Identifying the most suitable machine learning approach for a road digital twin - a systematic literature review

Identifying the most suitable machine learning approach for a road digital twin; a systematic literature review

Kun Chen[1, 2, +], Mehran Eskandari Torbaghan[1], Mingjie Chu[3], Long Zhang[3], Alvaro Garcia-Hernández[2]
[1]*School of Engineering, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom*
[2]*Nottingham Transportation Engineering Centre, Faculty of Engineering, The University of Nottingham, NG7 2RD, United Kingdom*
[3]*Department of Electrical and Electronic Engineering, University of Manchester, Manchester, UK*
[+]*corresponding author (KXC005@student.bham.ac.uk)*

Abstract
Road infrastructure systems have been suffering from ineffective maintenance strategies, exaggerated by budget restrictions. A more holistic road asset management approach enhanced by data-informed decision making through effective condition assessment, distress detection, future condition predictions can significantly enhance maintenance planning, prolonging asset life. Recent technology innovations such as Digital Twins have great potentials to enable the needed approach for road condition predictions and a proactive asset management. To this end, machine learning techniques have also demonstrated convincing capabilities in solving engineering problems. However, none of them has been considered specifically within digital twins context. There is therefore a need to review and identify appropriate approaches for the usage of machine learning techniques within road digital twins.

This paper provides a systematic literature review of machine learning algorithms used for road condition predictions and discusses findings within the road digital twin framework. The results show that existing machine learning approaches are to some extent, suitable and mature to stipulate successful road digital twin development. Moreover, the review whilst identifying gaps in the literature, indicates several considerations and recommendations required on the journey to road digital twins, and suggests multiple future research directions based on the review summaries of machine learning capabilities.

1. Background

The maintenance and management of road infrastructure is of great significance in any country for achieving a sustainable social and economic development. Therefore, there is an ongoing research question on how to achieve high quality for the operation and maintenance of roads within an optimised asset management strategy when facing insufficient funding (Burningham & Stankevich, 2005). In the UK, the cost of road maintenance activities for local roads, managed by local authorities, was up to £3.3 billion in the years 2016 to 2017 (Haylen, 2019). The maintenance activities are mostly carried out based on a reactive approach which involves heavy manual labour, enough to maintain the serviceability of roads. However, the reactive approach is not as cost-effective as the proactive approaches (UK House of Commons, 2019). One of the reasons that proactive maintenance strategies are not well adopted is a lack of a holistic management approach based on the reliable and accurate deterioration modelling and asset whole lifecycle data management (Bowden et al., 2006). The most recent trend in the field of infrastructure maintenance is leaning towards intelligent models informed by monitoring data for proactively optimised maintenance strategies (MJ DeJong et al., 2019). This requires consideration of (i) maintenance activities covering the whole lifecycle of the asset, (ii) relevant data that can be influential factors of asset performance, (iii) the quantified optimum balance of costs across different stakeholders including infrastructure owner, manager, and user. Due to the complexity of road types, materials, structures, and their degradation mechanisms shaped by large number of factors during their service life, it is extremely challenging to create precise maintenance strategies with a high level of granularity.

On the other hand, recent developments in digital technologies such as Internet of Things (IoT), Big Data, Sensorisation, Machine Learning (ML) or Deep Learning (DL), and Artificial Intelligence (AI) have received increasing attention in resolving civil engineering problems (Karimzadeh, 2020). Among both academics and practitioners, road engineers have recognised that a large amount of data from a network of sensors can provide continuous and useful information on road behaviour and performance and is able to provide a more comprehensive understanding of the road status if combined well with visualisation (D Trousdale, 2019; Steyn, 2020). Specifically, with the growing number of real-time data from various sources, data analytics approaches have been leveraged to perform improved road deterioration modelling and to enable prediction with high accuracy for road performance (Piryonesi, 2019). With regards to road performance prediction, various ML neural networks and different DL algorithms have been explored to predict the road performance based on a data-driven method (Amirhossein Hosseini, 2020; Choi & Do, 2020). As a result of the evolvement of these technologies, a novel concept of digital twin has recently become a popular research area for architecture, engineering and construction management and has shown a great potential to support thorough intelligent automated decision-making as a visual-aided tool for the asset whole lifecycle management and therefore to optimise operation and maintenance strategies (Macchi et al., 2018). The report from National Infrastructure Commission - Data for the Public Good provided suggestions for the UK government toward a digital infrastructure and development of a national digital twin was the key element (NIC, 2017). The Gemini principles published by the UK Digital Framework Task Group and the Centre for Digital Built Britain also addressed the need of digital twins (Bolton, 2018). Various applications of digital twin concept have been investigated for different assets at different levels. For example, at a city level for buildings, Lu et al (2020) demonstrated successful implementation of a building digital twin which integrates heterogeneous data sources and supports decision-making processes in operational and maintenance management. As for the bridges, and at an asset level Ye et al (2019) created a digital twin for structural health monitoring of bridges which provided several benefits including efficient query of relevant data, integrated capabilities of data processing and interpretation as well as a collaborative environment for various stages of a bridge project. Yu et al (2020) presented a highway tunnel performance prediction model based on digital twin concept and ML technique, which illustrated a reliable data-driven management method for preventive maintenance. Comprehensibly, road being one of the fundamental assets of any national infrastructure system, digital twins for roads would improve the accuracy of prediction and through that the effectiveness of any associated decision-making. However, no previous research has thoroughly investigated the potential applications of digital twin and its involved capabilities for roads, especially with a focus on road condition prediction.

One key enabling element of digital twin is the ML or DL capability for classifications and prediction of deterioration trends by taking advantage of the large volume of data available in the industry 4.0 age

(Fuller et al., 2020). For instance, ML algorithms have been used to efficiently and effectively perform classification tasks on remotely sensed images (Maxwell et al., 2018). With regards to road management, they have been utilised in classifying the types of roads (Saleh & Otoum, 2020), and different road surfaces (Bibi et al., 2021; Menegazzo & von Wangenheim, 2021a) as well as forecasting future performance of roads (Marcelino et al., 2021).

Currently, digital twin has various definitions provided by different organisations and sectors. VanDerHorn & Mahadevan (2021) summarised more than 46 literatures related to digital twin concept and provided a relatively generalised and consolidated definition as follows, "a virtual representation of a physical system (and its associated environment and processes) that is updated through the exchange of information between the physical and virtual systems". Based on these definitions, there are multiple requirements for a digital twin which are summarised in Table 1. Considering current practices for road maintenance and the core functionalities and elements of an ideal pavement management system (PMS), a more specific list of requirements for road digital twin together with a detailed mapping with ML algorithms are presented in Table 1.

Table 1 Requirements for a general digital twin, road digital twin mapped against ML capacities / applications

		Requirements											
		Synchronization between Physical & digital[1,8,9,11]	Data quality & granularity[1,2,17]	Covers the entire lifecycle[3,6,13,16]	Real-time data[1,9,14]	Virtual representation[2,4,9]	Performance prediction[5,9]	Informs decision-making[6,8,13,14]	Probabilistic prediction[1,3,13]	Common data storage[3,12]	Coping with big data[7,13,17]	Project prioritisation[3,15]	Maintenance planning[1,10]
General digital twin													
Road digital twin													
ML capacities / applications	Automated data analysis												
	Data cleaning & pre-processing												
	Efficient handing of multidimensional data												
	Data-driven condition assessment												
	Classification algorithms for road defect detection												
	Performance prediction												
	Reinforcement learning												
	Probabilistic simulation												

[1]Barricelli et al (2019) [2]Lu et al (2019) [3]Ye et al (2019) [4]VanDerHorn & Mahadevan (2021) [5]Abbott Dean (2014) [6]Macchi et al (2018) [7]Gandomi & Haider (2015) [8]Kaur et al (2020) [9]Jones et al (2020) [10]Errandonea et al (2020) [11]Zhang (2019) [12]Fuller et al (2020) [13]Tao et al (2019) [14]Hofmann & Branding (2019) [15]Boje et al (2020) [16]He et al (2021) [17]Chen et al (2014)

While various ML algorithms have been developed for modelling road deteriorations and to predict its conditions, none of these has been developed, tailored nor applied for a digital twin and its needs. Therefore, this research first takes a detailed look at the literature related to ML algorithms as a foundational basis for the road digital twin, in order to map the requirement of the digital twins against the ML capacities and applications. ML algorithms could have potential to assist and cover a broad range of road digital twin functionalities depending on the user requirements, such as traffic management and safety assurance. However, this paper then narrows down to describe and discuss the ML approaches and algorithms, identified through a systematic literature review, that could fit into a road digital twin to satisfy the requirements of road condition prediction applications.

2. Systematic review

A systematic literature review has been carried out to critically appraise and syntheses research findings on ML models for road conditions prediction. The systematic review described in this paper addressed the follow question:

- What are the most suitable ML algorithms or approaches to address the requirements of a road digital twin for the application of road condition prediction?

2.1 Definitions

The systematic literature review followed a search protocol which is based on the definitions provided in Table 2.

Table 2: Search protocol definitions

Term	Definition	Reference
Digital Twin	Digital Twin is a virtual representation of manufacturing elements such as personnel, products, assets and process definitions, a living model that continuously updates and changes as the physical counterpart changes to represent status, working conditions, product geometries and resource states in a synchronous manner.	(ISO/ISO-AWI 23247, 2019)
Machine Learning	Machine Learning is a broad term encompassing a number of methods that allow the investigator to learn from the data. These methods may permit large real-world databases to be more rapidly translated to inform decision making.	(Brnabic and Hess, 2021)
Road Prediction Model	Prediction Model predicts future road conditions or performances, and it's used for assessing and prioritising maintenance treatment type and timing, and estimating life-cycle costs	(Kargah-Ostadi et al., 2019)

2.2 The adopted search approach

The first step was to define the search keywords to identify relevant articles according to three categories (Subject, Technology and Functionality), as described in Table 3, to identify the potentially relevant articles in different databases. The sources used in this systematic literature review were Google Scholar, ResearchGate, SAGE Journal, ScienceDirect (Elsevier), Scopus, Taylor and Francis, Transport Research International Documentation (TRID), Web of Science, and library database at the Universities of Birmingham, Nottingham, and Manchester.

Table 3: Study Search Keywords

Subject	Technology	Functionality
"Road" OR "Roads" OR "Pavement" OR "Pavements" OR "Highway" OR "Highways"	"Machine Learning" OR "Deep Learning" OR "Digital Twin" OR "Digital Twins" OR "Artificial Intelligence"	"Performance" OR "Condition" OR "Prediction" OR "Forecast" OR "Forecasting"

A systematic literature review software (EPPI-Reviewer™ Web version) (UCL, 2019) facilitated the review and was used for screening, coding, analysing, and storing retrieved articles (Thomas and Brunton, 2010). In particular, ML based priority screening function within the EPPI-Reviewer™ was utilised to help screening the large volume of articles as part of the review process (Tsou et al., 2020). The adopted systematic literature review process is summarised in Figure 1.

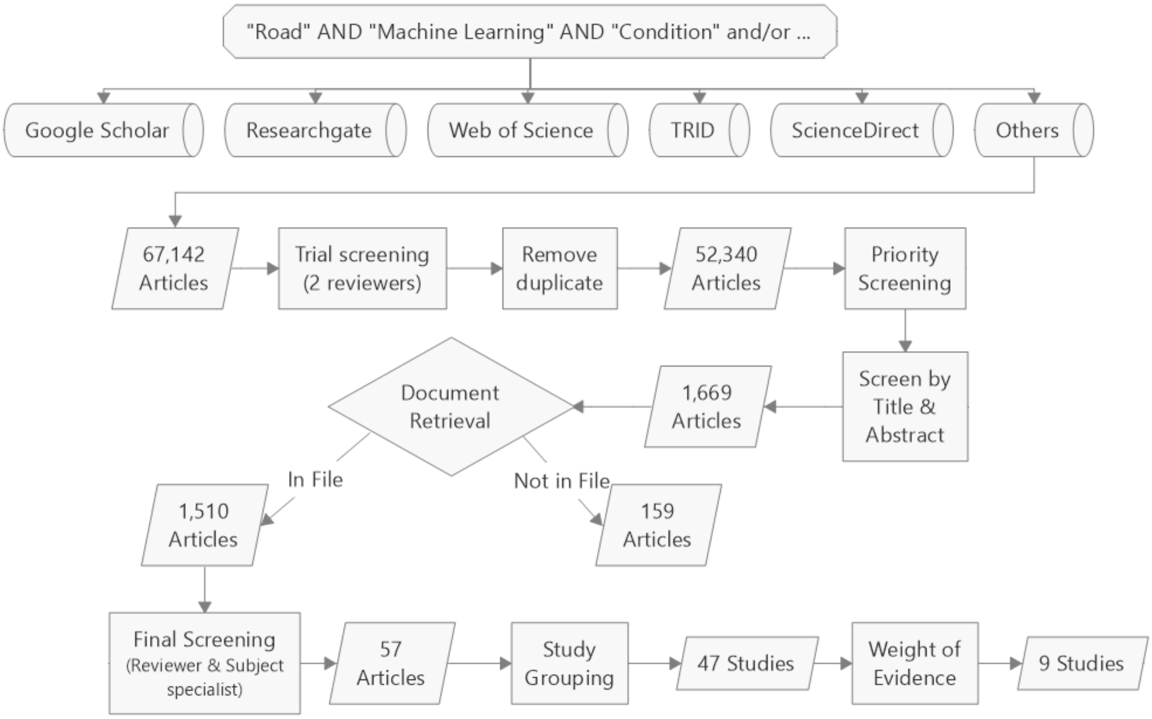

Figure 1. The systematic literature review process

2.3 Weight of evidence: assessing the quality of studies

A weight of evidence (WoE) framework has been developed and used to assess the quality and suitability of the included studies in three categories - research soundness, appropriateness of study design to review question, and relevance of evidence focus to review question (see Table 4). The studies finally included and considered were required to score high rating in at least two categories and one medium rating on the third. If the study satisfies at least one of the points of the description ratings, the corresponding ratings are applied.

Table 4 Weight of evidence (Gough et al., 2017)

Weight of Evidence	Ratings	Descriptions
Soundness	High	▪ Explicit and detailed methods and results on data collection, cleaning, and analysis ▪ Comprehensive model development, optimisation and evaluation based on ML principles ▪ Compelling ML algorithm prediction accuracy (R2 > 90%) ▪ Critical comparisons with other studies or models
	Medium	▪ Satisfactory methods and results on data collection, cleaning, and analysis ▪ Detailed model development and evaluation based on ML principles with satisfactory optimisation ▪ Satisfactory ML algorithm prediction accuracy (80% < R2 < 90%) ▪ Limited comparisons with other studies or models
	Low	▪ Unsatisfactory methods and results on the data used ▪ No optimisation on model development and evaluation ▪ Unsatisfactory ML algorithm prediction accuracy (R2 < 80%) ▪ No critical comparisons with other works
Appropriateness of study design	High	▪ Data comes from real world with extensive coverage. Frequency of collected data is high (weekly or monthly) ▪ The prediction model input parameters cover at least the following five categories: Traffic, Climate, Performance, Structure, Material, and makes performance prediction for one and more years ▪ Model outputs include one or more of the following aspects: distress condition, maintenance planning, and road lifecycle
	Medium	▪ Data comes from real world with limited coverage. Frequency of data collected is medium (yearly) ▪ The prediction model input parameters cover at least three of the below categories: Traffic, Climate, Performance, Structure, Material, and makes performance prediction for one and more years ▪ Model outputs include at least one of the following aspects: distress condition, maintenance planning, and road lifecycle
	Low	▪ Data comes from laboratory or simulation. Frequency of data collected is low (by-yearly) ▪ The prediction model input parameters cover one or two of the categories of Traffic, Climate, Performance, Structure, Material, and makes prediction for one year only ▪ Model outputs include only one of the following aspects: distress condition, maintenance planning, and road lifecycle
Relevance of the study focus	High	▪ The length of road sections chosen is greater than 100 meters ▪ The historical data covers more than 15 - 20 years
	Medium	▪ The length of road sections is less than 100 and greater than 50 ▪ The historical data covers more than 10 - 15 years
	Low	▪ The length of road sections is less than 50 ▪ The historical data coverage is less than 10 years

2.4 Synthesis of evidence

The data were synthesised to indicate the most suitable ML algorithms to best predict road condition considering within a road digital twin environment. More specifically, section 3 summaries different ML models, their inputs, output, data being used that would potentially have an impact on the model performance; and section 4 provides a further discussion on how these findings and insights could be contributed to the development of a road digital twin.

3. Summary of main findings

This section reports detailed findings on the selected studies after the Weight of Evidence stage. Although it is not within the scope of this review, to facilitate review findings summary, a potential road digital twin framework has been suggested (Figure 2) based on the large amount of literature reviewed, with a focus on road condition prediction within a road digital twin application.

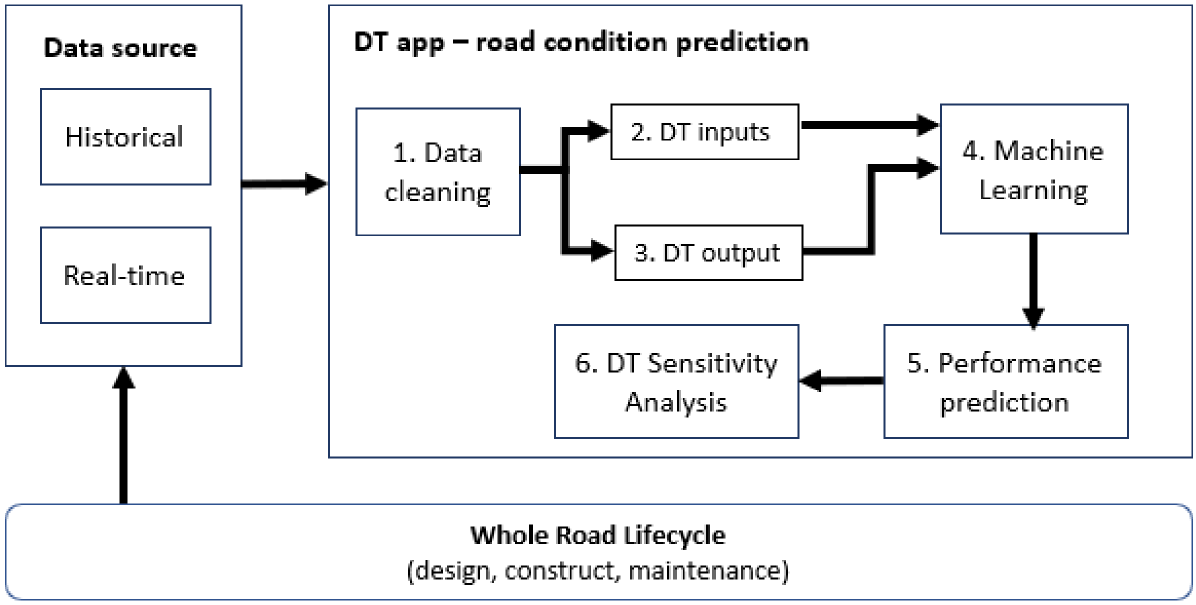

Figure 2. Digital Twin (DT) framework for road condition prediction

Nine high quality studies based on the WoE analysis were identified to answer the primary review question. The given ranking for the nine selected studies during the WoE stage are presented in Table 5.

Table 5 Details of WoE analysis on selected studies

Selected studies	Research soundness	Study design Appropriateness	Study focus relevance	Ranking Justification
(Tabatabaee et al., 2013)	High	High	High	• Clearly explained the utilised ML algorithm structure and the associated mathematical equations. • Collected large amount of data with over 10,000 records enabled by sensor data for over 15 years • Considered more than 5 categories of data sources including structure, performance, traffic, climate, material, and maintenance records and plans
(Ziari et al., 2015)	High	High	High	• Used advanced data processing method, i.e., group method of data handling (GMDH) and performed in-depth model optimisation with large scale. • Collected sources of data are broad, covering performance, structure, traffic, climate, and material • The collected data covers large spectrum of characteristics, for over 17 years
(Ziari et al., 2016)	High	High	High	• Used the same data as (Ziari et al., 2015), and provided clear explanation on the ML algorithm, including cross-validation method on the testing data
(Alharbi, 2018)	High	Medium	High	• Demonstrated clear ML algorithm training and testing process with loss functions • Collected annual data for over 17 years with over 1,000 records in total, aimed at multiple output predictions (roughness, cracking and rutting) • Rated Medium in the study design appropriateness criteria mainly due to the limited considered number of data sources • Conducted model sensitivity analysis to understand the parameters affecting the performance
(Abdelaziz et al., 2018)	High	Medium	High	• Provided clear hidden layer structure for the ML algorithm for model optimisation • Emphasised much on data processing and cleaning, as well as maintenance impacts, ensuring high quality data • Regarding model inputs, it mainly considered previous condition index as the fundamental information
(Fathi et al., 2019)	High	Medium	High	• Combined two ML algorithms and showed significant improvement with R^2 from 64% to 92%. • Collected data over 20 years across four climate zone in the US • It used cross-validation techniques in the ML model training, testing process to avoid overfitting issues in ML • Put an emphasis on material properties as data sources for prediction
(Marcelino et al., 2019)	High	Medium	High	• Combined two databases using transferring learning algorithms to make predictions with 99.5% accuracy for multi-year prediction • The types of considered data are performance, traffic, climate, structure
(Bukharin et al., 2021)	High	Medium	High	• Showed detailed ML model structure and evaluated model performance with existing baseline models • Used around 15,000 data records for over 20 years, and made multiple year predictions (e.g., 2-5 years) with the minimum accuracy of 98% • Considered different types of data including geometry, performance, traffic
(Gong et al., 2018)	Medium	High	High	• Combined data from LTPP data with the information from Mechanistic-Empirical Pavement Design Guide (MEPDG) to improve prediction • Provided clear illustration on deep neural network, optimised with adaptive moment estimation, and focused on regularisation of overfitting of the model. It was rated Medium in the research soundness mainly due to the relatively lower model performance (e.g., less than R2 = 90%)

All the nine studies contribute to the understanding and definition of the requirements of a road DT as well as the suitability of up-to-date ML techniques. DT is a digital replica of the physical entity across its lifecycle, and therefore its prediction capacity for the road condition is vital, which could be achieved through a ML based approach. Various findings of the identified studies are synthesised and described in the following subsections. The following sub-sections describe findings in further details on the key elements shown in Figure 2.

3.1 Model input parameters

The systematic literature review indicates that researchers have considered a broad spectrum of different types of categories of parameters that could have an impact on the pavement performance and have used them as ML prediction model inputs. To summarise, the considered input parameters fall under below categories: Pavement structure, material, historical performance, traffic, environment, as well as maintenance treatment records.

Six studies (Alharbi, 2018; Fathi et al., 2019; Gong et al., 2018; Marcelino et al., 2019; Ziari et al., 2015, 2016) have considered the total thickness of the pavement structure as one of the model inputs, the results of the studies suggest that the range of pavement thickness have an impact on pavement performances, with a particular emphasis on cracking. As for the material, asphalt mixture air voids and the measured asphalt content have been utilised by three studies to build prediction models (Fathi et al., 2019; Gong et al., 2018; Tabatabaee et al., 2013). Fathi et al (2019) reported that air voids are one of the important attributes that significantly affect fatigue cracking along with pavement age and thickness. Annual average precipitation has been mostly considered to account for climate impacts on pavement performance (Alharbi, 2018; Gong et al., 2018; Marcelino et al., 2019; Ziari et al., 2015, 2016). This is an important factor for pavement condition forecasting to include the impact of changing climate and on road deterioration. The reviewed ML models have proved its capacities to quantify the impact of climate factors. For traffic loadings, equivalent single axis load (ESALs) has been used as input for four studies (Bukharin et al., 2021; Tabatabaee et al., 2013; Ziari et al., 2015, 2016) whereas annual average daily traffic (AADT) also has been utilised in four studies (Gong et al., 2018; Marcelino et al., 2019; Ziari et al., 2015, 2016). It appears that most studies either considered ESALs or AADT and not both at the same time. However, Ziari et al (2015) considered both and presented that AADT has slightly more importance on pavement performance prediction than ESAL by 10%. In addition, some studies chose not to consider ESAL as it can be assumed to be equivalent to age (Abdelaziz et al., 2018). Historical pavement performance metrics have also been identified as useful model inputs by various studies. For instance, three studies used previous rutting values to predict future road condition (Abdelaziz et al., 2018; Alharbi, 2018; Gong et al., 2018). Furthermore, Abdelaziz et al. (2018) and Alharbi (2018) have investigated other road defects as road condition indicators such as international roughness index (IRI) and crack rating for model development. Most studies have concluded that initial condition data is the most significant factor that determines the model performance (Alharbi, 2018; Fathi et al., 2019; Gong et al., 2018; Tabatabaee et al., 2013; Ziari et al., 2015). Last but not least, although maintenance data is an equally important information and a useful additional data source, few studies have utilised the data due to the potential difficulty to access them. Among the chosen studies, only one study has considered maintenance treatments as an input for the prediction model and have produced satisfying results ($R^2 = 98\%$) (Tabatabaee et al., 2013). Abdelaziz et al. (2018) also analysed maintenance impacts on the IRI values before and after the treatment, but it has only been used to identify and evaluate the corrected IRI values instead of maintenance type itself being prepared as an input to the model. It would be useful to understand ML's capacity in handing maintenance effects when making road condition predictions.

3.2 Model output parameters

Four studies (Abdelaziz et al., 2018; Marcelino et al., 2019; Ziari et al., 2015, 2016) have analysed IRI as the prediction model target output, noting that it has been used by many world transportation agencies as an indicator for pavement performance as well as maintenance and rehabilitation initiation factor (Abdelaziz et al., 2018). Moreover, Tabatabaee et al. (2013) have predicted present serviceability index (PSI) which is highly correlated with the IRI. The rest of the studies have focused on the forecasting of the future condition of the road with specific distresses, and in particular, cracking (Alharbi, 2018; Bukharin et al., 2021; Fathi et al., 2019) and rutting (Alharbi, 2018; Gong et al., 2018). In addition, this review also found that the output of ML model for different applications could potentially be expanded to detect and locate road defects (Ahmadi et al., 2021; Alzraiee et al., 2021; Arya, Maeda, Ghosh, Toshniwal, & Sekimoto, 2021), measure road temperature or wetness (Milad et al., 2021; Morris & Yang, 2021), classify roads based on intelligent inertial sensors (Lee et al., 2021; Menegazzo & von Wangenheim, 2021b), prioritise and plan maintenance (Gao et al., 2021; Han et al., 2020), as well as determine the remaining service life of pavements (Citir et al., 2021; Karballaeezadeh et al., 2019; Nabipour et al., 2019) depending on the requirements of the road DT.

3.3 Data used for research

Various data have been explored and used to develop prediction models using ML technique with unique data characteristics, pre-processing methods, and training and testing data split percentage, which are described in the following sub-sections.

3.3.1 Database

Six studies (Abdelaziz et al., 2018; Fathi et al., 2019; Gong et al., 2018; Marcelino et al., 2019; Ziari et al., 2015, 2016) have used US Long-Term Pavement Performance (LTPP) (Highway Administration and Pavement Performance program, 2017) which is a public database that provides state-of-the-art pavements information. The other three studies (Alharbi, 2018; Bukharin et al., 2021; Tabatabaee et al., 2013) used data sources from the pavement management systems (PMSs) owned by the US state transportation agencies. In addition, two studies (Gong et al., 2018; Marcelino et al., 2019) combined LTPP data with US National Cooperative Highway Research Program (NCHRP) permanent deformation model and reports and Portuguese national road administration database accordingly to increase the total amount of data. The reason behind the LTPP database being extensively researched by the pavement scholars is because of its availability for public, and its comprehensive coverage of various types and categories of data, such as structure and material, climate, traffic, and performance.

The number of considered road sections and data records also vary significantly for different studies. Fathi et al. (2019), Gong et al. (2018), and Ziari et al. (2015, 2016) used data from 26 to 88 road sections and obtained over 200 records for ML model development. Meanwhile Abdelaziz et al. (2018), and Alharbi (2018) collected data from around 500 sections resulting in average of 1,792 records to train and test the ML algorithms. However, other researchers (Bukharin et al., 2021; Marcelino et al., 2019; Tabatabaee et al., 2013) utilised over 14,000 data entries for preparation of next stages such as data pre-processing, model development and validation. The number of years covered in the utilised data ranged from five to 31 years.

3.3.2 Data pre-processing

Road condition data often includes various measurement errors and missing values, and hence it is essential to perform data pre-processing in order to achieve a relatively complete, clean, and high-quality data set (Ziari et al., 2015). This would help ML algorithms to better identify and understand the data and to eventually resolve the classification and regression problems for future road condition predictions.

Four studies (Abdelaziz et al., 2018; Bukharin et al., 2021; Ziari et al., 2015, 2016) have particularly addressed the issue of missing data. The methods used to fill missing values were 1) linear interpolation, 2) removal 3) regression model. In addition, Bukharin et al. (2021) performed removal on irrational behaviours of the crack, ride, or rut ratings data. To increase the accuracy and the large variance in the collected data, Ziari *et al*. (2015) used data normalisation technique to prevent bias due to large differences between minimum and maximum values in datasets.

3.3.3 Split on training and testing data

Before developing any ML model, it is important to prepare specific data for training and testing purposes. The size of training data in the considered studies varies from 58% (Bukharin et al., 2021) to 90% (Ziari et al., 2016). However, it can be concluded that the most common practice has been using 80% of the data for training and 20% of the testing, which is the case for four studies (Fathi et al., 2019; Gong et al., 2018; Marcelino et al., 2019; Ziari et al., 2015).

3.4 ML algorithms and configurations

Among the nine chosen studies, two main approaches have been explored to achieve satisfactory prediction model results. One approach was to use single ML algorithm to train and learn from the data, while alternatively a combined ML algorithms method was used for model development. For the usage of single ML algorithm, neural networks with various neuron configurations in hidden layers have been adopted (Abdelaziz et al., 2018; Alharbi, 2018; Gong et al., 2018; Ziari et al., 2015). Among these studies, Ziari et al. (2015) with a three-layer neuron structure achieved the highest prediction accuracy (R^2 = 96.8% and Root Mean Squared Error (RMSE) = 0.14). In addition to neural networks, other algorithms such as support vector machines with Pearson VII Universal kernel (Ziari et al., 2016) and boosting-based learning algorithm - TrAdaBoost (Marcelino et al., 2019) have been utilised and also produced a very high prediction accuracy (in average R^2 > 85%). As for the combined ML algorithms approach, it has resulted in a higher accuracy compared to the single ML algorithms (in average R^2 > 95%). As part of this combined ML approaches, the output of one ML algorithm model was used as an inputs into another ML algorithm with the purpose of improving the overall prediction accuracy. The explored algorithms are support vector machine with recurrent neural network (RNN) (Tabatabaee et al., 2013); random forest with artificial neural network (ANN) (Fathi et al., 2019); and long-term short memory neural network plus ANN (Bukharin et al., 2021). Detailed ML algorithms and model accuracy for all nine studies are presented in Table 6 below.

Table 6: Summary on ML algorithms and relevant accuracies

Author(s)	ML Technique(s)	Model Accuracy
Tabatabaee (2013)	Support Vector Classifier + Recurrent Neural Network	R^2 = 98% RMSE = 0.135
Ziari (2016)	Support Vector Regression	R^2 = 91.69% RMSE = 0.2259
Ziari (2015)	Artificial Neural Network + Group Method of Data Handling (GMDH)	**Short term (1 year):** R^2 = 96.8% RMSE = 0.14 **Short term (2 years):** R^2 = 97.2% RMSE = 0.167 **Long term (lifecycle):** R^2 = 97.9% RMSE = 0.377
Marcelino (2020)	Boosting - TrAdaBoost algorithm for regression	R^2 = 99.5%
Gong (2018)	Deep neural networks	R^2 = 86.7% RMSE = 1.403
Fathi (2019)	Random Forest + ANN	R^2 = 91% RMSE = 10.9
Bukharin (2021)	LSTM + ANN	Year 1: R^2 = 99% Year 2: R^2 = 98% Year 3: R^2 = 98% Year 4: R^2 = 99% Year 5: R^2 = 98%
Alharbi (2018)	ANN	R^2 = 92% RMSE = 8.42
Abdelaziz (2020)	ANN	R^2 = 75%*

*This article was selected despite a lower overall accuracy than the threshold defined in the weight of evidence stage because the author presented the model goodness-of-fit with 86% in the paper

3.5 Types of Prediction and Accuracy

From the perspective of different prediction duration, most studies produced one year or one step prediction, which is a basic capability of any ML models. But several studies have developed the ML algorithm for the use of multi-year or multi-step predictions. For example, Ziari et al. (2015) forecasted the IRI values for the next one year (R^2 = 96.8%), two years (R^2 = 97.2%) and the whole lifecycle, i.e., three or over three years, (R^2 = 97.9%). Moreover, the study conducted by Bukharin et al. (2021) produced next 1 – 5 years condition predictions on crack ratings with a continuous high accuracy (in average R^2 = 98.5%). Similarly, (Marcelino et al., 2019) predicted the next one to four years with relatively little loss in performance (in average R^2 = 99.5%). One insight worth noting is the use of transfer learning which enables the model to learn and extract the knowledge from multiple different databases and thereby increasing the overall prediction capacity.

3.6 Model Sensitivity analysis

Sensitivity analysis is another key component for any ML model development and analysis process. Five studies have provided a relatively in-depth analysis on the importance of each input parameter to the model output (Alharbi, 2018; Fathi et al., 2019; Gong et al., 2018; Tabatabaee et al., 2013; Ziari et al., 2015). Table 7 shows a summary of the most important input parameters ranking from #1 (most relevant) to #5 (least relevant) for different considered ML models.

Table 7: Summary for the sensitivity analysis for the considered ML models

No.	Study	ML Algorithm(s)	Model output	Input (importance)				
				#1	#2	#3	#4	#5
1	Tabatabaee et al. (2013)	Support Vector Classifier + RNN	Next year's PSI	PSI	Maintenance is required	Structure	Season	Age
2	Ziari et al. (2015)	ANN + Group Method of Data Handling (GMDH)	IRI	Annual Average Precipitation	AADT	Annual Average Daily Truck Traffic	Pavement thickness	ESAL
3	Gong et al. (2018)	Deep neural networks	Rutting	Predicted rutting in the asphalt concrete layer	Air voids in asphalt concrete	Predicted rutting in the granular base layer	AADT	Annual average precipitation
4	Fathi et al. (2019)	Random Forest + ANN	Alligator deterioration index (ADI)	Age	Asphalt content of the mix	Voids in mineral aggregate (VMA)	Pavement Thickness	Air Voids
5	Alharbi (2018)	ANN	Riding index	Age	Temperature	Previous IRI	Thickness	Precipitation
			Cracking index	Temperature	Previous cracking	Age	Structural number	N/A
			Rutting index	Temperature	Thickness	Age	Previous rutting	N/A

4. Discussions

Based on the findings of systematic literature review, it can be understood that various databases, data analytics techniques and ML algorithms have been developed, trailed, and tested for modelling road deterioration and to predict the road condition. Having said that, none of the existing data analysis and ML algorithms has been developed, tailored nor applied for road DT context to accommodating its needs. Therefore, this section explores further on the review findings and discusses the potential applications of different types of ML algorithms, their configurations as well as the optimisations that could be applied within a DT framework especially formed on the basis of the high-level requirements defined by the Centre for Digital Built Britain (CDBB) (Bolton, 2018) as well as multiple definitions and requirements summarised by VanDerHorn and Mahadeva (2021).

4.1 Data collection and analysis

As described in Figure 2, data source module contains both historical and real-time data, using multiple sources and types of data are important requirements for a road DT to ensure high accuracy in showing the status of the asset and predicting its future condition. Possible data sources are historical road inspection surveys, existing public pavement performance database (e.g., LTPP (FHWA, 1995)), public image datasets for road damage analysis (Arya, Maeda, Ghosh, Toshniwal, Omata, et al., 2021; Atikur, 2020; Maciej Serda, 2013; Varma et al., 2018), ground penetrating radar (GPR) surveys, data from existing pavement management systems, IoT devices, smartphones, cameras, and intelligent sensors based on distributed acoustic sensing (DAS) which enable the collection of a broad range of types of data (Liehr et al., 2019). These various sources of data can then be fed into the road DT for further processing, analysing, and modelling for the specific applications of the DT.

It also can be inferred from this review that various types of ML algorithms are capable of being the enabler of a road DT regarding future condition predictions. However, all selected studies have focused on time static data which is collected at a fixed time or frequency. Large amount of data as well as a more frequent data (real-time or near real-time) is expected to be an essential part of a DT according to various definitions of the concept (Abramovici et al., 2017; Demkovich et al., 2018; Negri et al., 2017). The higher sampling frequency could potentially result in the improvement on the prediction accuracy of different ML methods and possibly help to identify pavement condition changes at an early stage although it may not be easy to determine a sampling ratio in practice due to variable factors influencing road performance (e.g., different ages of the roads and different road conditions). However, the higher sampling could lead to big data, which further requires large storage and computations. Therefore, there is a need for further research on ML pavement performance prediction model development using big data concept and IoT platforms (Steyn, 2020) to understand the capabilities of different ML algorithms to deal with higher frequency data collection mechanisms, especially on the trade-off between the accuracy and sampling frequency (and intervals).

4.2 Data characteristics and ML algorithms choice

As it is mentioned in the definition and DT general requirements summarised in Table 1, DT is built on data, while the data itself plays a decisive role in the development of a DT. Data pre-processing and cleaning remain to be a fundamental factor in achieving high quality data and subsequently reliable ML prediction modelling in a road DT scenario where data could become larger in volume, variety, and velocity. Based on the review, an important finding is that the performance of ML algorithm itself is largely dependent on the data in terms of its structure, values, and patterns, where the same algorithm and configurations could result in completely different prediction accuracy if applied on a different dataset. This might be another potential research area where a road DT needs to have a process or algorithm in place to intelligently identify the most suitable ML algorithm(s) for a given dataset. The selection also depends on the needs of various users of the DT. The process or algorithm could provide a platform, like a data lake (Redeker et al., 2021), where all potential types of ML algorithms can be trialled to select the one with highest accuracy as the output. The findings from this review suggest that ANN, RNN, Long-short-term memory (LSTM) and a boosting algorithm AdaBoost.RT would be helpful in a road DT environment as it gives high prediction accuracies for not only short-term prediction (e.g., 1 or 2 years) but also long-term multiple year predictions (e.g., 3, 5 or 10 years) (Bukharin et al., 2021; Marcelino et al., 2019; Ziari et al., 2015).

Another fundamental key requirement and characteristic of a DT from Table 1 is its dependency on quality and variety of data, as well the data science pipeline. Having said that, none of the studies has considered to perform feature selection first on the identified model input parameters, which is an important component as part of the standard process within ML prediction model development. This could be one of the future research areas to improve the accuracy and reliability of the developed ML models.

4.3 Road lifecycle analysis based on DT

According to the defined road DT requirements in Table 1, the ability to communicate and reflect the characteristics and functionalities across the entire lifecycle within different phases of the corresponding physical asset is a crucial factor for DTs. Although the review result only presents the use of ML algorithms for road condition predictions, various ML or DL algorithms have been applied and could be further improved on other aspects of road asset management lifecycle. These applications could also be considered in the road DT development. For instance, ML algorithms such as Conventional Neural Network (CNN) can be used on image data to perform on-going structural condition health monitoring or assessment (Azimi and Pekcan, 2020) and to automatically detect different types of pavement distresses using k-nearest neighbour (kNN) algorithm (Du et al., 2020). In addition, to the application of DT for health monitoring, distress identification, and performance prediction, it can also be used for the maintenance treatments and rehabilitation strategy selections using DL or reinforcement learning algorithms, which has been proven more effective in optimising maintenance planning and budget allocations than traditional methods (Gao et al., 2021b; Yao et al., 2020).

While the management of the road during its operation phase is important, the benefit of using ML capabilities within a DT environment for other phases of a road lifecycle cannot be ignored. Therefore, future research directions could be on the development or improvement of DTs for road design, construction, and demolition. Furthermore, as an ideal DT should be able to reflect the whole lifecycle development of the physical asset, there is a great value in understanding the framework and processes of integrating the usage of different ML, DL or reinforcement learning algorithms at different stages of the lifecycle of the road DT that altogether could enable its purpose, and fully realise the asset values that would result in highest operation efficiency and the lowest cost for the asset owners and end-users.

4.4 Future pavement management system based on road DT

Furthering the discussion from Section 4.3, this paper identified a large number of articles on the applications of ML algorithms for various road applications. Therefore, this section focuses on discussing road DT applications in improving pavement management systems. As mentioned by Kulkarni et al (2003), the existing key elements of pavement management systems are 1) data collection and management; 2) pavement performance prediction; 3) economic analysis 4) priority evaluation 5) optimisation and 6) institutional issues. For each element, a specific road DT application can be developed, enabled by one or multiple ML algorithms. For example, a road DT for pavement health monitoring can be created using the data from multiple sensors such as pressure cells, deflectometers, strain gauges, thermocouples, moisture sensors, fibre-optic sensors, non-destructive testing surveys or other IoT devices, to enable a constant monitoring of the pavement. This in turn enables to detect the formation of road distresses such as internal cracks (di Graziano et al., 2020). ML or DL algorithms such as genetic expression programming models (Majidifard et al., 2020) have been developed to perform classification tasks to evaluate road surface conditions over time. With all the available data fed into the road DT, additional applications can be created to use the data leveraging that the identified ML or DL models in this paper can provide. The ML algorithms can be used to firstly create more precise and dynamic prediction model which self-validates, and updates based on constant data inputs. Secondly, with data from maintenance activities, construction history and asset owner policy, finance, and user cost, road DT can then be used to generate and analyse 'what-if' scenarios on maintenance planning. This will enhance asset maintenance prioritisation at network or project level, by achieving maintenance optimisation through taking into consideration the associated social and economic factors (Yao et al., 2020). Maintenance optimisation could be achieved by understanding the right timing to perform the most cost-effective maintenance treatment on the right location(s) of road network and sections with the highest priorities. However, to enable the innovation of road DTs, the sharing of various data across different departments and even organisations would be the number one challenge due to data privacy and data security issues (Marai et al., 2020). Aforementioned applications could be also presented in 3D pavement models via BIM-based PMS for improved visualisation compared to current 2D or GIS-based visualisation (D'amico et al., 2022).

4.5 Other factors

Road condition data often suffers poor quality regardless of its collection frequencies, while periodical inspection surveys often have missing values due to machine faults or human errors (Ziari et al., 2015). In addition, as sensor monitoring data also is subjected to some level of inaccuracies influenced by various factors and conditions (Oliveira et al., 2021), it would be necessary to take into consideration these uncertainties into the road DT framework. This can be done, considering a probabilistic approach for the DT environment, thereby achieving a more realistic and accurate model output. For instance, Yao et al. (2021) proposed a probabilistic ML model – Bayesian neural network to address the issue of uncertainty, which could be adopted within a DT context. Another possibility to address poor quality data could be combining physics-based modelling with ML algorithms to generate extra data from physics-based model to compliment the data quality issues faced normally by ML applications (Arias Chao et al., 2020; Willard et al., 2020).

It is worth mentioning that although the purpose of this systematic literature review was to be inclusive and avoid missing valuable relevant articles, given the fact that DT and ML are recent technological advancements, the articles before and including the year 2011 have not been considered in this review.

There are various functionalities for a road DT, depending on the user requirements, which might or might not be addressed by ML algorithms. For instance, a road DT could consider traffic and its management, and/or enhancing safety by minimising car accidents (variable speed limits) while ML has the potential to assist achieving these. However, this paper only focused on ML capacity for predicting road condition.

5. Conclusion

This study is the first systematic literature review summarising the ML algorithms used in road condition predictions, and specifically, attempting to identify the most suitable ML approach within a road DT framework considering the attributes and characteristics of DTs. From the review's results, it can be concluded that in order to fulfil the purpose of future performance prediction of a road, the corresponding DT should include at least the following categories of data inputs:

1) Existing performance index;
2) Existing distress condition;
3) Materials;
4) Structure;
5) Climate;
6) Traffic

The main data outputs of the ML models could be 1) Future performance index and 2) Future distress condition. In addition, this study comprehensively reviewed different types of ML algorithms and approaches applied for multiple functionalities that a road DT should include from the perspectives of road condition assessment, road defect detection, future performance prediction as well as maintenance planning and optimisations. This review has provided a thorough synthesis of findings on ML techniques used for pavement performance prediction. Several articles have been identified to show promising adaptability with high ML prediction capability that could fulfil the road DT requirements.

From this systematic review, it can be concluded that the ML algorithms which have recently been applied in resolving various aspects of pavement management problems are the foundational building blocks for a road DT and they do certainly demonstrate enough maturity within DTs' context. The insights obtained from this review indicate that there are multiple potential future research directions for a successful road DT development. For example, real-time data analytics capability, optimum choice of ML algorithms, integration of various ML algorithms for different DT capacities and the inclusion of probability ML considerations are some of the research areas. Further exploration and development in DTs for road would allow transportation agencies to more intelligently and efficiently, design, construct, operate and maintain their road infrastructure systems.

Acknowledgements

The financial support from the UK Collaboratorium for Research in Infrastructure and Cities (UKCRIC), through the UKCRIC ECR project funding is acknowledged with gratitude.

References

Abbott, D. (n.d.). *Applied Predictive Analytics : Principles and Techniques for the Professional Data Analyst*. 453.

Abdelaziz, N., Abd El-Hakim, R. T., El-Badawy, S. M., & Afify, H. A. (2018). *International Roughness Index prediction model for flexible pavements*. *21*(1), 88–99.

Abramovici, M., Göbel, J. C., & Savarino, P. (2017). Reconfiguration of smart products during their use phase based on virtual product twins. *CIRP Annals*, *66*(1), 165–168. https://doi.org/10.1016/J.CIRP.2017.04.042

Ahmadi, A., Khalesi, S., & Golroo, A. (2021). *An integrated machine learning model for automatic road crack detection and classification in urban areas*. https://doi.org/10.1080/10298436.2021.1905808

Alharbi, F. (2018). Predicting pavement performance utilizing artificial neural network (ANN) models. *Graduate Theses and Dissertations*. https://lib.dr.iastate.edu/etd/16703

Alzraiee, H., Ruiz, A. L., & Sprotte, R. (2021). Detecting of Pavement Marking Defects Using Faster R-CNN. *Journal of Performance of Constructed Facilities*, *35*(4), 04021035. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001606

Amirhossein Hosseini, S. (2020). *Data-driven framework for modeling deterioration of pavements in the state of Iowa the state of Iowa*. https://lib.dr.iastate.edu/etd

Arias Chao, M., Kulkarni, C., Goebel, K., & Fink, O. (2020). Fusing Physics-based and Deep Learning Models for Prognostics. *Reliability Engineering and System Safety*, *217*. https://doi.org/10.1016/j.ress.2021.107961

Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., Omata, H., Kashiyama, T., Seto, T., Mraz, A., & Sekimoto, Y. (2021). *RDD2020: An Image Dataset for Smartphone-based Road Damage Detection and Classification*. *1*. https://doi.org/10.17632/5TY2WB6GVG.1

Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., & Sekimoto, Y. (2021). RDD2020: An annotated image dataset for automatic road damage detection using deep learning. *Data in Brief*, *36*, 107133. https://doi.org/10.1016/J.DIB.2021.107133

Atikur, C. (2020). *Pothole Dataset*. https://public.roboflow.com/object-detection/pothole

Azimi, M., & Pekcan, G. (2020). Structural health monitoring using extremely compressed data through deep learning. *Computer-Aided Civil and Infrastructure Engineering*, *35*(6), 597–614. https://doi.org/10.1111/MICE.12517

Barricelli, B. R., Casiraghi, E., & Fogli, D. (2019). A survey on digital twin: Definitions, characteristics, applications, and design implications. In *IEEE Access* (Vol. 7). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2019.2953499

Bibi, R., Saeed, Y., Zeb, A., Ghazal, T. M., Rahman, T., Said, R. A., Abbas, S., Ahmad, M., & Khan, M. A. (2021). Edge AI-Based Automated Detection and Classification of Road Anomalies in VANET Using Deep Learning. *Computational Intelligence and Neuroscience*, *2021*. https://doi.org/10.1155/2021/6262194

Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic Construction Digital Twin: Directions for future research. In *Automation in Construction* (Vol. 114, p. 103179). Elsevier B.V. https://doi.org/10.1016/j.autcon.2020.103179

Bolton. (2018). *The Gemini Principles*. https://doi.org/10.17863/CAM.32260

Bowden, S., Dorr, A., Thorpe, T., & Anumba, C. (2006). Mobile ICT support for construction process improvement. *Automation in Construction*, *15*(5), 664–676. https://doi.org/10.1016/j.autcon.2005.08.004

Brnabic, A., & Hess, L. M. (2021). Systematic literature review of machine learning methods used in the analysis of real-world data for patient-provider decision making. *BMC Medical Informatics and Decision Making*, *21*(1), 1–19. https://doi.org/10.1186/s12911-021-01403-2

Bukharin, A. W., Yang, Z., & Tsai, Y. (James). (2021). Five-Year Project-Level Statewide Pavement Performance Forecasting Using a Two-Stage Machine Learning Approach Based on Long Short-Term Memory. *Transportation Research Record: Journal of the Transportation Research Board*, 036119812110171. https://doi.org/10.1177/03611981211017132

Burningham, S., & Stankevich, N. (2005). *Why road maintenance is important and how to get it done*.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(2), 171–209. https://doi.org/10.1007/S11036-013-0489-0/FIGURES/6

Choi, S., & Do, M. (2020). Development of the road pavement deterioration model based on the deep learning method. *Electronics (Switzerland)*, *9*(1), 3. https://doi.org/10.3390/electronics9010003

Citir, N., Ceylan, H., Kim, S., & Kaya, O. (2021). Incorporating the Effectiveness of Preservation and Rehabilitation Techniques on Flexible Pavement Service Life Predictions Using Machine Learning Approach. *Airfield and Highway Pavements 2021: Pavement Materials and Sustainability - Selected Papers from the International Airfield and Highway Pavements Conference 2021*, 365–377. https://doi.org/10.1061/9780784483510.033

D Trousdale. (2019). The Sensorisation of Infrastructure Management and Maintenance: Enabling Better Service Delivery. *26th World Road Congress World Road*. https://trid.trb.org/view/1738705

D'amico, F., Ciampoli, L. B., di Benedetto, A., Bertolini, L., & Napolitano, A. (2022). Integrating Non-Destructive Surveys into a Preliminary BIM-Oriented Digital Model for Possible Future Application in Road Pavements Management. *Infrastructures 2022, Vol. 7, Page 10*, *7*(1), 10. https://doi.org/10.3390/INFRASTRUCTURES7010010

Demkovich, N., Yablochnikov, E., & Abaev, G. (2018). Multiscale modeling and simulation for industrial cyber-physical systems. *Proceedings - 2018 IEEE Industrial Cyber-Physical Systems, ICPS 2018*, 291–296. https://doi.org/10.1109/ICPHYS.2018.8387674

di Graziano, A., Marchetta, V., & Cafiso, S. (2020). Structural health monitoring of asphalt pavements using smart sensor networks: A comprehensive review. *Journal of Traffic and Transportation Engineering (English Edition)*, *7*(5), 639–651. https://doi.org/10.1016/J.JTTE.2020.08.001

Du, R., Qiu, G., Gao, K., Hu, L., & Liu, L. (2020). Abnormal Road Surface Recognition Based on Smartphone Acceleration Sensor. *Sensors 2020, Vol. 20, Page 451*, *20*(2), 451. https://doi.org/10.3390/S20020451

Errandonea, I., Beltrán, S., & Arrizabalaga, S. (2020). Digital Twin for maintenance: A literature review. In *Computers in Industry* (Vol. 123, p. 103316). Elsevier B.V. https://doi.org/10.1016/j.compind.2020.103316

Fathi, A., Mazari, M., Saghafi, M., Hosseini, A., & Kumar, S. (2019). Parametric study of pavement deterioration using machine learning algorithms. *Airfield and Highway Pavements 2019: Innovation and Sustainability in Highway and Airfield Pavement Technology - Selected Papers from the International Airfield and Highway Pavements Conference 2019*, 31–41. https://doi.org/10.1061/9780784482476.004

FHWA. (1995). *LTPP InfoPave - Home*. https://infopave.fhwa.dot.gov/

Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital Twin: Enabling Technologies, Challenges and Open Research. *IEEE Access*, *8*, 108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/10.1016/J.IJINFOMGT.2014.10.007

Gao, L., Yu, Y., Ren, Y. H., & Lu, P. (2021a). Detection of Pavement Maintenance Treatments using Deep-Learning Network: *Https://Doi.Org/10.1177/03611981211007846*, *2675*(9), 1434–1443. https://doi.org/10.1177/03611981211007846

Gong, H., Sun, Y., Mei, Z., & Huang, B. (2018). Improving accuracy of rutting prediction for mechanistic-empirical pavement design guide with deep neural networks. *Construction and Building Materials*, *190*, 710–718. https://doi.org/10.1016/J.CONBUILDMAT.2018.09.087

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews*. Sage.

Han, C., Ma, T., Xu, G., Chen, S., & Huang, R. (2020). *Intelligent decision model of road maintenance based on improved weight random forest algorithm*. https://doi.org/10.1080/10298436.2020.1784418

Haylen, A. (2019). *Local road maintenance in England*. https://commonslibrary.parliament.uk/research-briefings/cbp-8383/

He, B., Liu, L., & Zhang, D. (2021). Digital Twin-Driven Remaining Useful Life Prediction for Gear Performance Degradation: A Review. *Journal of Computing and Information Science in Engineering*, *21*(3). https://doi.org/10.1115/1.4049537

Highway Administration, F., & Pavement Performance program, L.-T. (2017). *The Long-Term Pavement Performance Program*.

Hofmann, W., & Branding, F. (2019). Implementation of an IoT- And cloud-based digital twin for real-time decision support in port operations. *IFAC-PapersOnLine*, *52*(13), 2104–2109. https://doi.org/10.1016/j.ifacol.2019.11.516

ISO/ISO-AWI 23247. (2019). *ISO - ISO 23241:2021 - Automation systems and integration — Digital twin framework for manufacturing — Part 1: Overview and general principles*. https://www.iso.org/standard/75066.html

Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the Digital Twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, *29*, 36–52. https://doi.org/10.1016/j.cirpj.2020.02.002

Karballaeezadeh, N., Dineva, A., Mosavi, A., Nabipour, N., Shamshirband, S., & Mohammadzadeh, D. (2019). *Hybrid Machine Learning Model of Support Vector Machine and Fruit Fly Optimization Algorithm for Prediction of Remaining Service Life of Flexible Pavement*. https://doi.org/10.20944/PREPRINTS201910.0238.V1

Kargah-Ostadi, N., Zhou, Y. (Mina), & Rahman, T. (2019). *Developing Performance Prediction Models for Pavement Management Systems in Local Governments in Absence of Age Data: 2673*(3), 334–341. https://doi.org/10.1177/0361198119833680

Karimzadeh, A. (2020). *Predictive Analytics for Roadway Maintenance: A Review of Current Models, Challenges, and Opportunities*. Civil Engineering Journal. https://civilejournal.org/index.php/cej/article/view/2005/pdf

Kaur, M. J., Mishra, V. P., & Maheshwari, P. (2020). The Convergence of Digital Twin, IoT, and Machine Learning: Transforming Data into Action. In *Internet of Things* (pp. 3–17). Springer International Publishing. https://doi.org/10.1007/978-3-030-18732-3_1

Kulkarni, R. B., Miller, R. W., & Kulkarni, R. B. (2003). *Pavement Management Systems Past, Present, and Future*.

Lee, D., Kim, J. C., Kim, M., & Lee, H. (2021). Intelligent Tire Sensor-Based Real-Time Road Surface Classification Using an Artificial Neural Network. *Sensors 2021, Vol. 21, Page 3233*, *21*(9), 3233. https://doi.org/10.3390/S21093233

Liehr, S., Münzenberger, S., & Krebber, K. (2019). Wavelength-Scanning Distributed Acoustic Sensing for Structural Monitoring and Seismic Applications. *Proceedings 2019, Vol. 15, Page 30*, *15*(1), 30. https://doi.org/10.3390/PROCEEDINGS2019015030

Lu, Q., Parlikad, A. K., Woodall, P., Don Ranasinghe, G., Xie, X., Liang, Z., Konstantinou, E., Heaton, J., & Schooling, J. (2020). Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus. *Journal of Management in Engineering*, *36*(3), 05020004. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000763

Macchi, M., Roda, I., Negri, E., & Fumagalli, L. (2018). Exploring the role of Digital Twin for Asset Lifecycle Management. *IFAC-PapersOnLine*, *51*(11), 790–795. https://doi.org/10.1016/J.IFACOL.2018.08.415

Maciej Serda. (2013). Synteza i aktywność biologiczna nowych analogów tiosemikarbazonowych chelatorów żelaza. *Uniwersytet Śląski*, 343–354. https://doi.org/10.2/JQUERY.MIN.JS

Majidifard, H., Adu-Gyamfi, Y., & Buttlar, W. G. (2020). *Deep Machine Learning Approach to Develop a New Asphalt Pavement Condition Index*.

Marai, O. el, Taleb, T., & Song, J. S. (2020). Roads Infrastructure Digital Twin: A Step Toward Smarter Cities Realization. *IEEE Network*. https://doi.org/10.1109/MNET.011.2000398

Marcelino, P., de Lurdes Antunes, M., Fortunato, E., & Gomes, M. C. (2019). Transfer learning for pavement performance prediction. *International Journal of Pavement Research and Technology 2019 13:2*, *13*(2), 154–167. https://doi.org/10.1007/S42947-019-0096-Z

Marcelino, P., de Lurdes Antunes, M., Fortunato, E., & Gomes, M. C. (2021). Machine learning approach for pavement performance prediction. *International Journal of Pavement Engineering*, *22*(3), 341–354. https://doi.org/10.1080/10298436.2019.1609673

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). *Implementation of machine-learning classification in remote sensing: an applied review Implementation of machine-learning classification in remote sensing: an applied review*. https://doi.org/10.1080/01431161.2018.1433343

Menegazzo, J., & von Wangenheim, A. (2021a). Road surface type classification based on inertial sensors and machine learning: A comparison between classical and deep machine learning approaches for multi-contextual real-world scenarios. *Computing*, *103*(10), 2143–2170. https://doi.org/10.1007/S00607-021-00914-0/TABLES/17

Milad, A., Adwan, I., Majeed, S. A., Yusoff, N. I. M., Al-Ansari, N., & Yaseen, Z. M. (2021). Emerging Technologies of Deep Learning Models Development for Pavement Temperature Prediction. *IEEE Access*, *9*, 23840–23849. https://doi.org/10.1109/ACCESS.2021.3056568

MJ DeJong, JM Schooling, & GMB Viggiani. (2019). *International Conference on Smart Infrastructure and Construction 2019 (ICSIC) | Cambridge Centre for Smart Infrastructure & Construction*. https://www.icevirtuallibrary.com/doi/book/10.1680/icsic.64669

Morris, C., & Yang, J. J. (2021). A machine learning model pipeline for detecting wet pavement condition from live scenes of traffic cameras. *Machine Learning with Applications*, *5*, 100070. https://doi.org/10.1016/J.MLWA.2021.100070

Nabipour, N., Karballaeezadeh, N., Dineva, A., Mosavi, A., Mohammadzadeh S., D., & Shamshirband, S. (2019). Comparative Analysis of Machine Learning Models for Prediction of Remaining Service Life of Flexible Pavement. *Mathematics 2019, Vol. 7, Page 1198*, *7*(12), 1198. https://doi.org/10.3390/MATH7121198

Negri, E., Fumagalli, L., & Macchi, M. (2017). A Review of the Roles of Digital Twin in CPS-based Production Systems. *Procedia Manufacturing*, *11*, 939–948. https://doi.org/10.1016/J.PROMFG.2017.07.198

NIC. (2017). *Data for the public good*.

Oliveira, J. R. de, Lima, E. R. D., Almeida, L. M. D., & Wanner, L. (2021). Improving sensor data quality with predictive models. *7th IEEE World Forum on Internet of Things, WF-IoT 2021*, 735–740. https://doi.org/10.1109/WF-IOT51360.2021.9595020

Piryonesi, S. M. (2019). *The Application of Data Analytics to Asset Management: Deterioration and Climate Change Adaptation in Ontario Roads*.

Qiuchen Lu, V., Parlikad, A. K., Woodall, P., Ranasinghe, G. D., & Heaton, J. (2019). Developing a Dynamic Digital Twin at a Building Level: using Cambridge Campus as Case Study. *International Conference on Smart Infrastructure and Construction 2019 (ICSIC)*, 67–75. https://doi.org/10.1680/icsic.64669.067

Redeker, M., Weskamp, J. N., Rössl, B., & Pethig, F. (2021). *Towards a Digital Twin Platform for Industrie 4.0 Industrie 4.0 Communication View project Intelligent Automation Technologies View project Towards a Digital Twin Platform for Industrie 4.0*. https://doi.org/10.1109/ICPS49255.2021.9468204

Saleh, Y., & Otoum, N. (2020). *Road-Type Classification through Deep Learning Networks Fine-Tuning*. *19*(1). https://doi.org/10.1142/S0219649220400201

Steyn, W. J. V. (2020). Selected implications of a hyper-connected world on pavement engineering. *Springer*, *13*(6), 673–678. https://doi.org/10.1007/s42947-020-6012-7

Tabatabaee, N., Ziyadi, M., & Shafahi, Y. (2013). Two-Stage Support Vector Classifier and Recurrent Neural Network Predictor for Pavement Performance Modeling. *Journal of Infrastructure Systems*, *19*(3), 266–274. https://doi.org/10.1061/(ASCE)IS.1943-555X.0000132

Tao, F., Zhang, H., Liu, A., & Nee, A. Y. C. (2019). Digital Twin in Industry: State-of-the-Art. *IEEE Transactions on Industrial Informatics*, *15*(4), 2405–2415. https://doi.org/10.1109/TII.2018.2873186

Thomas, J., & Brunton, J. (2010). *EPPI-Reviewer 4: Software for Research Synthesis*.

Tsou, A. Y., Treadwell, J. R., Erinoff, E., & Schoelles, K. (2020). Machine learning for screening prioritization in systematic reviews: comparative performance of Abstrackr and EPPI-Reviewer. *Systematic Reviews*, *9*(1), 1–14.

UCL. (2019). *EPPI-Reviewer Web (Beta)*. Social Science Research Unit, UCL Institute of Education: London. https://eppi.ioe.ac.uk/eppireviewer-web/home

UK House of Commons. (2019). *Local roads funding and maintenance: filling the gap Tenth Report of Session 2017-19 Report, together with formal minutes relating to the report*. www.parliament.uk/transcom

VanDerHorn, E., & Mahadevan, S. (2021). Digital Twin: Generalization, characterization and implementation. *Decision Support Systems*, *145*. https://doi.org/10.1016/J.DSS.2021.113524

Varma, G., Subramanian, A., Namboodiri, A., Chandraker, M., & Jawahar, C. v. (2018). IDD: A Dataset for Exploring Problems of Autonomous Navigation in Unconstrained Environments. *Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019*, 1743–1751. https://doi.org/10.48550/arxiv.1811.10200

Willard, J., Jia, X., Steinbach, M., Kumar, V., & Xu, S. (2020). *Integrating Physics-Based Modeling With Machine Learning: A Survey*. *1*, 34. https://doi.org/10.1145/1122445.1122456

Yao, L., Dong, Q., Jiang, J., & Ni, F. (2020). Deep reinforcement learning for long-term pavement maintenance planning. *Computer-Aided Civil and Infrastructure Engineering*, *35*(11), 1230–1245. https://doi.org/10.1111/MICE.12558

Yao, L., Leng, Z., Jiang, J., & Ni, F. (2021). *Modelling of pavement performance evolution considering uncertainty and interpretability: a machine learning based framework*. https://doi.org/10.1080/10298436.2021.2001814

Ye, C., Butler, L., Bartek, C., Iangurazov, M., Lu, Q., Gregory, A., Girolami, M., & Middleton, C. (2019). A Digital Twin of Bridges for Structural Health Monitoring. *In: Proceedings of the 12th International Workshop on Structural Health Monitoring. Stanford University (2019) (In Press).* . https://web.stanford.edu/group/sacl/workshop/IWSHM2019/specialsession.html

Yu, G., Zhang, S., Hu, M., & Ken Wang, Y. (2020). Prediction of highway tunnel pavement performance based on digital twin and multiple time series stacking. *Advances in Civil Engineering*, *2020*. https://doi.org/10.1155/2020/8824135

Zhang, J. (2019). Bridging the Physical, the Digital, and the Social. *IEEE Internet of Things Magazine*, *2*(1), 8–9. https://doi.org/10.1109/miot.2019.8835418

Ziari, H., Maghrebi, M., Ayoubinejad, J., & Waller, S. T. (2016). Prediction of pavement performance: Application of support vector regression with different kernels. *Transportation Research Record*, *2589*, 135–145. https://doi.org/10.3141/2589-15

Ziari, H., Sobhani, J., Ayoubinejad, J., & Hartmann, T. (2015). Prediction of IRI in short and long terms for flexible pavements: ANN and GMDH methods. *International Journal of Pavement Engineering*, *17*(9), 776–788. https://doi.org/10.1080/10298436.2015.1019498