 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cyber-physical systems (CPS) are intended to integrate communication, control, and computational elements with physical processes, aiming to improve the effective monitoring and control of physical components. As the systems are designed to interact with the physical world, monitoring and controlling the physical processes shall generate a variety of data. The data is used to make decisions and control the behavior of CPS [1–3]. As shown in Figure 1, examples of CPS include autonomous vehicles in smart transportation, the industrial control system (ICS) in smart manufacturing, wearable sensors in medical CPS, etc. [4–8]. In CPS, vast amounts of data will be generated, representing space and time while operating themselves. Such data is generally denoted as CPS spatial-temporal data [9–11].

In CPS, the spatial-temporal data represents the physical world, and its evolution over time and space. CPS generates large amounts of data as they monitor and control physical processes in real-world objects. This data is used in CPS to make decisions and control behavior, such as predicting future events, detecting anomalies, and optimizing resource allocation, among others. For example, the spatial-temporal data collected by sensors in the smart transportation system detects obstacles and makes vehicle trajectory decisions.

ICS, a crucial component of smart manufacturing systems, utilizes spatial-temporal data to monitor and manage physical processes. The effective management and analysis of spatial-temporal data in CPS require sophisticated techniques and algorithms, including spatial data mining, temporal data analysis, and geographic information systems. The effective use of spatial-temporal data in CPS can significantly improve its performance, reliability, safety, and security [12].

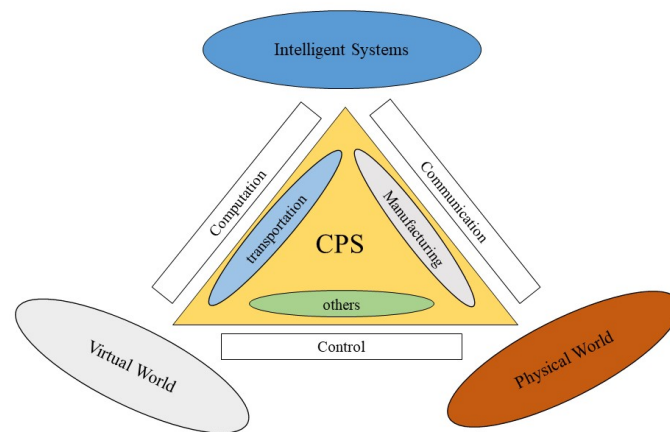


Figure 1. Example of the CPS Architecture.

Several techniques have evolved over the last decades to learn the complex patterns and changing dynamics of spatial-temporal data. This includes temporal time-series analysis (ARIMA, SARIMA, etc.) [13], spatial data analysis (spatial regression and autocorrelation, kriging, etc.) [14], signal processing approaches (Fourier and wavelet analysis, Kalman filtering, etc.) [15], and machine learning approaches (regression analysis, support vector machine, etc.) [16,17]. Nonetheless, most of these methods are not well-suited for handling large, dynamic, non-stationary data, which depicts its space and how it changes over time [18–21]. Deep Neural Networks (DNNs), on the other hand, are applicable due to their ability to handle immense amounts of data and their capability of modeling complex relationships, both spatially and temporally [22]. For example, DNNs predict future events based on spatial-temporal data in CPS, including traffic patterns in smart cities, resource demand, and fault or intrusion detection in smart manufacturing systems, among others [23–26].

The existing research efforts revealed numerous successes in the application of DNNs, including anomaly detection [27,28], resource management [24], predictive maintenance [29–32], multi-modal data fusion [33], real-time decision making [34,35], and spatial-temporal data processing [36,37], among others. For example, Luo et al. [28] reviewed the applications of deep learning for anomaly detection in CPS, outlining areas, where deep learning has achieved promising results and areas that need improvement. Likewise, Zhang et al. [38] reviewed the historical and state-of-the-art applications of deep learning in energy CPS (i.e., frequency analysis and control in power systems). Carvalho et al. [39] systematically reviewed machine learning applications in general to predictive maintenance of industrial CPS to determine the best-performing models and the areas of challenge. However, more research needs to be done to review the applications of DNNs in CPS with spatial-temporal datasets.

In line with Rowe et al. [40], the survey strategy adopted for this paper explored and selected the relevant journals and articles submitted to highly reputable venues that are accessible online. Research databases like Google Scholar, IEEE Xplore, ACM Digital Library, Science Direct, and Springer were used. We focused on titles, abstracts, keywords, and articles that include ‘deep neural networks’, ‘spatial-temporal data’, and ‘cyber-physical

systems'. Even though the 'AND' operation was used within the terms, i.e., 'deep neural networks' AND 'spatial-temporal data' AND 'cyber-physical systems,' it had more precise results. Furthermore, each paper was examined carefully to ensure that the selected DNNs were applied to spatial-temporal datasets in CPS for the experiments. In other words, each article failing to satisfy the requirements falls under the research exclusion category.

This survey paper systematically reviews the applications of DNNs in handling spatial-temporal data in CPS. The research areas are outlined where DNNs have successfully handled spatial-temporal data in CPS and the emerging areas that require improvements as well. The major contributions in this paper are as follows.

- The applications of Deep Neural Networks (DNNs) - convolutional, recurrent, and graphs in handling spatial-temporal data in CPS are systematically reviewed.
- A three-dimensional problem space that considers: CPS (transportation, manufacturing, and others), Target (spatial-temporal data processing, anomaly detection, predictive maintenance, resource allocation, realtime decisions, and multi-modal data fusion), and DNN scheme (CNNs, RNNs, and GNNs) is proposed.
- Future research directions concerning data quality, strict performance assurance and reliability, safety, and security resilience have been outlined.

The remainder of this paper is organized as follows. The background of DNNs, spatial-temporal data, as well as CPS are reviewed in Section 2. In Section 3, the state-of-the-art DNNs to handle spatial-temporal data in CPS are explored. In Section 4, the existing research efforts on applying DNN to different CPS application domains are reviewed. In Section 5, several challenges and future research directions are outlined. Finally, Section 6 concludes the paper.

2. Preliminaries

This section briefly discusses deep neural networks (DNNs), spatial-temporal data, and CPS, respectively.

2.1. Deep Neural Networks (DNN)

Generally speaking, DNN is comprised of multiple layers of interconnected neurons that process input data to make a decision (prediction, classification, etc.). The term "deep" means many layers, enabling them to learn increasingly complex features from the input data. DNNs are useful and popular nowadays for their ability to learn and extract vital features from data without the help of any domain experts. They have been used in different domains for various tasks and purposes (e.g., security, industrial component recognition, integrated design of components to optimize the overall system performance) [41–43]. Training DNNs involves adjusting the weights of connections between neurons to minimize the prediction error with optimization algorithms. Some of the algorithms under DNNs are CNN in computer vision, graph neural networks (GNNs) for graph-structured data, recurrent neural networks (RNNs) in sequential data problems, natural language processing (NLP) in processing text data, and many other applications.

2.2. Spatial-Temporal Data

It describes the location and time of an observation. The data changes as the location and time change [44], as shown in Figure 2. It is common in different domains, including climate science [45], transportation [46], manufacturing [47], ecology [48], and social sciences [49], among others [50]. Examples of spatial-temporal data include climate data that tracks weather patterns across regions and time [45], traffic data that records vehicle movement and traffic patterns across different times and locations [51], and social media data that captures the space and time of user activity [52]. The characteristics of spatial-temporal data typically result in more complex data correlations than conventional methods can handle. Additionally, they are frequently firmly self-correlated, and unlike traditional data, data samples are often not produced independently. It is more challenging to process spatial-temporal data than traditional stationary data. For instance, interpreting spatial-

temporal data is more complicated than interpreting pictures, where researchers could rely on visual inspections [44]. The data can be represented as raster images, trajectories, and many more. Understanding and analyzing spatial-temporal data is difficult due to the complexity and interdependent nature of the spatial and temporal dimensions. Traditional statistical methods have been applied, but they are not the final panacea, which calls for more advanced machine learning techniques, such as the DNNs.

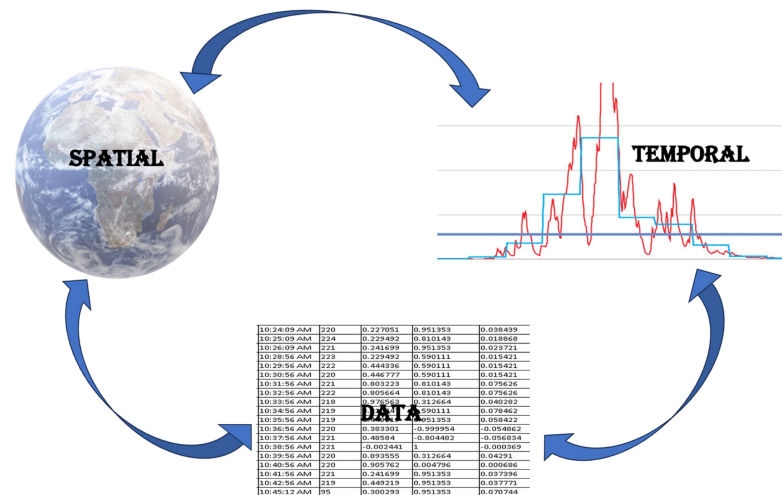


Figure 2. Spatial-Temporal Data.

2.3. Cyber Physical Systems (CPS)

It merges the physical world with the virtual world, driven by information communication technology, to create a new intelligent system that enables effective interaction with the environment. As a vertical architecture, CPS has various application domains, including transportation, manufacturing, and healthcare, among others [3,53]. It collects information about the physical world via sensors and responds to the system via actuators, and other parts. In the monitoring and control of physical systems, information collected needs to be transmitted to the processing unit (cloud, edge servers, etc.). The processing unit analyzes the data, makes decisions, and sends instructions to the actuators to control the physical system. CPS aims to revolutionize a wide range of industries by improving the performance that they render. For example, Industry 4.0 is the vision of CPS to realize the revolution of manufacturing processes in the industrial domain.

3. DNNs in CPS-Based Spatial-Temporal Data

In this section, the state-of-the-art DNNs used to handle spatial-temporal data in CPS are explored. As shown in Figure 3, a three-dimensional framework is proposed, in which X-axis indicates the CPS domains (transportation, manufacturing, and others), the Y-axis displays the targets (spatial-temporal data processing, anomaly detection, predictive maintenance, resource allocation, real-time decision, and multi-modal data fusion, among others), and the Z-axis reveals the different types of DNNs (i.e., CNNs, RNNs, and GNNs). Note that the purpose of mapping (say X_i , Y_j , Z_k) is to categorize the existing effort in the designed 3D framework. Recall that, in the X-axis, denote the transportation CPS as X_1 , manufacturing CPS as X_2 , etc.; in the Y-axis, denote Y_1 as data processing, Y_2 as anomaly detection, Y_3 as predictive maintenance, Y_4 as resource allocation, Y_5 as real-time decisions, and Y_6 as multi-modal data fusion; in Z-axis, denote CNNs as Z_1 , RNNs as Z_2 and GNN as Z_3 , respectively. Given that a specific DNN (say RNN) is applied to anomaly detection in transportation CPS, the corresponding effort can be categorized as (X_1, Y_2, Z_2) in the defined framework.

The proposed 3D framework can be used as a framework to summarize the existing research efforts concerning DNNs in handling spatial-temporal data in CPS. The designed framework can be used to categorize the existing research efforts, and help readers better understand the intersections among different DNNs with different application targets under different CPS. Furthermore, this designed framework is a generic one and can be extended to include more CPS, targets, and DNN techniques.

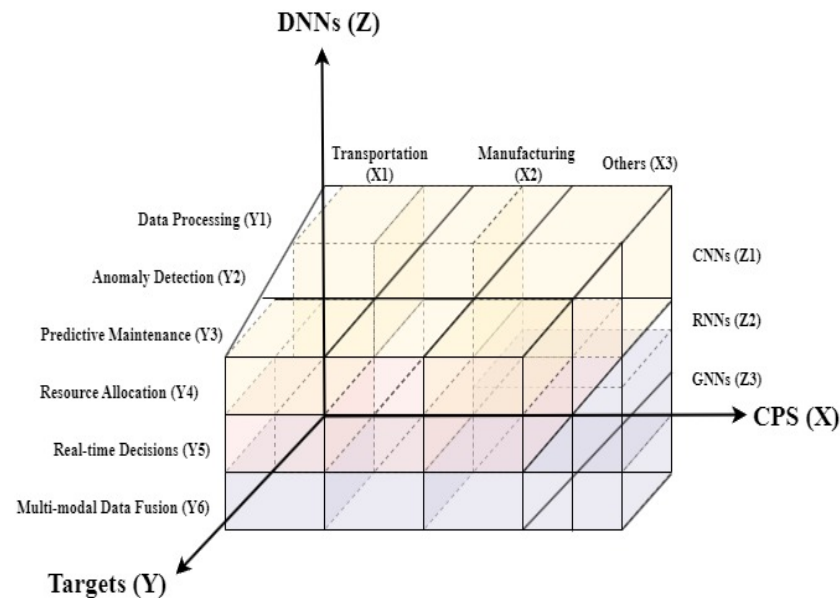


Figure 3. Problem Space for DNNs in CPS Spatial-Temporal Data.

3.1. The Problem Space

As denoted by Figure 3, six targets are defined to represent the fundamental research objectives that the DNNs have been used to handle spatial-temporal data in CPS. The targets emerged from the adopted research strategy described above and the careful consideration of the various goals achieved by the research conducted. For example, traffic speed, flow, and congestion prediction are achieved by processing road traffic data in transportation CPS. Research with these kinds of goals is categorized under spatial-temporal data processing. While those to detect or prevent the occurrence of attacks are categorized under anomaly detection. Real-time decisions go to autonomous vehicles or industrial control scenarios. However, in manufacturing CPS, for example, predicting equipment failures are classified as predictive maintenance, while the allocation of production logistics counts under resource allocation. Researchers that use data from various sources of different types are categorized under multi-modal data fusion.

The defined targets are elaborated further below:

- Spatial-temporal Data Processing (Y_1): It involves using DNNs to process and analyze spatial-temporal data in CPS, including feature extraction, classification, regression, and future event predictions.
- Anomaly Detection (Y_2): It involves using DNNs to detect anomalies based on spatial-temporal data in CPS, such as faults, attacks, or intrusions.
- Predictive Maintenance (Y_3): It involves using DNNs to predict the performance and maintenance needs of CPS according to spatial-temporal data, such as predicting the failure of smart manufacturing systems.
- Resource Allocation (Y_4): It involves using DNNs to optimize the allocation of resources in CPS, such as allocating network resources in communication networks and production logistics in manufacturing CPS.
- Real-time Decision (Y_5): It makes real-time decisions with spatial-temporal data, controlling the trajectory of autonomous vehicles.

- Multi-modal Data Fusion (Y_6): It includes employing DNNs to process data from multiple sources in CPS, such as combining data from sensors and communication networks to provide insightful information to the decision process.

3.2. CPS Application Domains

CPS can be classified into several domains based on the application area and the type of interactions between the physical and cyber components. In this context, CPS is categorized into application domains such as transportation, industrial manufacturing, and others.

3.2.1. Transportation CPS

It is popularly known as smart transportation. Physical systems are integrated with computational and communication techniques to realize various goals, such as improving traffic flow, toll collection, reducing congestion, smart parking, enhancing safety, safe pedestrian crossing, reducing carbon emissions, and autonomous driving, among others. Transportation CPS is a sophisticated, heterogeneous system that intends to offer effective services connected to various modes of transportation and traffic management. Transportation CPS manages a variety of new data sources, including geospatial transportation area, connected vehicle, roadside unit (RSU), and traffic network data. It also empowers users to be better informed and use transportation systems in a more innovative, safer, and organized fashion.

Smart transportation technology can offer services such as utilizing cameras to enforce traffic regulations, dialing 911 in the event of a car accident, and tracking the speed limit of vehicles, among others. There are several forms of security and privacy issues related to transportation CPS, targeting its essential elements such as IoT devices (sensors, actuators, microcontrollers, etc.), cloud services, and location-based services, among others. Examples of security issues include data tampering, man-in-the-middle (MITM) attacks, eavesdropping, impersonation, distributed denial of service (DDoS) attack, and artificial intelligence (AI)-based attacks, etc. In addition, model inversion, model poisoning, model evasion, and model extraction are typical ways of attacking AI models, which leads to severe impacts on driverless cars. While location privacy and commuter privacy are among the privacy concerns. To address the security and privacy issues in transportation CPS, both industry experts and researchers in academia are actively engaged in research to solve the problems. Nonetheless, despite these challenges, transportation CPS has the potential to revolutionize transportation systems, making them safer, more efficient, and more sustainable.

Transportation CPS can be categorized into vehicle transportation (which involves the integration of computing, communication, and physical systems within a vehicle); infrastructure transportation CPS (which consists of the integration of computing, communication, and physical procedures in transportation infrastructure); and system-level transportation CPS (which involves the integration of computing, communication, and physical systems at the system level).

Vehicle transportation CPS includes sensors, actuators, embedded systems, autonomous driving, etc. The main goal of vehicle transportation CPS is to improve safety, energy efficiency, and user experience. Transportation infrastructure for the transportation CPS includes traffic lights, sensors, cameras, communication networks, and other components that enable traffic monitoring, control, and management. The main objectives of transportation infrastructure CPS are to increase safety, improve traffic flow, and lessen congestion. System-level transportation CPS includes data analytics, simulation, optimization, and control algorithms that enable efficient and effective transportation planning, operation, and management. The main goal of system-level CPS is to improve the overall performance and sustainability of transportation systems. Real-time requirements for processing large amounts of data, secure communication mediums, interoperability among different systems, and effective collaboration between different stakeholders are among the challenges

faced by transportation CPS nowadays. However, despite these challenges, transportation CPS has the potential to revolutionize transportation systems, making them safer, more efficient, and more sustainable.

As shown in Figure 4, using the autonomous car as an example in the transportation CPS domain, the sensors/actuators will collect and send spatial-temporal data of the moving vehicle. With the help of the DNN model, the autonomous vehicle will obtain updated instructions about its environment.

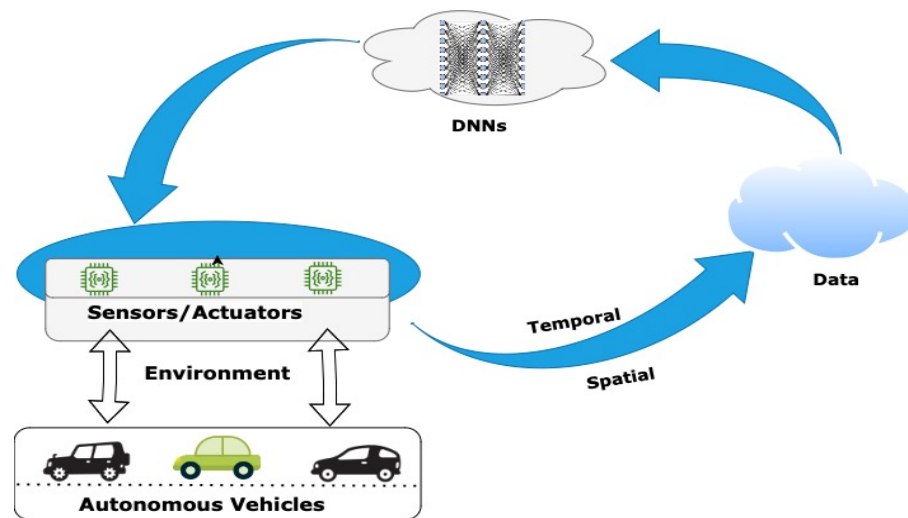


Figure 4. An Illustrative Example of Spatial-Temporal Data in Transportation CPS.

3.2.2. Manufacturing CPS

It is also known as smart manufacturing or Industry 4.0, which integrates physical processes with advanced computing and communication technologies to optimize manufacturing production and operation processes. The systems monitor and regulate physical processes using sensors and actuators while using data analytics and machine learning algorithms to streamline operations, boost productivity, and reduce costs. It applies to various manufacturing processes, including assembly lines, material handling, quality control, maintenance, and supply chain management. The Industrial Internet of Things (IIoT) and machine learning are the critical enabling blocks of manufacturing CPS. The former connects machines, sensors, and other devices to a network to collect and share data. This allows for real-time monitoring and control of equipment and the ability to control and adjust machines remotely. At the same time, the latter makes use of the data generated by the system. This includes predictive analytics, which can forecast equipment failures and maintenance needs, and prescriptive analytics, which can optimize manufacturing processes and improve product quality. With real-time data and analytics, industrial-based manufacturing CPS helps manufacturers optimize their production processes, reduce energy consumption, minimize waste, as well as enhance product quality.

3.3. DNN Techniques

There are different variants of DNNs. The most prominent techniques to handle spatial-temporal data in CPS are recurrent neural networks (RNNs), convolutional neural networks (CNNs), and graph neural networks (GNNs). Note that while RNN, CNN, and GNN are discussed, other DNN techniques can be expanded in the framework designed.

3.3.1. CNN

It refers to the class of DNNs, which is mainly used for image recognition and computer vision tasks. The essential attribute of this model is its ability to learn the spatial hierarchy of features from the input data, making it highly effective for tasks requiring an understanding

of the visual context of images. CNNs comprise several layers, including convolution, pooling, and fully connected layers. Each layer performs a specific function to extract features from the input data. The convolution layer applies filters or kernels to the input image, which slide over the entire image to capture specific features such as edges or corners. Each filter produces a feature map, representing a particular feature's presence in the input image. The pooling layer downsamples the feature maps obtained from the convolution layer to reduce their dimensionality and make subsequent computations more efficient. The Max Pooling layer can choose the maximum value from each sub-region of the feature map. After several convolution and pooling operations, the output is passed to the fully connected layer, which performs a non-linear transformation of the feature maps to produce a set of probabilities for each possible class. The final output is compared with the actual labels, and the network weights are updated using backpropagation, aiming to minimize the gap between the predicted and actual output. CNNs have been applied to various tasks, including object identification, facial recognition, and image classification. They have also been enhanced nowadays to accommodate other types of data (speech, text, video, etc.), making them a potent tool for numerous machine learning problems.

3.3.2. RNN

It refers to the DNNs designed to process sequential data, such as time series or text data, by maintaining an internal memory or state. The basic idea behind RNNs is that the output at a given time step is impacted not only by the current input but also by the previous inputs and the current state of the network. The network comprises a series of interconnected recurrent cells, which process the input at each time step and update the internal state of the network. Each cell takes the current and previous states as input, produces an output, and passes a new state to the next cell in the sequence. The back-propagation through time method trains RNNs, a variant of the standard backpropagation algorithm used to update the network weights based on the error signal at each time step. RNNs are suitable for application areas such as language modeling, speech recognition, etc. They are particularly effective for tasks requiring an understanding of the temporal dependencies of sequential data. Nonetheless, it is a challenging issue to train and prone to vanishing and exploding gradients. Various mechanisms have been developed to address the drawbacks, such as gradient clipping and regularization techniques (e.g., dropout).

RNN has two representative variants: one is LSTM and the other is GRU, which are briefly explained below.

- **LSTM:** It is an RNN variant designed to deal with challenges in RNNs in handling sequential data. Since conventional RNNs suffer from the “*vanishing gradient problem*” that limits their capability of capturing long-term dependencies between input and output sequences. LSTMs improve on that by remembering and selectively forgetting information over longer time horizons, making them effective for modeling sequences of variable lengths, such as natural language processing (text or speech). An LSTM consists of four main components: a memory cell, three gating units (an input, a forget, and an output gate), and an activation function. The memory cell stores information over long periods and passes it on to the next time step. The input gate controls data flow into the memory cell according to the current and previous outputs. Based on the input and output from the previous and current cycles, the forget gate determines the data to be erased from the memory cell. The output gate determines the output based on the current input and the current state of the memory cell. The activation function, mostly a hyperbolic tangent or sigmoid function, is used to compute the cell's current state. At each iteration, the LSTM unit receives an input vector and a hidden state vector from the previous time step and produces an output vector and a new hidden state vector. The input vector is passed through the input gate, and the output gate determines the output vector. The forget gate determines what information to keep from the previous hidden state, and the memory cell updates its internal state based on the input and the last hidden state. The updated memory cell is then passed on

to the next time step. LSTMs have been effective in various applications, including speech recognition, machine translation, image captioning, and music composition. They are often combined with other DNNs, such as CNNs or attention mechanisms, to realize performance on given tasks.

- GRU: It is a gating mechanism for RNN. It was introduced to serve as a simplified version of the complex LSTM. Like other RNNs, GRU processes sequential input data, such as text or time series, by maintaining a hidden state that captures information about the past sequence elements. It also uses a gating mechanism to selectively update and reset the hidden state, enabling it to capture longer-term dependencies in the sequence. GRU has two gating mechanisms: the reset gate and the update gate. The reset gate decides the past hidden state to forget, while the update gate makes the current input to incorporate into the new hidden state. GRU is a powerful and flexible neural network architecture that can capture long-term dependencies in sequential data and deal with the “*vanishing gradient problem*” that bedevils standard RNNs. It has been used in various applications, including NLP, speech/voice recognition, time series prediction, etc.

3.3.3. GNN

GNN is an architecture designed to operate on graph-structured data, such as traffic networks, social networks, molecular structures, etc. The framework incorporates information about the graph’s structure and the relationships between nodes into their computations. The fundamental idea behind GNNs is to represent each node in the graph as a vector or tensor and use message passing between nodes to update these representations based on the graph’s structure. At each iteration of the message-passing process, each node aggregates information from its neighbors and updates its representation based on a learned update function. There are different variations on the basic GNN architecture, depending on the specific problem being addressed [54]. For example, some GNNs use graph convolutional layers to learn local features of the graph, while others use attention mechanisms to learn global relationships between nodes. Some GNNs are designed to handle dynamic graphs that change over time, while others are designed for heterogeneous graphs with nodes and edges of different types. GNNs have been applied to address various problems, including traffic prediction, social network analysis, recommendation systems, and drug discovery, among others.

4. CPS Application Domains

As shown in Figure 3, several representative CPS application domains; transportation, manufacturing, and others are considered. Note that the transportation CPS and manufacturing CPS are two key examples to illustrate the designed framework outlined in Section 3 and show the existing efforts on applying DNN in representative CPS.

4.1. Transportation-Based CPS

We now review the recent literature on the DNNs that capture the data’s latent spatial-temporal features in the transportation-based CPS with respect to traffic forecasting, threat detection, data inconsistency identification, and autonomous vehicle collision prediction.

4.1.1. Traffic Forecasting

Traffic forecasting predicts future traffic patterns in an urban or city transportation system, which is necessary for traffic control, navigation systems management, and transportation planning. Accurate traffic forecasting aids in reducing congestion, enhancing safety, and maximizing the use of available transportation resources. Nowadays, DNNs learn the latent relationships and patterns within the traffic data to generate predictions based on those patterns. This motivated the effort of Zhou et al. [37], who proposed a “*wide-attention and deep-composite (WADC) model*”. To investigate its performance, they used CNN-LSTM to train the model with traffic flow spatial-temporal datasets. The result

revealed that it outperformed other models. Similarly, Guo et al. [55] proposed a “*graph attention-temporal convolutional network (GATCN)*” to forecast traffic speed in the short term. Graph attention and temporal convolution networks are combined to form each layer in the GATCN to apprehend the hidden spatial-temporal relationships concurrently. Likewise, Ma et al. [56] proposed a capsule network (CapsNet) and Nested LSTM (NLSTM) for network speed prediction, in which CapsNet was considered to extract extensive spatial features from roadway networks, while NLSTM was leveraged to capture traffic state hierarchical temporal dependencies.

Furthermore, Yan et al. [57] aimed to achieve an accurate and adaptable scheme for traffic flow prediction by proposing a graph-based network model. The model employed a fully connected layer to create a matrix from traffic data. LSTM was applied to the data to capture the temporal dependency, while ChebNet captured the spatial dependency. The spatial-temporal attributes were further combined for accurate traffic flow forecasting. Han et al. [58] stated that graph-based neural networks could be applied to enhance forecasting of traffic speed. To this end, they proposed a scheme that can learn time-specific spatial dependencies and a dynamic graph convolution module that aggregates hidden states of neighboring nodes to focal nodes using dynamic adjacency matrices and message passing. According to their study, the proposed scheme could offer clear and interpretable spatial relationships between road segments.

Furthermore, Tian et al. [59] proposed a multi-step prediction model that integrates CNN with an attention mechanism. In this way, the spatial-temporal dependencies and forecast traffic conditions of road networks could be captured. With the self-adaptive node embedding, the model is capable of extracting the latent spatial relationships in the data even without prior knowledge of the graph topology. Li et al. [60] observed that spatial-temporal correlations among road networks are changeable and complex. They proposed a model to achieve a dynamic traffic flow prediction model. Their proposed model comprises an adaptive mechanism block that preprocesses the data, improves its quality, and passes it to the multi-sensor data correlation convolution block to learn the dynamic temporal and spatial correlation among roads.

There are other related efforts. For example, Bai et al. [23] aimed for an effective traffic jam forecasting strategy in smart cities by proposing a “*Relative Position Congestion Tensor (RPCT)*” and a predictor for the “*Position Congestion Tensor*”. The proposed schemes leveraged the concept of relative locations to realize congestion matrices on regional traffic networks and convert them into spatial-temporal tensors. ConvLSTM was used to forecast future traffic congestion across the entire road network. Likewise, Lin et al. [61] discussed the significance of accurate traffic condition predictions in intelligent transportation systems. To this end, they proposed a “*graph convolution gated RNN (GCGRNN)*” to analyze multistep traffic volume by automatically determining the spatial-temporal dependencies in historical traffic data, where GCGRNN is based on encoder-decoder RNN and a data-driven graph filter. One benefit of their approach is that graph convolution is not dependent on a predefined adjacency matrix.

In the case of flight networks, there are some related studies. For example, Cai et al. [62] proposed an approach to carrying out the flight delay forecast. Their designed approach leveraged graph convolutional neural networks (GCN), which capture the insightful information of the airport network. In their study, an adaptive graph convolutional block was embedded in the proposed scheme so that the hidden spatial interactions in an airport network could be exposed. As another example, Peng et al. [63] observed that CNNs, GCNs, and RNNs were the most frequently utilized for extracting spatial-temporal features from traffic networks. They added that dynamic graphs could be more effective at reflecting the spatial-temporal features of the traffic network, but generating graph structures from data can be difficult. Thus, they proposed a long-term traffic flow prediction scheme that relies on GCN-LSTM to extract the spatial-temporal features for carrying out prediction. Furthermore, they developed a network of graph convolutional policies using the principle of reinforcement learning to create dynamic graphs when static ones are lacking because of

data sparsity problems. These efforts can be mapped to the cube $\langle X_1, Y_1, Z_1/Z_2/Z_3 \rangle$ in Figure 3 and Table 1. It means that, in those efforts, all the representative DNNs (CNNs, RNNs and GNNs) Z_1, Z_2 , and Z_3 are utilized for traffic prediction (Y_1) in transportation CPS (X_1).

4.1.2. Threat Detection

Threat detection involves ensuring the safety and reliability of transportation CPS by training models with the standard system behavior data to detect deviations from the said standard as anomalies. Some of these anomalous data may be targeted at cyber-attacks, trigger equipment failures, or cause environmental disturbances. For example, Kong et al. [64] proposed a framework combining trajectory data with environmental perception to detect outliers in driving behavior. The framework is comprised of trajectory processing, classification, and a mix of spatial-temporal-cost environments. Karim et al. [65] aimed to improve traffic safety by predicting accidents early on using video data recorded by dashboard cameras to study a dynamic spatial-temporal attention (DSTA) network model. The presented model combines both the dynamic temporal and spatial attention modules to focus on the most informative segments of a video and the spatial regions of frames. The gated recurrent unit module predicts the probability of a future accident.

Likewise, Diao et al. [66] aimed to prevent traffic accidents by proposing CRFAST-GCN, a multi-branch spatial-temporal attention graph convolution network that extracts long- and short-term dependencies, semantic similarity, and periodicity. Furthermore, Chen and Lv [67] considered improving the safety, performance, and development of intelligent transportation systems for autonomous vehicle in smart cities using digital twins and AI-based technologies. An architecture was proposed to use the 5G network so that resource load balancing scheduling could be provided to secure the transmission of autonomous vehicle data. A spatial-temporal graph convolution network technique was designed to forecast traffic flow in road networks, as well as real-space analysis of the compound traffic condition in the area of the road network using the concept of digital twin. These efforts can be mapped to the cube $\langle X_1, Y_2, Z_3 \rangle$ areas in Figure 3 and Table 1. It means that in these efforts, GNNs (Z_3) are utilized to detect the threats (Y_2) of transportation CPS (X_1).

Table 1. DNNs in Transportation CPS.

Research Objectives	Research Papers
$\langle X_1, Y_1, (\text{Transportation-Data Processing}), Z_1, Z_2, Z_3 \rangle$	[23,37,55–59,61–63,65–71]
$\langle X_1, Y_2, (\text{Transportation-Anomaly Detection}), Z_3 \rangle$	[64,65,67]
$\langle X_1, Y_3, (\text{Transportation-Predictive Maintenance}) \rangle$	N/A (No much research conducted in this direction)
$\langle X_1, Y_4, (\text{Transportation-Resource Allocation}) \rangle$	N/A (No much research conducted in this direction)
$\langle X_1, Y_5, (\text{Transportation-Real-time Decisions}), Z_3 \rangle$	[60,68,69]
$\langle X_1, Y_6, (\text{Transportation-Multi-modal Fusion}) \rangle$	N/A (No much research conducted in this direction)

4.1.3. Data Inconsistency Identification

Data inconsistency identification entails identifying and resolving inconsistencies in spatial-temporal datasets to ensure their accuracy and reliability. Related to this, Liang et al. [68] proposed a spatial-temporal aware data recovery network (STAR) to address the real-time spatial-temporal data imputation problem in a cooperative intelligent transportation system. The model is geared to handle the three types of data recovery tasks in real time and with inductive inference. Likewise, To infer missing values in the spatiotemporal input data, Kong et al. [69] proposed a novel paradigm for imputing traffic data. The model dramatically decreased the imputation error while increasing imputation

accuracy compared with the state-of-the-art. Additionally, the correlated information extracted from historical observations is used to deal with missing values. These efforts can be mapped to the cube $\langle X_1, Y_5, Z_3 \rangle$ area in Figure 3 and Table 1.

4.1.4. Autonomous Vehicle Collision Prediction

Autonomous vehicle collision prediction entails forecasting the likelihood of a collision between an autonomous vehicle and another object, such as a pedestrian or vehicle. Related to such an effort, Malawade et al. [70] proposed a spatial-temporal scene-graph embedding technique (SG2VEC), which adopts GNNs and LSTM layers to predict future autonomous vehicle accidents with the assistance of visual scene perception. Likewise, Sun et al. [71] adopted GNN and RNN to propose a global scheme called GST-GAT for traffic prediction. The framework leveraged “*global interaction + node query*” as a coherent way of information flow between nodes, which captures the interaction between traffic road networks that is spatial-temporal. These efforts can be mapped to the cube $\langle X_1, Y_5, Z_3 \rangle$ area in Figure 3 and Table 1.

Table 2 houses the identified research gaps and the contributions made by the reviewed efforts in the transportation domain.

Table 2. Summary of the Reviewed Contributions in Transportation (X_1) CPS.

Ref, Year	Framework	Gap	Contribution
[23], 2021	ConvLSTM	Traffic jam forecasting models were location specific	Created a model that covers the entire smart city
[37], 2020	CNN-LSTM	Traffic forecasting models were vehicle-type based	Proposed a generalized model applicable to all smart vehicles
[55], 2020	GATCN	Inaccurate prediction of spatial-temporal features from traffic data	An improved prediction of the hidden spatial-temporal features that lies in the data
[57], 2021	ChebNet-LSTM	Inadequate accuracy, deficient adaptability and inferior real-time	Proposed an enhanced real-time, adaptable prediction scheme
[58], 2021	DGNN	Previous models were built based on static adjacency matrix	Proposed a dynamic graph construction method
[59], 2021	STAWnet	Previous models were built based on the static dependency within the pre-defined structure	Designed a self attentive model that requires no prior knowledge of the graph topology
[61], 2021	GCGRNN	CNN models are inefficient in handling structure-varying data	Proposed a graph convolution approach that is independent of a pre-defined adjacency matrix
[62], 2021	GCN	Previous models were single airport specific	Proposed a flight delay prediction of airport networks framework
[63], 2021	GCN-LSTM	Previous models were static graphs based	Addressed data defects problems caused by the static graphs
[64], 2021	GNNs	Current methods were not security driven	Designed a taxi driving fraud detection system
[71], 2021	GST-GAT	Previous methods neglects the non-Euclidean nature of the road network	an improved dynamic spatial-temporal correlation method that captures the relevant characteristics of the traffic network
[65], 2022	DSTA	Conventional methods neglects the spatial-temporal features that exist in the data	Proposed Dashcams video data to predict/detect accidents
[66], 2022	CRFAST-GCN	Traditional forecasting methods neglects the semantic similarity between traffic nodes that degrades accuracy	Designed a model that extracts long and short-term dependencies, semantic similarity, and periodicity
[67], 2022	CNN-DT	Current accidents prediction methods do not operate in real-time	Presented digital-twins AI-enabled autonomous cars prediction model

Table 2. *Cont.*

Ref, Year	Framework	Gap	Contribution
[68], 2022	STAR	Imputing the missing entries in spatial-temporal traffic data is challenging	addressed the transport data corruption problem in real-time
[70], 2022	SG2VEC	Traditional collision prediction methods were expensive	Proposed an improved future autonomous vehicles accidents prediction method
[69], 2023	DGCRIN	Missing data imputation methods do not account for the dynamic spatial dependencies of the road network over time and the effective utilization of the diverse data	Designed an improved dynamic imputation method

4.2. Manufacturing CPS

In this section, research efforts that apply DNNs to apprehend the latent spatial-temporal attributes of manufacturing CPS data (real-time monitoring of factory logistics, production resource allocation, threat detection, etc.) are discussed.

4.2.1. Real-Time Monitoring of Factory Logistics

Wu et al. [34] proposed a scheme that integrates industrial IoT with digital twin technology to enable timely spatial-temporal traceability and visibility of manufacturing resources for efficient factory logistics. In their study, an LSTM network-based genetic indoor-tracking model was created and utilized to locate product trolleys with Bluetooth low energy and ultra-wide band technology. The extracted spatial-temporal features were used to activate location-based services for operational efficiency. This effort can be mapped to the cube $\langle X_2, Y_5, Z_2 \rangle$ as shown in Figure 3 and Table 3.

Table 3. DNNs in Manufacturing CPS.

Research Objectives	Research Papers
$\langle X_2, Y_5 \text{ (Industrial-Real-time Decisions)}, Z_2 \rangle$	[34]
$\langle X_2, Y_4 \text{ (Industrial-Resource Allocation)}, Z_3 \rangle$	[24]
$\langle X_2, Y_2 \text{ (Industrial-Anomaly Detection)}, Z_1, Z_3 \rangle$	[25,72,73]
$\langle X_2, Y_3 \text{ (Industrial-Predictive Maintenance)} \rangle$	[74–76]
$\langle X_2, Y_1 \text{ (Industrial-Data Processing)} \rangle$	N/A (No much research conducted in this direction)
$\langle X_2, Y_6 \text{ (Industrial-Multi-modal Data Fusion)} \rangle$	N/A (No much research conducted in this direction)

4.2.2. Production Resources Allocation

Zhao et al. [24] proposed a model that improves production logistics efficiency through effective resource allocation. The model adopts dynamic knowledge graph modeling and the digital twin spatial-temporal mapping method to learn and represent the spatial-temporal values and relationships among the resources. A graph algorithm is employed to allocate the resources. This effort can be mapped to the cube $\langle X_2, Y_3, Z_3 \rangle$ as shown in Figure 3 and Table 3.

4.2.3. Threat Detection

Anomaly detection mechanisms in manufacturing CPS are only effective if the non-linear spatial-temporal features of the industrial processing data are considered [25]. In their study, the authors proposed a method based on spatial-temporal modeling (AD-RoSM) for detecting FDIA in ICS [25]. Their proposed scheme employs a neural-based state estimation

model that utilizes CNN for time-related modeling and a mechanism for carrying out space-related modeling. In this way, the spatial-temporal correlations within the process data can be described explicitly. Yang et al. [72] proposed a graph representation-based scheme for the detection of multivariate time series anomalies in highly complex industrial processes. Their proposed model is capable of improving the existing techniques by offering spatial-temporal feature extraction and decision criteria based on spatial-temporal graph modeling with no predefined topological priors and a discriminative decision boundary. HiSTAR was shown to provide the expected anomaly detection performance and anomaly localization outcomes. Likewise, Liu et al. [73] adopted CNN on manufacturing spatial-temporal data to identify abnormal production processes. Their study was based on a pasting process in lead-acid battery production as a case study. The CNN-based approach was designed to recognize abnormal processes by analyzing spatial-temporal data from sensors. These efforts can be mapped to the cube $\langle X_2, Y_2, Z_1, \text{ and } Z_3 \rangle$ as shown in Figure 3 and Table 3.

4.2.4. Predictive Maintenance

Li et al. [74] proposed a convolutional network model that mines deterioration information in order to anticipate the remaining usable life of a machine. Their designed scheme models the sensor network by taking into account both the spatial-temporal dependencies of the sensors. It adopts a hierarchical graph representation layer to model spatial dependencies, a bi-directional LSTM to model temporal dependencies, as well as a regularized self-attention graph pooling for effective information fusion. Yang et al. [75] proposed SuperGraph, a feature extraction technique for diagnosing rotating machinery faults. The technique adopts graph theory-based spectrum analysis so that a spatial-temporal graph can be constructed and a Laplacian matrix-based feature vector can be derived. GCN was further utilized to learn the latent features. Shcherbakov et al. [77] proposed a hybrid multi-task learning framework by integrating CNN and LSTM to reflect the relatedness of functional life prediction with the health status detection process for complex multi-object systems in the CPS environment. The CNN extracts significant spatial-temporal features from raw multi-sensory input data and compresses the condition monitoring data, while the LSTM captures the temporal dependencies. As another example, Zhang et al. [76] proposed an equipment fault prediction technique using spatial-temporal graph information. Their proposed scheme has the potential to stop fatal damage and reduce equipment maintenance costs. Their experimental results showed that their approach is capable of offering precise short-term and long-term fault prediction. These efforts can be mapped to the cube $\langle X_2, Y_3, Z_2, \text{ and } Z_3 \rangle$ as shown in Figure 3 and Table 3.

There are other related efforts concerning predictive maintenance. For instance, Xiong et al. [78] discussed the importance of human-robot collaboration (HRC) in smart manufacturing processes and the role of human action recognition in enabling HRC. In their study, a method based on optical flow and CNN transfer learning was proposed. Their proposed scheme leverages the optical flow to extract time-related information from video images and simultaneously parse spatial-temporal information with a two-stream CNN structure. Transfer learning was also leveraged to establish feature extraction capability by pre-training the model on a non-manufacturing specific dataset and transferring the gained knowledge to the target domain of assembly tasks, which have limited training samples. Zheng et al. [79] addressed the problems of scene recognition in underground coal mining using CNN, LSTM, and an attention mechanism. Jia et al. [80] proposed a data-driven method using a graph convolution network to model the compound and time-varying characteristics of the process industry. The technique tends to capture the relationships among variables. The model was trained with regularization terms so that distinctive localized spatial-temporal correlations can be learned and time-series properties can be derived using temporal convolution.

Furthermore, Li et al. [81] proposed CLSTMA, a hybrid model that integrates CNN, LSTM, and an attention mechanism to monitor water quality in a wastewater treatment system. In their proposed model, a sequential fusion CNN, LSTM, and attention mechanism

were used to predict water quality and assist in the reduction of energy and emissions. Their proposed scheme captures the fused spatial features using CNN, LSTM for the temporal information, and variable-weighted calculations using the attention mechanism. Likewise, Guo et al. [82] employed historical energy consumption time series and previous knowledge of material flow to propose a spatial-temporal deep learning network (STDLN) framework, which merges a GCN and a GRU and forecasts the energy consumption of nodes.

In order to enhance maintenance practices in production CPS, Bampoula et al. [83] adopted autoencoders to conduct predictive maintenance so that maintenance planning can be enabled based on real-time machine operation. Table 3 and Table 4 summarize the identified research gaps and the contributions made by the reviewed efforts in the manufacturing domain.

Table 4. Summary of the Reviewed Contributions in Manufacturing (X_2) CPS.

Ref, Year	Framework	Gap	Contribution
[24], 2022	DSTKG	The Dynamites of the operating environment makes the efficient allocation of production logistics challenging	Proposed a framework for the adequate allocation of smart production logistics
[73], 2020	CNN	Abnormal manufacturing processes are not well examined	Extensible recognition framework for identifying abnormal manufacturing processes
[78], 2020	CNN	Accurate prediction of the evolving human activities in Human-Robot-Collaboration (HRC) was challenging	Optical flow CNN-based transfer learning technique was leveraged to promote HRC in smart manufacturing system
[79], 2020	CNN-LSTM	previous methods were inaccurate	An improved scene identification framework for underground coal mining
[74], 2021	HAGCN	Machineries remaining useful life prediction models do not consider their operating environment	Considered the environment and its dynamic features for improved accuracy
[75], 2021	SuperGraph	Achievement of a generic feature extraction method for vibration signals was challenging	Graph theory was found supportive and successfully applied for rotating machinery fault diagnosis
[83], 2021	LSTM-Autoencoders	Data quality challenges and preventive maintenance strategies	Proposed predictive maintenance for steel industry production processes
[81], 2021	CLSTMA	Effective feature extraction methods, data size and scalability challenges	An improved scheme for monitoring water quality in a wastewater treatment system
[82], 2022	GCN-GRU	Early energy prediction of complex nodes in smart manufacturing systems remains a challenge	Framework for predicting energy consumption behaviour of multiple nodes concurrently
[34], 2022	LSTM-GITA	Real-time monitoring of factory logistics is under-investigated	IIoT-DT based method of monitoring factory logistics
[25], 2022	AD-RoSM	Accuracy problems for anomaly detection schemes in ICS	An improved anomaly detection method for ICS
[72], 2022	HiSTAR	Spatial-temporal feature extraction was challenging for multivariate industrial anomaly detection schemes	A graph-theory concept was proposed for improved feature extraction and detection accuracy
[77], 2022	CNN-LSTM	Health assessment of complex systems remains a challenge	A method for evaluating the health status of smart CPS
[76], 2022	Markov graph	Accuracy problems in predicting faults for electromechanical instruments	An improved scheme aims to stop fatal damage and reduce equipment maintenance costs
[80], 2023	GCN	Compound and time-varying characteristics of the process industry are not well investigated	Graph theory was used to capture the inherent relationships among the affected variables for an improved accuracy

4.3. Other CPS

Apart from transportation CPS and manufacturing CPS, there are other types of CPS in different application domains, such as smart cities, medical CPS, aviation CPS, etc.

4.3.1. Flood Prediction

Related to smart cities as an important application domain of CPS, Chen et al. [84] proposed a flood process prediction model based on CNN using a decade's worth of historical data collected by smart sensors in city infrastructure. To predict the peak of the flood and its arrival time, the model takes rainfall spatial-temporal, geographical, and trend features into account. The model was presented to predict stream flow by integrating the rainfall spatial-temporal feature obtained through analyzing the historical stream flow and the digital elevation model data. These efforts can be mapped to the cube $\langle X_3, Y_1, Z_1, \text{ and } Z_3 \rangle$ as shown in Figure 3 and Table 5.

Table 5. DNNs in Other CPS.

Research Objectives	Research Papers
$\langle X_3, Y_1 \text{ (Other CPS-Data Processing), } Z_1, Z_2, Z_3 \rangle$	[84,85]
$\langle X_3, Y_2 \text{ (Other CPS-Anomaly Detection), } Z_1 \rangle$	[86]
$\langle X_3, Y_3 \text{ (Other CPS-Predictive Maintenance), } Z_1 \rangle$	N/A (No much research conducted in this direction)
$\langle X_3, Y_4 \text{ (Other CPS-Resource Allocations) } \rangle$	N/A (No much research conducted in this direction)
$\langle X_3, Y_5 \text{ (Other CPS-Real-time Decisions), } Z_1, Z_2 \rangle$	[8,87,88]
$\langle X_3, Y_6 \text{ (Other CPS-Multi-modal Data Fusion) } \rangle$	N/A (No much research conducted in this direction)

4.3.2. CPS-Data Processing

Related to smart cities and aviation CPS, Jiang et al. [85] proposed a GNN-based approach for predicting air mobility to enable the control and decision-making process in the airport of things. In their study, a spatial-temporal GCNN was employed to capture the latent characteristics of the graph-structured data. Their proposed approach was validated using airline on-time performance data and found to be effective in predicting spatial-temporal air mobility. This effort can be mapped to the cube $\langle X_3, Y_1, Z_1, \text{ and } Z_3 \rangle$ as shown in Figure 3 and Table 5.

4.3.3. Physical Attack Detection

Related to smart cities, Pan et al. [86] adopted ConvLSTM to propose a method for detecting threats (from cyber or physical spaces) against cyber-physical surveillance cameras. The technique uses a new video frame interpolation to detect video anomalies in spatial-temporal feeds. This effort can be mapped to the cube $\langle X_3, Y_2, Z_1, Z_2 \rangle$ as shown in Figure 3 and Table 5.

4.3.4. Real-Time Fire Identification Systems

Also related to smart cities, Zhang et al. [87] developed a real-time fire identification system that uses an IoT sensor network, cloud server, AI engine, and user interface to collect, store, process, and display complex building fire information. Their designed system also leveraged Conv-LSTM neural network. The neural network was trained based on given numerical data and validated in a fire test room with successful results. This effort can be mapped to the cube $\langle X_3, Y_5, Z_1, Z_2 \rangle$ as shown in Figure 3 and Table 5.

4.3.5. Medical CPS

Wang et al. [8] developed a framework (PhysiQ) that uses passive sensory detection to track and objectively assess people's off-site physical therapy exercises in real-time using a smartwatch. The system used a multi-task spatial-temporal Siamese neural network to evaluate the effectiveness of exercises based on absolute and relative quality. Exercises were assessed by PhysiQ using metrics (i.e., range of motion, stability, and repetition).

Ge et al. [89] adopted RNN-LSTM with an attention mechanism to determine the specific variable patterns in a medical application. Likewise, Pan et al. [88] proposed a temporal-based Swin Transformer network (TSTNet) for the surgical video workflow recognition problem. These efforts can be mapped to the cube $\langle X_3, Y_5, Z_1, Z_2 \rangle$ as shown in Figure 3 and Table 5.

Table 6 outlined the identified research gaps and the contributions made by the reviewed efforts in the other CPS domain.

Table 6. Summary of the Reviewed Contributions in Other (X_3) CPS.

Ref, Year	Framework	Gap	Contribution
[8], 2023	PhysiQ	Lack of adequate and convenient methods of tracking exercises done off-site	Proposed a framework that tracks and assesses people's off-site physical therapy exercises in real time using a smartwatch.
[86], 2019	ConvLSTM	Integrity attacks against the physical configuration of cyber-physical devices are underinvestigated	Method of detecting threats against cyber-physical surveillance cameras
[84], 2021	CNN	Both traditional and data-driven methods are inefficient for flood process forecasting	Proposed an effective method of flood process prediction
[85], 2021	GNN	Previous methods neglected the propagation of traffic perturbations among airports	Air mobility prediction model for effective control and decision-making in the airport of things network
[87], 2022	Conv-LSTM	Modern firefighting systems need to be integrated with the state-of-the-art technologies	Proposed a real-time fire identification system
[89], 2022	RNN-LSTM	Methods for status monitoring and evaluating patients in ICUs neglect the temporal features of their operating environment	A method for determining the longitudinal variable patterns associated with the higher risk of medical ICU patient mortality.

5. Challenges and Future Research Directions

Despite DNNs having achieved remarkable success in handling spatial-temporal data in CPS, there are some limitations and challenging issues that require attention in future research. As for the limitations, it is affirmed that DNNs are not the final panacea to all spatial-temporal data problems in CPS, which calls for integrating other sophisticated machine learning and data analytics techniques (continuous learning, and transfer learning, among others). Similarly, in the areas where DNNs have been successfully applied, it is also realized that they raise additional challenging issues to CPS, ranging from longer training times to insufficient training data, which in turn conflicts with the strict performance requirements in CPS.

Note that the challenges and future research directions listed in this section are not only based on the thorough literature review of this topic but also based on our research experience and vision in this topic. As future research directions, we outline three fundamental challenges: Data Quality Assurance, Strict Performance Assurance, and Reliability, Safety, and Security Resilience, which consider both data quality that affects the effectiveness of DNNs and the performance requirements of CPS. Therefore, the purpose of this section is to present the limitations and challenges examined, which are later supported by future research directions from our vision, and we believe that those challenges should be addressed by the research community. Other technical challenges that can affect the application of DNN in spatial-temporal data in CPS are high computational power, problem complexity, and the learning hyperparameter.

- **Performance:** Real-time communication could be impacted by the latency caused by several protocols, especially when event-driven communication and detection are involved. Some protocols influence the performance of DNNs while handling CPS spatial-temporal data, i.e., by affecting data transmission, size, latency, reliability,

and synchronization. For example, network protocols (UDP and TCP by determining the reliability and latency of data transmission), data serialization protocols (JSON, protocol buffers by affecting the data size, encoding/decoding overhead), compression protocols (by scaling/shrinking the data size during transmission to improve the network performance), real-time communication protocols (MQTT, DDS by providing low-latency, publish-subscribe messaging for timely data delivery), and synchronization protocols (PTP, NTP by ensuring time synchronization in distributed systems, which aides coordinated processing).

Cross-platform sensor-actuator communication remains a challenging issue, and it is important to design a comprehensive quality of service framework and satisfy the performance requirements of CPS. In addition, sensor failure is another remaining challenge because most CPS heavily rely on sensing data for the sake of control and motoring purposes. The entire CPS will not function well if there are failures of some sensors within the ecosystem. Thus, the deployment model shall be thoroughly studied to guarantee the robustness of sensor deployment (e.g., coverage, connectivity). Therefore, it is critical to design a holistic solution to ensure the overall performance of CPS by considering all components and their integration as one complex system. The realization of the performance satisfaction of CPS systems depends on the performance with respect to computing, control, and communication. Thus, it is critical to design the modeling and optimization techniques to integrate all components (sensing, networking, computing, and data). Some existing research efforts have been conducted to address the integration of some components (sensing, control, networking, and data). Nonetheless, to enhance the performance of CPS, how all its components interact and interplay jointly, leading to a unified design and optimization strategy, is worth investigation.

- **Security:** CPS has unique system requirements and security challenges. Specifically, the confidentiality, integrity, and availability (CIA) security paradigm has been widely used to design security standards for information technology-driven systems. For example, availability is a crucial property regarding security and an essential requirement of a CPS. Different threats (DoS attacks, malware propagation, etc.) could affect the availability of CPS. Under this situation, computing and networking components in the CPS shall employ effective mitigation measures so that malicious computing requests and traffic can be detected in time and the impact of such attacks can be effectively mitigated. For CPS integrity, an ML model that depends on real-time data inputs is critical for the realization of highly dependable and trustworthy CPS (transportation infrastructure, manufacturing infrastructure, etc.). Data fidelity is crucial for the CPS, as it is the information that can accurately simulate and direct the physical system in response to environmental changes accurately and quickly. In CPS, an adversary could compromise the integrity of sensing data by intercepting the communication channel using either a man-in-the-middle (MITM) attack or the commands transmitted by the programmable logic controllers. Thus, security measures (device authentication, etc.) shall be in place to prevent unauthorized users from changing data. Although solutions based on cryptography have been promoted in the context of CPS, such as those that use TLS, HMACs, or other authentication and integrity guarantees. Historically, such countermeasures have not been widely used due to hardware restrictions and the relative computational cost of deploying protocols and mechanisms. Table 7, summarizes the challenges of utilizing the DNN in CPS spatial-temporal data.

Based on the requirements of performance and security in CPS, we consider the following fundamental research challenges that are required for further research concerning the performance of DNNs and the performance requirement of CPS.

- **Data Quality Assurance for Effective DNNs:** The spatial-temporal data in CPS is characterized as complex (generated from multiple sources of sensors, microcontrollers, etc.), incomplete (measurement errors, missing values, outliers, etc.), noisy data (real-time streaming data), challenging to interpret, and unavailability in some cases, among others. There is a need to address these challenges, i.e., by developing better data collection techniques, missing data imputation and normalization methods, and new feature extraction procedures that can effectively capture the relevant latent spatial-temporal information in large and complex datasets.

Similarly, explainable AI can be leveraged to develop more precise, interpretable, and explainable DNNs that provide the detailed underlying features and relationships driving results or decisions. On the other hand, transfer learning can be leveraged with existing knowledge and pre-trained models to improve the accuracy and efficiency of DNNs. For example, the transfer of knowledge from related spatial-temporal datasets within or across the different CPS domains would be beneficial for improving ML model efficiency and supporting the CPS co-design initiative.

- **Strict Performance Assurance for CPS:** Most models that handle spatial-temporal data in CPS are highly complex, combining two or more DNNs for a given task (CNN-LSTM, GCN-GTN, etc.). This calls for the use of multiple layers and many parameters, leading to a longer training time and hindering real-time performance in practice. In a nutshell, the computational complexity, which translates to communication delays, and the dynamic nature of CPS data are among the factors hindering the achievement of real-time performance in various CPS domains. This calls for the design of new efficient model architectures that require few parameters to handle CPS spatial-temporal data.

The targeted model architectures can significantly reduce the computational requirements and memory footprint of DNN models, making them more suitable for real-time tasks. Such architectures can include lightweight models, such as MobileNet and ShuffleNet, with smaller parameters that can be executed on CPS-resource-constrained devices. Alternatively, using specialized hardware, such as field-programmable gate arrays, can further optimize the execution of DNN models. Furthermore, the targeted model architectures can enable the deployment of DNN models in edge devices (sensors, actuators, smart cameras, etc.), which can process data locally within the network edge to reduce communication latency with the cloud. This can be critical for CPS applications, which require real-time decision-making and control. Similarly, developing and using continuous learning techniques that can adapt to the changing CPS data in real time can improve the performance, accuracy, and reliability of DNN models as well.

- **Reliability, Safety, and Security Resilience Insurance for DNNs and CPS:** Reliability entails predicting, detecting, and mitigating failures, while safety guarantees the system by dealing with unexpected failures. Security resilience entails preventing security threats posed by adversaries. As for reliability and safety, CPS could fail due to hardware or software faults resulting in inconsistent spatial-temporal data, which might lead the DNN models to make incorrect predictions or even shut down the system. There is a need for models that adopt spatial-temporal data to predict, detect, and mitigate hardware and software failures in CPS. Similarly, methods for evaluating the reliability and availability of DNN models and systems, such as reliability metrics and failure analysis, are critical too. Developing strategies for implementing fault-tolerant techniques and self-healing CPS cost-effectively and efficiently is critical for the predictive maintenance of CPS.

On the other hand, data breaches and theft can compromise the confidentiality and integrity of spatial-temporal data in CPS. Furthermore, threats can lead the models to make incorrect predictions. In addition, malware and other cyberattacks can infect CPS and disrupt its regular operations. There is a need for strategies to predict, detect, and prevent attacks (defensive distillation, adversarial training, etc.). Additionally,

introducing methods to ensure the privacy and confidentiality of spatial-temporal data in CPS (such as access control, encryption, and privacy-preserving machine learning techniques) is necessary. Finally, procedures for detecting and mitigating malware and other cyberattacks on CPS systems (e.g., intrusion detection and network segmentation) shall be considered as well.

Table 7. Challenges of Using DNN in CPS Spatial-Temporal Data.

S/NO	Performance	Security
1	Longer training time	Data Fidelity issue
2	High computational power	Attack on availability
3	Insufficient training data	Sensing data attack
4	Problem complexity	Cost of protocols deployment
5	Learning hyperparameter	PLC compromise

6. Final Remarks

CPS combines computational, control, and communication components with physical processes. It is made to communicate with the physical world, keep track of and manage operational physical processes, and produce data. Within the operation of CPS, “spatial-temporal data” refers to the data used to describe the physical world and how it has changed over time. Decisions are made with spatial-temporal data to regulate the behavior and operation of CPS. This paper systematically reviewed the applications of DNNs, i.e., convolutional, recurrent, and graph neural networks, in handling spatial-temporal data in CPS. Additionally, an extensive literature survey was conducted to determine the areas, in which DNNs have successfully taken spatial-temporal data in representative CPS and the emerging areas that require attention. A generic three-dimensional framework was proposed by considering the type of CPS, target (spatial-temporal data processing, anomaly detection, predictive maintenance, resource allocation, real-time decisions, and multi-modal data fusion), and DNN schemes (CNNs, RNNs, and GNNs). Finally, research areas that need further investigation, such as performance and security, were identified. Additionally, data quality assurance, strict performance assurance, reliability, safety, and security resilience challenges were outlined as future research challenges and opportunities.

In the future, this line of research could be extended by conducting several case studies to address the areas that have not received sufficient attention (i.e., “N/A”) as depicted by Tables 1–5. Similarly, other attention mechanisms like the Transformers could be explored further and employed in this domain to compare their performance with the DNNs considered in this survey.

Author Contributions: Conceptualization, A.A.M., W.L. and W.Y.; methodology, A.A.M., A.H., W.L. and W.Y.; writing—original draft preparation, A.A.M., A.H. and W.Y.; writing—review and editing, W.L., F.L. and W.Y.; problem space and formalization, A.A.M., A.H., W.L. and W.Y.; supervision, W.L., F.L. and W.Y.; project administration, W.L., F.L. and W.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable, the study does not report any data.

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

References

1. Lee, E.A. Cyber physical systems: Design challenges. In Proceedings of the 2008 11th IEEE international symposium on object and component-oriented real-time distributed computing (ISORC), Orlando, FL, USA, 5–7 May 2008; pp. 363–369.
2. Xu, H.; Liang, F.; Yu, W. Internet of Things: Architecture, Key Applications, and Security Impacts. In *Encyclopedia of Wireless Networks*; Springer: Cham, Switzerland, 2020; pp. 672–681.

3. Lin, J.; Yu, W.; Zhang, N.; Yang, X.; Zhang, H.; Zhao, W. A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications. *IEEE Internet Things J.* **2017**, *4*, 1125–1142. [\[CrossRef\]](#)
4. Mitchell, R.; Chen, I.R. A survey of intrusion detection techniques for cyber-physical systems. *ACM Comput. Surv. (CSUR)* **2014**, *46*, 1–29. [\[CrossRef\]](#)
5. Liu, X.; Qian, C.; Hatcher, W.G.; Xu, H.; Liao, W.; Yu, W. Secure Internet of Things (IoT)-based smart-world critical infrastructures: Survey, case study and research opportunities. *IEEE Access* **2019**, *7*, 79523–79544. [\[CrossRef\]](#)
6. Zahran, B.; Hussaini, A.; Ali-Gombe, A. IIoT-ARAS: IIoT/ICS Automated risk assessment system for prediction and prevention. In Proceedings of the Eleventh ACM Conference on Data and Application Security and Privacy, Virtual Event, 26–28 April 2021; pp. 305–307.
7. Ramasamy, L.K.; Khan, F.; Shah, M.; Prasad, B.V.V.S.; Iwendi, C.; Biamba, C. Secure smart wearable computing through artificial intelligence-enabled internet of things and cyber-physical systems for health monitoring. *Sensors* **2022**, *22*, 1076. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Wang, H.D.; Ma, M. PhysiQ: Off-site Quality Assessment of Exercise in Physical Therapy. *Proc. Acm Interact. Mob. Wearable Ubiquitous Technol.* **2023**, *6*, 1–25. [\[CrossRef\]](#)
9. Tan, Y.; Vuran, M.C.; Goddard, S. Spatio-temporal event model for cyber-physical systems. In Proceedings of the 2009 29th IEEE International Conference on Distributed Computing Systems Workshops, Montreal, QC, Canada, 22–26 June 2009; pp. 44–50.
10. Mohammadi, M.; Al-Fuqaha, A.; Sorour, S.; Guizani, M. Deep learning for IoT big data and streaming analytics: A survey. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 2923–2960. [\[CrossRef\]](#)
11. Liu, X.; Xu, H.; Liao, W.; Yu, W. Reinforcement learning for cyber-physical systems. In Proceedings of the 2019 IEEE International Conference on Industrial Internet (ICII), Orlando, FL, USA, 1–12 November 2019; pp. 318–327.
12. Iqbal, R.; Doctor, F.; More, B.; Mahmud, S.; Yousuf, U. Big Data analytics and Computational Intelligence for Cyber-Physical Systems: Recent trends and state of the art applications. *Future Gener. Comput. Syst.* **2020**, *105*, 766–778. [\[CrossRef\]](#)
13. Hao, W.; Yang, T.; Yang, Q. Hybrid statistical-machine learning for real-time anomaly detection in industrial cyber-physical systems. *IEEE Trans. Autom. Sci. Eng.* **2021**, *20*, 32–46. [\[CrossRef\]](#)
14. Chaojun, G.; Yang, D.; Jirutitijaroen, P.; Walsh, W.M.; Reindl, T. Spatial load forecasting with communication failure using time-forward kriging. *IEEE Trans. Power Syst.* **2014**, *29*, 2875–2882. [\[CrossRef\]](#)
15. Mardia, K.V.; Goodall, C.; Redfern, E.J.; Alonso, F.J. The kriged Kalman filter. *Test* **1998**, *7*, 217–282. [\[CrossRef\]](#)
16. Tien Bui, D.; Tuan, T.A.; Klempe, H.; Pradhan, B.; Revhaug, I. Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. *Landslides* **2016**, *13*, 361–378. [\[CrossRef\]](#)
17. Ge, X.; Shi, L.; Fu, Y.; Mueen, S.; Zhang, Z.; He, H. Data-driven spatial-temporal prediction of electric vehicle load profile considering charging behavior. *Electr. Power Syst. Res.* **2020**, *187*, 106469. [\[CrossRef\]](#)
18. Hamdi, A.; Shaban, K.; Erradi, A.; Mohamed, A.; Rumi, S.K.; Salim, F.D. Spatiotemporal data mining: A survey on challenges and open problems. *Artif. Intell. Rev.* **2022**, *55*, 1441–1488. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Du, P.; Bai, X.; Tan, K.; Xue, Z.; Samat, A.; Xia, J.; Li, E.; Su, H.; Liu, W. Advances of four machine learning methods for spatial data handling: A review. *J. Geovis. Spat. Anal.* **2020**, *4*, 1–25. [\[CrossRef\]](#)
20. Nikparvar, B.; Thill, J.C. Machine learning of spatial data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 600. [\[CrossRef\]](#)
21. Bao, Y.; Huang, J.; Shen, Q.; Cao, Y.; Ding, W.; Shi, Z.; Shi, Q. Spatial-Temporal Complex Graph Convolution Network for Traffic Flow Prediction. *Eng. Appl. Artif. Intell.* **2023**, *121*, 106044. [\[CrossRef\]](#)
22. Wang, S.; Cao, J.; Yu, P. Deep learning for spatio-temporal data mining: A survey. *IEEE Trans. Knowl. Data Eng.* **2020**, *34*, 3681–3700. [\[CrossRef\]](#)
23. Bai, M.; Lin, Y.; Ma, M.; Wang, P.; Duan, L. PrePCT: Traffic congestion prediction in smart cities with relative position congestion tensor. *Neurocomputing* **2021**, *444*, 147–157. [\[CrossRef\]](#)
24. Zhao, Z.; Zhang, M.; Chen, J.; Qu, T.; Huang, G.Q. Digital twin-enabled dynamic spatial-temporal knowledge graph for production logistics resource allocation. *Comput. Ind. Eng.* **2022**, *171*, 108454. [\[CrossRef\]](#)
25. Li, S.; Liu, J.; Pan, Z.; Lv, S.; Si, S.; Sun, L. Anomaly Detection based on Robust Spatial-temporal Modeling for Industrial Control Systems. In Proceedings of the 2022 IEEE 19th International Conference on Mobile Ad Hoc and Smart Systems (MASS), Denver, CO, USA, 19–23 October 2022; pp. 355–363.
26. Zahran, B.; Hussaini, A.; Ali-Gombe, A. Security of IT/OT Convergence: Design and Implementation Challenges. *arXiv* **2023**, arXiv:2302.09426.
27. Alwan, A.A.; Ciupala, M.A.; Brimicombe, A.J.; Ghorashi, S.A.; Baravalle, A.; Falcarin, P. Data quality challenges in large-scale cyber-physical systems: A systematic review. *Inf. Syst.* **2022**, *105*, 101951.
28. Luo, Y.; Xiao, Y.; Cheng, L.; Peng, G.; Yao, D. Deep learning-based anomaly detection in cyber-physical systems: Progress and opportunities. *ACM Comput. Surv. (CSUR)* **2021**, *54*, 1–36. [\[CrossRef\]](#)
29. Iqbal, R.; Maniak, T.; Doctor, F.; Karyotis, C. Fault detection and isolation in industrial processes using deep learning approaches. *IEEE Trans. Ind. Inform.* **2019**, *15*, 3077–3084. [\[CrossRef\]](#)
30. Cinar, E.; Kalay, S.; Saricicek, I. A Predictive Maintenance System Design and Implementation for Intelligent Manufacturing. *Machines* **2022**, *10*, 1006. [\[CrossRef\]](#)
31. Castano, F.; Cruz, Y.J.; Villalonga, A.; Haber, R.E. Data-driven insights on time-to-failure of electromechanical manufacturing devices: A procedure and case study. *IEEE Trans. Ind. Inform.* **2022**, *19*, 7190–7200. [\[CrossRef\]](#)

32. Ghasemkhani, B.; Aktas, O.; Birant, D. Balanced K-Star: An Explainable Machine Learning Method for Internet-of-Things-Enabled Predictive Maintenance in Manufacturing. *Machines* **2023**, *11*, 322. [\[CrossRef\]](#)
33. Ding, C.; Sun, S.; Zhao, J. MST-GAT: A multimodal spatial-temporal graph attention network for time series anomaly detection. *Inf. Fusion* **2023**, *89*, 527–536.
34. Wu, W.; Shen, L.; Zhao, Z.; Li, M.; Huang, G.Q. Industrial IoT and long short-term memory network enabled genetic indoor tracking for factory logistics. *IEEE Trans. Ind. Inform.* **2022**, *18*, 7537–7548. [\[CrossRef\]](#)
35. Cruz, Y.J.; Rivas, M.; Quiza, R.; Haber, R.E.; Castaño, F.; Villalonga, A. A two-step machine learning approach for dynamic model selection: A case study on a micro milling process. *Comput. Ind.* **2022**, *143*, 103764.
36. Kim, T.Y.; Cho, S.B. Predicting residential energy consumption using CNN-LSTM neural networks. *Energy* **2019**, *182*, 72–81. [\[CrossRef\]](#)
37. Zhou, J.; Dai, H.N.; Wang, H.; Wang, T. Wide-attention and deep-composite model for traffic flow prediction in transportation cyber-physical systems. *IEEE Trans. Ind. Inform.* **2020**, *17*, 3431–3440. [\[CrossRef\]](#)
38. Zhang, Y.; Shi, X.; Zhang, H.; Cao, Y.; Terzija, V. Review on deep learning applications in frequency analysis and control of modern power system. *Int. J. Electr. Power Energy Syst.* **2022**, *136*, 107744. [\[CrossRef\]](#)
39. Carvalho, T.P.; Soares, F.A.; Vita, R.; Francisco, R.d.P.; Basto, J.P.; Alcalá, S.G. A systematic literature review of machine learning methods applied to predictive maintenance. *Comput. Ind. Eng.* **2019**, *137*, 106024. [\[CrossRef\]](#)
40. Rowe, F. What literature review is not: Diversity, boundaries and recommendations. *Eur. J. Inf. Syst.* **2014**, *23*, 241–255. [\[CrossRef\]](#)
41. Hatcher, W.G.; Yu, W. A survey of deep learning: Platforms, applications and emerging research trends. *IEEE Access* **2018**, *6*, 24411–24432. [\[CrossRef\]](#)
42. Liang, F.; Yu, W.; Liu, X.; Griffith, D.; Golmie, N. Toward Edge-Based Deep Learning in Industrial Internet of Things. *IEEE Internet Things J.* **2020**, *7*, 4329–4341. [\[CrossRef\]](#)
43. Xu, H.; Liu, X.; Yu, W.; Griffith, D.; Golmie, N. Reinforcement Learning-Based Control and Networking Co-Design for Industrial Internet of Things. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 885–898. [\[CrossRef\]](#)
44. Gao, N.; Xue, H.; Shao, W.; Zhao, S.; Qin, K.K.; Prabowo, A.; Rahaman, M.S.; Salim, F.D. Generative adversarial networks for spatio-temporal data: A survey. *ACM Trans. Intell. Syst. Technol. (TIST)* **2022**, *13*, 1–25. [\[CrossRef\]](#)
45. Chen, F.; Chen, J.; Huang, W.; Chen, S.; Huang, X.; Jin, L.; Jia, J.; Zhang, X.; An, C.; Zhang, J.; et al. Westerlies Asia and monsoonal Asia: Spatiotemporal differences in climate change and possible mechanisms on decadal to sub-orbital timescales. *Earth-Sci. Rev.* **2019**, *192*, 337–354. [\[CrossRef\]](#)
46. Shao, W.; Salim, F.D.; Gu, T.; Dinh, N.T.; Chan, J. Traveling officer problem: Managing car parking violations efficiently using sensor data. *IEEE Internet Things J.* **2017**, *5*, 802–810. [\[CrossRef\]](#)
47. Yang, Y.; Zhang, Y.; Cai, Y.D.; Lu, Q.; Koric, S.; Shao, C. Hierarchical measurement strategy for cost-effective interpolation of spatiotemporal data in manufacturing. *J. Manuf. Syst.* **2019**, *53*, 159–168. [\[CrossRef\]](#)
48. Feng, S.; Fan, F. Developing an Enhanced Ecological Evaluation Index (EEEI) Based on Remotely Sensed Data and Assessing Spatiotemporal Ecological Quality in Guangdong–Hong Kong–Macau Greater Bay Area, China. *Remote Sens.* **2022**, *14*, 2852. [\[CrossRef\]](#)
49. Kupilik, M.; Witmer, F. Spatio-temporal violent event prediction using Gaussian process regression. *J. Comput. Soc. Sci.* **2018**, *1*, 437–451. [\[CrossRef\]](#)
50. Rumi, S.K.; Salim, F.D. Modelling regional crime risk using directed graph of check-ins. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, Virtual Event, 19–23 October 2020; pp. 2201–2204.
51. Pan, Z.; Liang, Y.; Wang, W.; Yu, Y.; Zheng, Y.; Zhang, J. Urban traffic prediction from spatio-temporal data using deep meta learning. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 1720–1730.
52. Heikinheimo, V.; Di Minin, E.; Tenkanen, H.; Hausmann, A.; Erkkonen, J.; Toivonen, T. User-generated geographic information for visitor monitoring in a national park: A comparison of social media data and visitor survey. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 85. [\[CrossRef\]](#)
53. Lin, J.; Yu, W.; Yang, X.; Yang, Q.; Fu, X.; Zhao, W. A Novel Dynamic En-Route Decision Real-Time Route Guidance Scheme in Intelligent Transportation Systems. In Proceedings of the 2015 IEEE 35th International Conference on Distributed Computing Systems, Columbus, OH, USA, 29 June–2 July 2015; pp. 61–72. [\[CrossRef\]](#)
54. Liang, F.; Qian, C.; Yu, W.; Griffith, D.; Golmie, N. Survey of Graph Neural Networks and Applications. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 9261537. [\[CrossRef\]](#)
55. Guo, G.; Yuan, W. Short-term traffic speed forecasting based on graph attention temporal convolutional networks. *Neurocomputing* **2020**, *410*, 387–393. [\[CrossRef\]](#)
56. Ma, X.; Zhong, H.; Li, Y.; Ma, J.; Cui, Z.; Wang, Y. Forecasting transportation network speed using deep capsule networks with nested LSTM models. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 4813–4824. [\[CrossRef\]](#)
57. Yan, B.; Wang, G.; Yu, J.; Jin, X.; Zhang, H. Spatial-temporal chebyshev graph neural network for traffic flow prediction in iot-based its. *IEEE Internet Things J.* **2021**, *9*, 9266–9279. [\[CrossRef\]](#)
58. Han, L.; Du, B.; Sun, L.; Fu, Y.; Lv, Y.; Xiong, H. Dynamic and multi-faceted spatio-temporal deep learning for traffic speed forecasting. In Proceedings of the 27th ACM SIGKDD conference on Knowledge Discovery & Data Mining, Virtual Event, 14–18 August 2021; pp. 547–555.

59. Tian, C.; Chan, W.K. Spatial-temporal attention wavenet: A deep learning framework for traffic prediction considering spatial-temporal dependencies. *IET Intell. Transp. Syst.* **2021**, *15*, 549–561. [\[CrossRef\]](#)
60. Li, W.; Wang, X.; Zhang, Y.; Wu, Q. Traffic flow prediction over multi-sensor data correlation with graph convolution network. *Neurocomputing* **2021**, *427*, 50–63. [\[CrossRef\]](#)
61. Lin, L.; Li, W.; Zhu, L. Network-wide multi-step traffic volume prediction using graph convolutional gated recurrent neural network. *arXiv* **2021**, arXiv:2111.11337.
62. Cai, K.; Li, Y.; Fang, Y.P.; Zhu, Y. A deep learning approach for flight delay prediction through time-evolving graphs. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 11397–11407. [\[CrossRef\]](#)
63. Peng, H.; Du, B.; Liu, M.; Liu, M.; Ji, S.; Wang, S.; Zhang, X.; He, L. Dynamic graph convolutional network for long-term traffic flow prediction with reinforcement learning. *Inf. Sci.* **2021**, *578*, 401–416. [\[CrossRef\]](#)
64. Kong, X.; Zhu, B.; Shen, G.; Workneh, T.C.; Ji, Z.; Chen, Y.; Liu, Z. Spatial-temporal-cost combination based taxi driving fraud detection for collaborative internet of vehicles. *IEEE Trans. Ind. Inform.* **2021**, *18*, 3426–3436. [\[CrossRef\]](#)
65. Karim, M.M.; Li, Y.; Qin, R.; Yin, Z. A Dynamic Spatial-Temporal Attention Network for Early Anticipation of Traffic Accidents. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 9590–9600. [\[CrossRef\]](#)
66. Diao, C.; Zhang, D.; Liang, W.; Li, K.C.; Hong, Y.; Gaudiot, J.L. A Novel Spatial-Temporal Multi-Scale Alignment Graph Neural Network Security Model for Vehicles Prediction. *IEEE Trans. Intell. Transp. Syst.* **2022**, *24*, 904–914. [\[CrossRef\]](#)
67. Chen, D.; Lv, Z. Artificial intelligence enabled Digital Twins for training autonomous cars. *Internet Things Cyber-Phys. Syst.* **2022**, *2*, 31–41. [\[CrossRef\]](#)
68. Liang, W.; Li, Y.; Xie, K.; Zhang, D.; Li, K.C.; Souiri, A.; Li, K. Spatial-temporal aware inductive graph neural network for C-ITS data recovery. *IEEE Trans. Intell. Transp. Syst.* **2022**, *2022*, 1–12. [\[CrossRef\]](#)
69. Kong, X.; Zhou, W.; Shen, G.; Zhang, W.; Liu, N.; Yang, Y. Dynamic graph convolutional recurrent imputation network for spatiotemporal traffic missing data. *Knowl.-Based Syst.* **2023**, *261*, 110188. [\[CrossRef\]](#)
70. Malawade, A.V.; Yu, S.Y.; Hsu, B.; Muthirayan, D.; Khargonekar, P.P.; Al Faruque, M.A. Spatiotemporal scene-graph embedding for autonomous vehicle collision prediction. *IEEE Internet Things J.* **2022**, *9*, 9379–9388. [\[CrossRef\]](#)
71. Sun, B.; Zhao, D.; Shi, X.; He, Y. Modeling global spatial-temporal graph attention network for traffic prediction. *IEEE Access* **2021**, *9*, 8581–8594. [\[CrossRef\]](#)
72. Yang, J.; Yue, Z. Learning Hierarchical Spatial-Temporal Graph Representations for Robust Multivariate Industrial Anomaly Detection. *IEEE Trans. Ind. Inform.* **2022**, *2022*, 1–12. [\[CrossRef\]](#)
73. Liu, Y.; Zhao, Z.; Zhang, S.; Jung, U. Identification of abnormal processes with spatial-temporal data using convolutional neural networks. *Processes* **2020**, *8*, 73. [\[CrossRef\]](#)
74. Li, T.; Zhao, Z.; Sun, C.; Yan, R.; Chen, X. Hierarchical attention graph convolutional network to fuse multi-sensor signals for remaining useful life prediction. *Reliab. Eng. Syst. Saf.* **2021**, *215*, 107878. [\[CrossRef\]](#)
75. Yang, C.; Zhou, K.; Liu, J. SuperGraph: Spatial-temporal graph-based feature extraction for rotating machinery diagnosis. *IEEE Trans. Ind. Electron.* **2021**, *69*, 4167–4176. [\[CrossRef\]](#)
76. Zhang, X.; Long, Z.; Peng, J.; Wu, G.; Hu, H.; Lyu, M.; Qin, G.; Song, D. Fault Prediction for Electromechanical Equipment Based on Spatial-Temporal Graph Information. *IEEE Trans. Ind. Inform.* **2022**, *19*, 1413–1424. [\[CrossRef\]](#)
77. Shcherbakov, M.; Sai, C. A hybrid deep learning framework for intelligent predictive maintenance of Cyber-Physical Systems. *ACM Trans. Cyber-Phys. Syst. (TCPS)* **2022**, *6*, 1–22. [\[CrossRef\]](#)
78. Xiong, Q.; Zhang, J.; Wang, P.; Liu, D.; Gao, R.X. Transferable two-stream convolutional neural network for human action recognition. *J. Manuf. Syst.* **2020**, *56*, 605–614. [\[CrossRef\]](#)
79. Zheng, T.; Liu, C.; Liu, B.; Wang, M.; Li, Y.; Wang, P.; Qin, X.; Guo, Y. Scene recognition model in underground mines based on CNN-LSTM and spatial-temporal attention mechanism. In Proceedings of the 2020 International Symposium on Computer, Consumer and Control (IS3C), Taichung City, Taiwan, 13–16 November 2020; pp. 513–516.
80. Jia, M.; Xu, D.; Yang, T.; Liu, Y.; Yao, Y. Graph convolutional network soft sensor for process quality prediction. *J. Process Control* **2023**, *123*, 12–25. [\[CrossRef\]](#)
81. Li, X.; Yi, X.; Liu, Z.; Liu, H.; Chen, T.; Niu, G.; Yan, B.; Chen, C.; Huang, M.; Ying, G. Application of novel hybrid deep learning model for cleaner production in a paper industrial wastewater treatment system. *J. Clean. Prod.* **2021**, *294*, 126343. [\[CrossRef\]](#)
82. Guo, J.; Han, M.; Zhan, G.; Liu, S. A Spatio-Temporal Deep Learning Network for the Short-Term Energy Consumption Prediction of Multiple Nodes in Manufacturing Systems. *Processes* **2022**, *10*, 476. [\[CrossRef\]](#)
83. Bampoula, X.; Siaterlis, G.; Nikolakis, N.; Alexopoulos, K. A deep learning model for predictive maintenance in cyber-physical production systems using lstm autoencoders. *Sensors* **2021**, *21*, 972. [\[CrossRef\]](#) [\[PubMed\]](#)
84. Chen, C.; Hui, Q.; Xie, W.; Wan, S.; Zhou, Y.; Pei, Q. Convolutional Neural Networks for forecasting flood process in Internet-of-Things enabled smart city. *Comput. Netw.* **2021**, *186*, 107744. [\[CrossRef\]](#)
85. Jiang, Y.; Niu, S.; Zhang, K.; Chen, B.; Xu, C.; Liu, D.; Song, H. Spatial-temporal graph data mining for iot-enabled air mobility prediction. *IEEE Internet Things J.* **2021**, *9*, 9232–9240. [\[CrossRef\]](#)
86. Pan, J. Physical Integrity Attack Detection of Surveillance Camera with Deep Learning based Video Frame Interpolation. In Proceedings of the 2019 IEEE International Conference on Internet of Things and Intelligence System (IoTIS), Bali, Indonesia, 5–7 November 2019; pp. 79–85.

87. Zhang, T.; Wang, Z.; Zeng, Y.; Wu, X.; Huang, X.; Xiao, F. Building Artificial-Intelligence Digital Fire (AID-Fire) system: A real-scale demonstration. *J. Build. Eng.* **2022**, *62*, 105363. [[CrossRef](#)]
88. Pan, X.; Gao, X.; Wang, H.; Zhang, W.; Mu, Y.; He, X. Temporal-based Swin Transformer network for workflow recognition of surgical video. *Int. J. Comput. Assist. Radiol. Surg.* **2023**, *18*, 139–147. [[CrossRef](#)]
89. Ge, W.; Huh, J.W.; Park, Y.R.; Lee, J.H.; Kim, Y.H.; Zhou, G.; Turchin, A. Using deep learning with attention mechanism for identification of novel temporal data patterns for prediction of ICU mortality. *Inform. Med. Unlocked* **2022**, *29*, 100875. [[CrossRef](#)]

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