An Integrated Deep Learning-Based Approach for Automobile Maintenance Prediction with GIS Data

Chong Chen[a, b], Ying Liu[a*], Xianfang Sun[c], Carla Di Cairano-Gilfedder[d] and Scott Titmus[e]

[a] *Institute of Mechanical and Manufacturing Engineering, School of Engineering, Cardiff University, Cardiff, CF24 3AA, UK*

[b] *Guangdong Provincial Key Laboratory of Cyber-Physical System, Guangdong University of Technology, Guangzhou 510006, China*

[c] *School of Computer Science and Informatics, Cardiff University, Queen's Buildings, Cardiff CF24 3AA, UK*

[d] *Applied Research, BT Technology, Ipswich, IP5 3RE, UK*

[e] *Rivus Solutions (formerly BT Fleet), Solihull, B37 7YN, UK*

[*] *Corresponding author. E-mail: LiuY81@Cardiff.ac.uk*

Abstract

Predictive maintenance (PdM) can be beneficial to the industry in terms of lowering maintenance cost and improve productivity. Remaining useful life (RUL) prediction is an important task in PdM. The RUL of an automobile can be impacted by various surrounding factors such as weather, traffic and terrain, which can be captured by the geographical information system (GIS). Recently, most researchers have conducted studies of RUL modelling based on sensor data. Owing to the fact that the collection of sensor data is expensive, while maintenance data is relatively easy to obtain. This study aims to establish an automobile RUL prediction model with GIS data through a data-driven approach. In this approach, firstly, due to the data type and sampling rate of the maintenance data and GIS data are different, a data integration scheme was researched. Secondly, the Cox proportional hazard model (Cox PHM) was introduced to construct the health index (HI) for the integrated data. Then, a deep learning structure called M-LSTM (Merged-long-short term memory) network was designed for HI modelling based on the integrated data which contains both sequential data and ordinary numeric data. Finally, the RUL was mapped by predicted HI and the Cox PHM. An experimental study using a sizable real-world fleet maintenance dataset provided by a UK fleet company revealed the effectiveness of the proposed approach and the impact of the GIS factors on the automobiles under investigation.

Keywords: Predictive Maintenance; RUL prediction; Deep Learning; GIS; Machine Learning

1. Introduction

In the era of smart manufacturing, there is a high demand in the development of advanced PdM techniques. It is of great importance to the industry as it can improve the decision support for maintenance management [1]. With the advanced PdM techniques, the maintenance cost can be minimised and the productivity and profitability of a company can be leveraged [2]. The prediction of RUL is a challenging task in industry, such as automobile fleet management. If the RUL of an automobile can be accurately predicted, appropriate maintenance can be scheduled in advance to avoid serious failure, lower the maintenance cost and machine downtime.

Recently, various emerging PdM techniques have been proposed. There are two main research areas in PdM, which are condition-based PdM and statistical-based PdM [2]. Condition-based PdM is the prevailing type in PdM. It aims to model the equipment or system degradation based on sensor data [3]. However, the deployment of sensors and data collection needs extra investment, which may not be affordable for some companies. Without available sensor data, the implementation of conditioned based PdM is still challenging. Statistical PdM, as another type of PdM, also can be a useful tool in the industry. In comparison with the condition based PdM base on the real-time data, statistical-based PdM aims at failure prediction modelling base on the event data such as maintenance data [4], which is more available in the industry in comparison with the sensor data.

The automobile lifecycle is impacted by various factors such as automobile design, driver behaviour and working environment. For a fleet management company that possesses a large number of automobiles, the automobiles can be allocated in different areas. The geographical factors which may affect automobile lifecycle, such as weather, terrain, and traffic, are different from area to area. Hence, introducing these factors into the study of PdM can bring tangible

benefits to the fleet management company. GIS is a system designed to capture, store, manipulate, analyse, manage, and present various types of geographical data [5]. With the help of GIS, the GIS data which may relevant to the automobile lifecycle can be collected. After the GIS data is collected, an issue that needs to be addressed is integrating these data with maintenance data.

Data integration is an essential part of machine learning. The data integration approaches can be classified as three types: (1) Early integration: projecting or concatenating the different datasets to a larger dataset and feed it to a machine learning algorithm; (2) Intermediate integration: fusing different datasets via a joint machine learning algorithm such as deep neural networks; (3) Late integration: Training different machine learning models using different datasets, and taking a majority vote to determine the final output [6]. However, it is hard to directly implement the above data integration approaches for maintenance data and GIS data. The reasons are: Firstly, both maintenance dataset and GIS dataset may contain sequential data, ordinary numeric data and nominal data. The data types in both datasets are different; Furthermore, the sampling rate of the sequential data in both datasets is different, which requires data transformation before data integration. Hence, an appropriate data integration scheme needs to be explored to integrate maintenance data and GIS data. Deep learning, as a subset of machine learning, has gained increasing attention [7]. Deep learning components such as a fully-connected layer, convolutional layer, and long-short term memory layer are good at processing ordinary numeric data, image data, and sequential data, respectively. In the research of deep learning, merged neural network as an emerging neural network structure, can combine different sub neural networks and therefore enable the data integration of multi-source data [8, 9]. Hence, it can be a useful tool in modelling based on the maintenance dataset and GIS dataset.

Recently, deep learning algorithms have been widely studied in the studies of RUL modelling [10-17]. However, most of the existing studies are only based on sensor data. The GIS factors

that may impact the equipment or system lifecycle has not been considered. In this paper, an automobile RUL modelling approach with the consideration of GIS data is proposed. In this approach, maintenance data is first transformed and integrate with GIS data. Secondly, a Cox PHM is deployed to construct the health index (HI) of the integrated data. Thirdly, a merge neural network called M-LSTM network is designed for HI modelling. Finally, the predicted HI is used to map the RUL based on the Cox PHM obtained from step 2. The main contributions of this work are listed as follow: (1) Different from the existing RUL prediction approach based on sensor data which is expensive to obtain, an RUL prediction approach based on maintenance data is proposed; (2) This is the first time that GIS data was introduced into the study of PdM to improve the RUL prediction accuracy; (3) A data integration scheme was proposed to integrate maintenance data and GIS data from multi-source; (4) M-LSTM network, a multi-model neural network, was designed to learn and fuse the high-level representation of the sequential data and ordinary numeric data. The rest of this paper is organised as follows: The relevant studies in PdM using machine learning and GIS research using machine learning are reviewed in Section 2. Section 3 details the problem statement. Section 4 introduces the methodology of this study. Section 5 reports an experimental study. The experimental results are demonstrated in Section 6 and discussed in Section 7. Finally, the conclusions were drawn in Section 8.

2. Literature Review

2.1. The Studies of Predictive Maintenance Using Machine Learning

Deep learning has been prevailing in the research of PdM, especially in the condition based PdM. Various deep learning structures, such as transformers neural network [16], deep convolutional neural network [10], dual-LSTM network [17], dilated convolution neural network [11], deep adversarial neural networks [12], and multi-scale deep convolutional neural

network [13], were proposed and verified using NASA's Commercial Modular Aero-Propulsion System Simulation (C-MAPSS) dataset. Most of the deep learning structures developed for RUL modelling have only one training path. Al-Dulaimi et al. (2019) [14] proposed a noisy and hybrid CNN and BLSTM-based deep (NBLSTM) network that contains a CNN path and bi-directional LSTM path. Zhang et al. (2020) [15] proposed a merged LSTM network that contains two LSTM sub-networks. The inputs of both merged neural network mentioned above are the same.

In order to model the machinery RUL of dependent competing failure processes (DCFPs), Yan et al. (2021) [18] researched an RUL modelling approach with the consideration of both soft and hard failure. In this approach, the first passage time-based analytical expression of degradation is derived. Then, the offline RUL estimation and online parameters update are jointly conducted. In order to achieve long term accurate prediction of RUL. To reduce the effect of time-lagged correlations on the feature extraction, Zhang et al. (2018) [19] deployed a curve-registration method to evaluate the time lags among sensors. After the time lags of sensors were adjusted, the data was then used to establish an LSTM network model to predict upcoming failure. In order to model the degradation of manufacturing systems, a hidden Markov model with auto-correlated observation was proposed. The current state of this model depends on both the corresponding hidden system state and the previous observations. Besides, EM (expectation maximum) was adopted to estimate the unknown parameters. Based on the prediction of RUL, an optimised maintenance policy was developed [20]. Right-censored data represents the maintenance was scheduled before an asset completely failed. In the right-censored scenario, the maintenance time located arbitrarily before the failure. In order to model the asset degradation, a method called Relative Entropy Weibull-SAX was proposed using HI and HS degradation modelling method for multivariate asset data. In this approach, the HI of an asset can be constructed using relative entropy. The experimental results based on C-MAPSS

show that this method was able to represent the health stage of observed engine [21]. Li et al. (2021) [22] researched an RUL prediction method based on a multi-sensor data fusion model, which a state transition function and a Wiener process are deployed to express the degradation process of the system state. In order to deploy the training model in different operational scenarios, Zhang et al. (2021) [3] proposed a transfer learning approach using deep representation regularisation for RUL prediction. In this approach, the healthy state and degradation direction of the data from the target domain and source domain are aligned before the RUL is predicted.

Statistical based PdM is based on the event data such as maintenance data or log data. In order to reduce the machine downtime, a data-driven approach was proposed to predict failures of healthcare machine. In this study, raw system log information including time, id and description were extracted and then used to establish an SVM model for failure classification [23]. In order to address the data imbalance issue in log data, Dangut et al. (2021) [24] proposed a hybrid approach based on natural language processing techniques and ensemble learning to predict the aircraft component failure. Calabrese et al. (2020) [25] researched the Gradient Boosting algorithm for the fault classification of woodworking industrial machines based on log data. In this study, temporal feature engineering techniques were deployed to enhance the performance of the Gradient Boosting machine. Kobayashi et al. (2017) [26] proposed a method based on a graph-based algorithm to extract failures and their causes from network system log data. In this method, a graph-based algorithm called PC algorithm is firstly introduced to infer causal structure from event time series efficiently. Then, data pre-processing and post-processing methods from a set of log messages were proposed to improve the performance of the PC algorithm. It can be seen that the RUL prediction approaches in the existing literature are all based on sensor data, while the mainstream of statistical PdM focuses on failure classification or cause mining. Furthermore, the existing studies have not considered the impact of GIS factors on RUL modelling.

2.2. The studies of GIS Using Machine Learning

Pham et al. (2017) [5] combined several ensemble methods with multiple perceptron neural Networks to establish a landslide classification model. GIS features such as slope, slope aspect, elevation, curvature and plan curvature were adopted for modelling. In order to identify the contribution of the features to the landslide, a feature selection method called Relief-F method was used. Tehrany et al. (2014) [27] proposed a flood susceptibility mapping approach based on the data collected from the records of flood occurrence. The terrain features used in this study included flood inventory, slope, stream power index, topographic wetness index, and altitude, etc. The weight-of-evidence method was applied to measure the relevant weight of each factor. Then, the data contains these features were reclassified using the acquired weights, before the data is sent to the support vector machine model to evaluate the correlation between flood occurrence and each conditioning factor. Four types of kernel-based SVMs (linear, polynomial, radial basis function and sigmoid) were used for modelling. The results indicate that the RBF kernel-based SVM has achieved the best performance.

Naghibi et al. (2016) [28] investigated groundwater potential mapping using tree-based algorithms. The purpose of this study was to produce groundwater spring potential maps in Koohrang, Iran. These GIS factors include slope degree, slope aspect, altitude, topographic wetness index, lithology, and land use, etc. The groundwater spring potential was modelled and mapped using classification and regression tree, random forest, and boosted regression tree algorithms. Another study of potential groundwater mapping is that Rahmati et al. (2016) [29] deployed random forest and maximum entropy models for groundwater potential mapping is investigated at Mehran Region, Iran. Massawe et al. (2018) [30] proposed a mapping approach for soil taxa mapping based on heterogeneous data, which was collected from different sources including satellite image, digital elevation map and digital soil map. The collected features

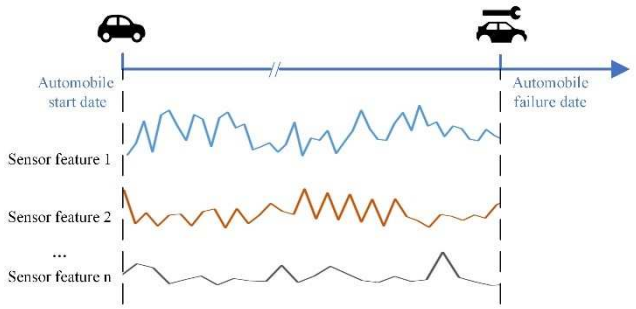
include soil classes, effects of living organisms (vegetation), terrain parameters and spatial location. Random forest and J48 algorithms were used to train the soil profile classification model separately.

In order to build a classifier for different elements using spectral and spatial data, a spectral-spatial feature-based classification (SSFC) framework was proposed to lower the dimension of spectral data and extract the features from spatial data. In this framework, a dimension reduction method called balanced local discriminant embedding (BLDE) was proposed to lower the dimension of hyperspectral images (spectral data with high resolution). CNN was used to extract the abstract features from spatial images. The features obtained from BLDE and CNN were then combined and used to train a classifier [31]. The rapid development of electric vehicles can significantly alleviate environmental problems and energy tension. Zhang et al. (2019) [32] proposed a multi-objective optimisation model based on particle swarm optimisation to plan the placement of the charging station of the electric vehicle. GIS was used in this study to identify the intersection of power system and traffic system maps. The intersections of the maps are the candidates of the charging stations. In the research based on GIS data, machine learning has become a useful tool. However, there is no existing study that concerns using GIS data in the study of PdM.

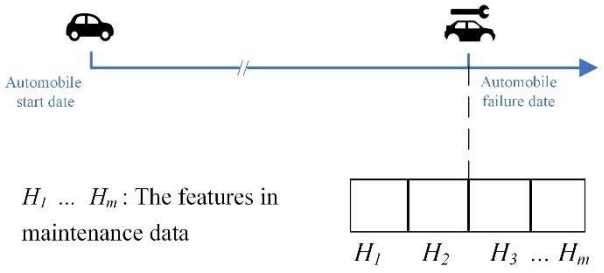
3. Problem Statement

PdM has become a big concern for various industry sectors. In an automobile fleet maintenance company, the prevailing maintenance policies are corrective maintenance and preventive maintenance. Corrective maintenance, also known as run to failure, will lead to high maintenance cost because it leads to considerable uncertainty in fleet management. Preventive maintenance can lower the maintenance cost to some extent, while the schedule check period needs to be well considered [33]. A too long and too short schedule check period leads to a high

accident rate and high maintenance cost, respectively. From the state-of-the-art, most of the research in PdM has been carried out based on the sensor data, which can reveal the health status of equipment at a certain time. In contrast, the maintenance data can only provide information when a failure occurs. The demonstration of sensor data and maintenance data in the automobile lifecycle is shown in Figure 1. The RUL prediction is the main focus of the study of PdM. When the RUL can be accurately predicted, the maintenance can be scheduled in an appropriate time. Since the maintenance data does not cover the information that can reveal the health status of equipment, it is challenging that using maintenance data to establish an equipment RUL prediction.



(a) Sensor data



(b) Maintenance data

Figure 1. The demonstration of sensor data and maintenance data

In PdM, most research has studied the equipment lifecycle either based on the sensor data or maintenance data. However, to the best of our knowledge, there is no research to consider the impact of GIS factor on the product lifecycle. For an automobile, its lifecycle can be affected by various surrounding factors such as weather, traffic, and terrain. With the enrichment of GIS

data, the automobile RUL prediction accuracy can be improved. Moreover, it can offer insights into how the GIS factors affect automobile lifecycle, which can help a fleet management company to optimise its maintenance policy. Our previous study introduced GIS data into automobile time-between-failure (TBF) modelling through a data-driven approach. The maintenance data and GIS data were directly concatenated and fed into a deep forest algorithm for training. The experimental results showed that prediction accuracy was promoted with the enrichment of GIS data [34]. However, the data integration of maintenance data and GIS data can be further investigated.

The identification of the data characteristics is important to the integration of maintenance data and GIS data. Automobile maintenance data is originated from the automobile maintenance record. In comparison with the sensor data, which can reveal the automobile health status, maintenance data records the automobile information in a maintenance event. Each data entry in maintenance data contains the basic automobile information such as mileage, age, last time to repair and model in the automobile start date (i.e. the date automobile after maintenance or first used). The output of each data entry is the next TBF. Hence, a maintenance data entry is collected in a certain timestamp. Besides, maintenance data can be further classified as two specific types which are sequential data and ordinary numeric data. The data which has sequential property is relevant to the maintenance history, such as the repaired time, automobile age, and mileage. Because the next TBF of an automobile tends to be shorter than the last TBF, the sequential features in the last TBF are relevant to the next TBF [35]. The ordinary numeric features in maintenance data are constants such as model. In our previous study, all the ordinary numeric features are considered as sequential features. The data which contained these features were fed into an LSTM network for training. The algorithm performance in terms of model correlation coefficient and root mean square error is better than that of a fully-connected neural network [36]. However, when the number of ordinary numeric features increases, the algorithm performance of LSTM network could be compromised. Hence, a new data integration technique

needs to be explored for automobile RUL modelling based on these two types of data.

A maintenance data entry can only be collected when maintenance is implemented. Different from maintenance data, GIS data can be collected continuously and chronologically. For example, the weather data can be collected daily, weekly, and monthly, etc. If a GIS feature is collected monthly in a year, it is a one-dimension array. If multiple GIS features are collected in a long period according to a certain sampling frequency, a two-dimensions array can be obtained. Figure 2 shows the automobile maintenance data and GIS data. The features in maintenance dataset, which contains both sequential and ordinary numeric features, is a one-dimension array.

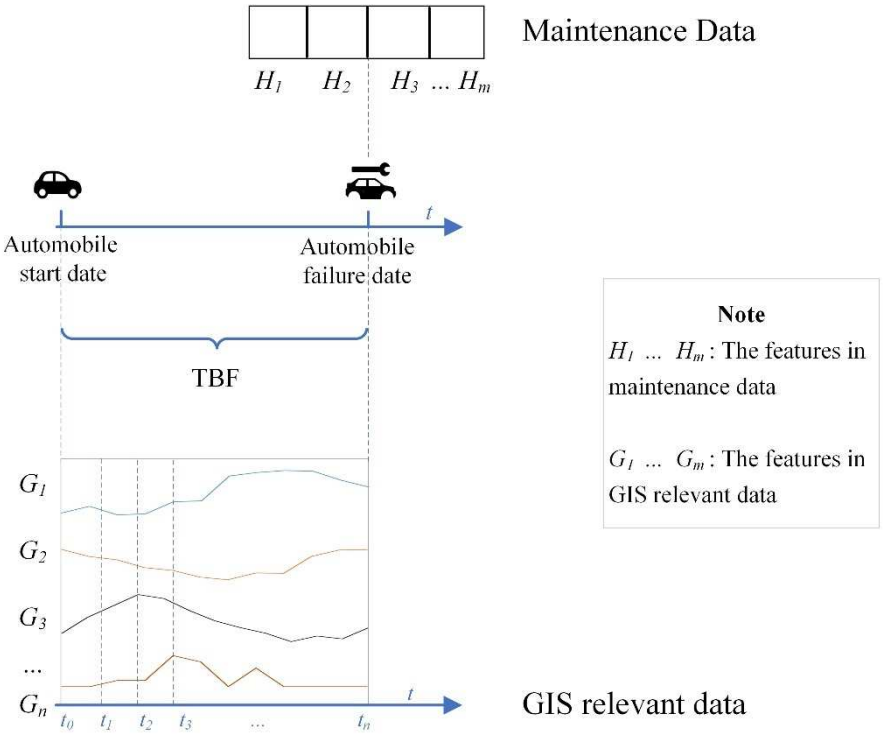


Figure 2. The automobile maintenance data and GIS data

Hence, the research question in this study is: how to integrate the maintenance data and GIS data for automobile RUL prediction? In this study, we propose a new approach that contains a novel deep learning structure to integrate data from different sources and establish an automobile RUL prediction model.

4. Methodology

In our proposed approach, firstly, the maintenance data needs to be collected from the garage of a fleet management company, while the GIS data needs to be collected according to the automobile working area. The dataset is split into two parts, where the first part is used for data integration, and the second part is used for HI construction for the integrated data. Secondly, the maintenance data and GIS data are integrated. In this stage, both datasets are classified into two parts which are sequential data and ordinary numeric data. Both types of data from maintenance data and GIS data are concatenated. The details of data integration are introduced in Section 4.2. Thirdly, Cox PHM is deployed to estimate the HI using the maintenance data in conjunction with the statistical GIS data (i.e. the features of mean and standard deviation). Fourthly, a merged neural network called M-LSTM network is designed for automobile HI modelling. Finally, the predicted HI obtained from M-LSTM network is used to map the RUL of an automobile according to Cox PHM. Figure 3 shows the flow chart of the proposed method.

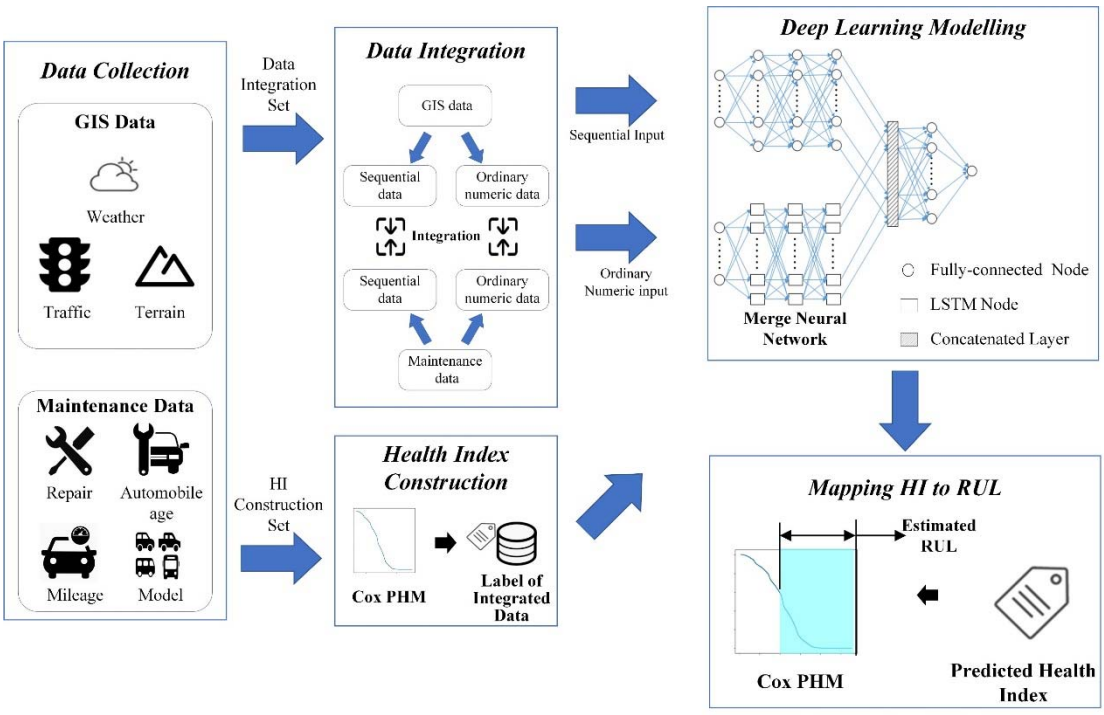


Figure 3. The flow chart of the proposed approach

4.1. GIS Data Collection

Automobile lifecycle can be affected by various factors includes weather, traffic and terrain. A fleet management company processes a large number of automobiles which the conditions of the working environment can be very different in terms of the GIS factors mentioned above. Hence, the GIS data of a certain working area need to be summarised and extracted using GIS software. Weather data such as temperature, rainfall, and sunshine hours of a certain working area can be obtained from the weather observation stations within the working area. Automobile lifecycle also can be affected by traffic condition. In an area with heavy traffic, the frequency of acceleration and deceleration tends to be higher, which may accelerate the failure of the automobile. Traffic data of a certain working area, such as traffic flow statistics, can be collected from the traffic department. The terrain is another aspect that can impact automobile life. In the mountainous area, automobiles have to frequently accelerate and decelerate, which speed up the failure of the automobile. The terrain data regarding elevation and slope in a certain working area can be analysed and extracted from the elevation map in GIS software. The taxonomic graph of the maintenance data and GIS data is shown in Figure 4. It is interesting to note that some GIS factors such as weather and traffic are changing along the time, while other factors such as terrain are relatively stable in a period. In other words, weather and traffic data can be considered as sequential data, while terrain data can be considered as ordinary numeric data in the study of PdM.

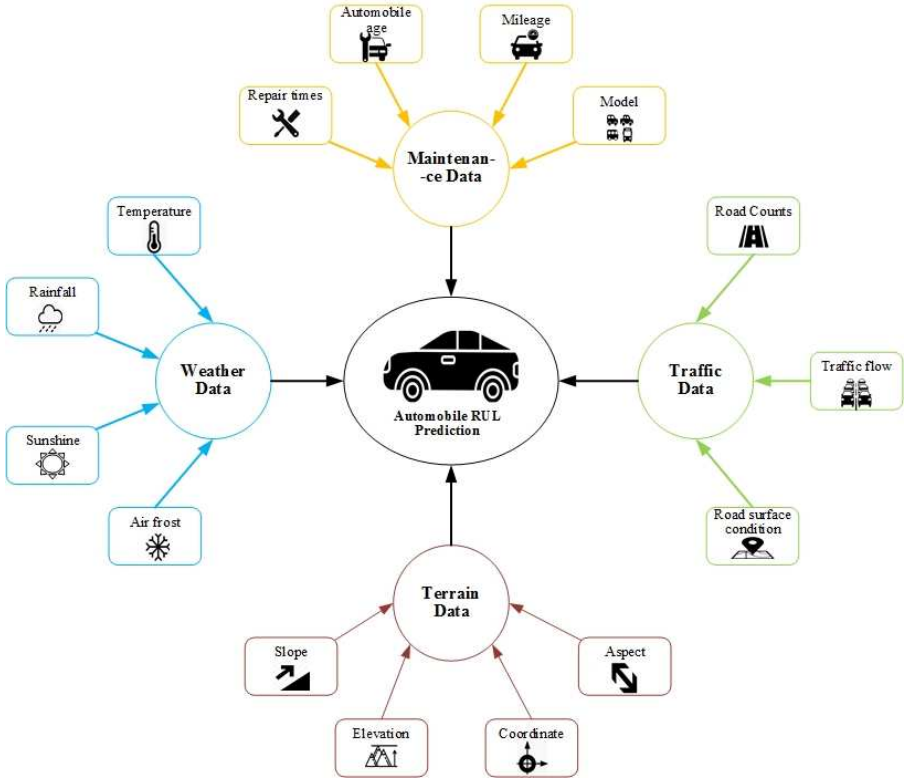


Figure 4. The taxonomic graph of the maintenance data and GIS data

4.2. Data Integration

After the maintenance data and GIS data is collected, they need to be integrated for automobile RUL modelling. Figure 5 shows the data integration of maintenance data and GIS data. Both maintenance data and GIS dataset consist of ordinary numeric data and sequential data. Firstly, the sequential data in both maintenance data and GIS dataset needs to be integrated. A maintenance data entry is denoted as $\{x_i, y_i\}$, where x_i is the sequential part of the input vector, y_i is the TBF (days). The sequential part of a GIS data entry is denoted as g_i. y_i ranges from a month to a few years, while g_i represents the average GIS condition in a month (30 days). The GIS data entries corresponding to x_i and y_i can be expressed by:

$$G_i = \{g_{i(1)}, g_{i(2)}, g_{i(3)}, \ldots\ldots, g_{i(n)}\} \quad (1)$$

where $n = \frac{y_i}{30}$

To integrate the sequential part of maintenance data and GIS data. y_i needs to be segmented by a month and x_i needs to be approximated correspondingly, which can be expressed by:

$$y_i \rightarrow \{y_{i(1)}, y_{i(2)}, y_{i(3)}, \dots, y_{i(n)}\} \tag{2}$$

$$x_i \rightarrow \{x_{i(1)}, x_{i(2)}, x_{i(3)}, \dots, x_{i(n)}\} \tag{3}$$

where $n = \frac{y_i}{30}$

After the sequential part of the maintenance data is transformed, it is concatenated with the sequential part of the GIS data. z_n is denoted as the data entry of the integrated sequential dataset *Z*, which can be expressed by:

$$z_n = \{[x_{i(n)}, g_{i(n)}], y_{i(n)}\} \tag{4}$$

The ordinary numeric data in maintenance dataset is transformed from nominal data such as automobile model, garage and area. Autoencoder, a deep learning algorithm, is used to transform the one-hot representation generated by nominal data into robust representation, which was detailed in our previous work [37]. An ordinary numeric data entry in maintenance dataset is denoted as o_i, while an ordinary numeric data entry in GIS dataset is denoted as h_i. Due to the transformation of the sequential data, the data volume is expanded. Both o_i and h_i need to be duplicated to match the data size of the sequential dataset. Finally, the ordinary numeric data in maintenance dataset and GIS dataset is concatenated.

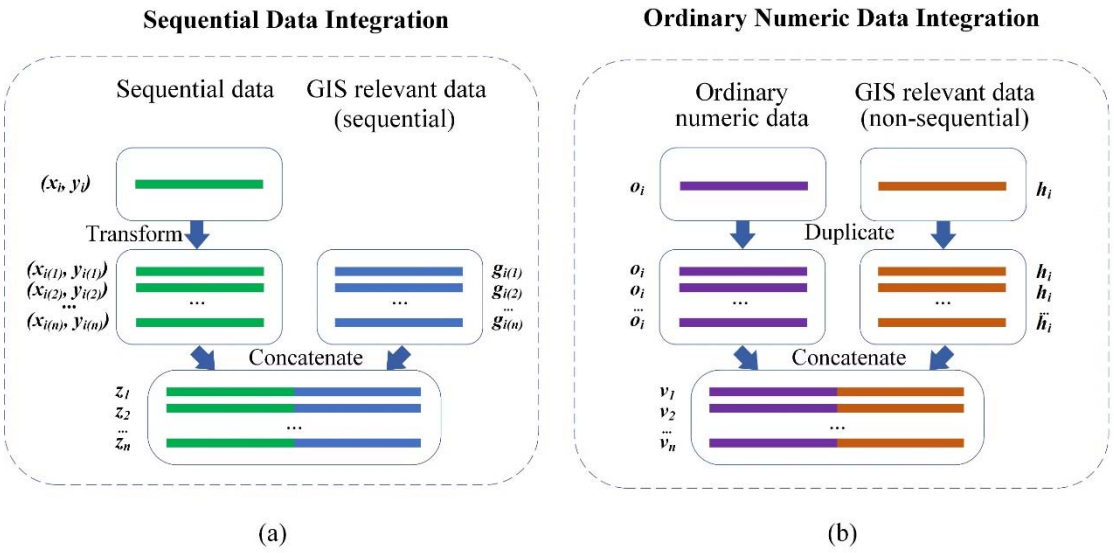


Figure 5. Data integration of maintenance data and GIS data

4.3. Health Index Construction

After the maintenance data and GIS data is integrated, the HI need to be estimated. The data label y_i is segmented by month and become time-to-failure (TTF). If TTF is used as the data label, automobile degradation is assumed to follow a linear distribution, which is not a common pattern in automobile failure modelling. An appropriate data label that can represent the automobile health condition needs to be determined. It can be seen from the literature that various methods have been deployed to construct the HI of machine, system or component. However, most of the existing methods are used to estimate the HI based on sensor data, which is strongly correlated to the actual health condition of machine, system or component. When the sensor data is not available, most of the existing methods are not available. Cox PHM, as a prevailing statistical method in reliability analysis, is used to estimate the degradation trend of the automobile based on maintenance data.

Cox PHM is used to analyse the relationship between different features and hazard

function. The covariate is denoted as β_p and the input vector is denoted as X_p. The Cox PHM is denoted as:

$$h(t, X) = h_0(t)\exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p) \quad (5)$$

where $h_0(t)$ is the maximum likelihood estimator proposed by Breslow [38].

Then a Cox PHM is established using the data in the HI construction set. For the construction of Cox PHM, the mean and standard deviation (StD) of the sequential GIS data (before automobile start to run) are first extracted. Then, the extracted GIS features are adopted in conjunction with the features in maintenance data to build a Cox PHM. The data in data integration set is fed into the Cox PHM to obtain the hazard function of the automobile. The hazard function of a Cox PHM is shown in Figure 6. Since the automobile TBF is segmented by *n* times and therefore *n* survival times are obtained. The corresponding HI in the integrated dataset is approximated by sending the survival time to the hazard function. When a HI is predicted, the hazard function is used to map the RUL.

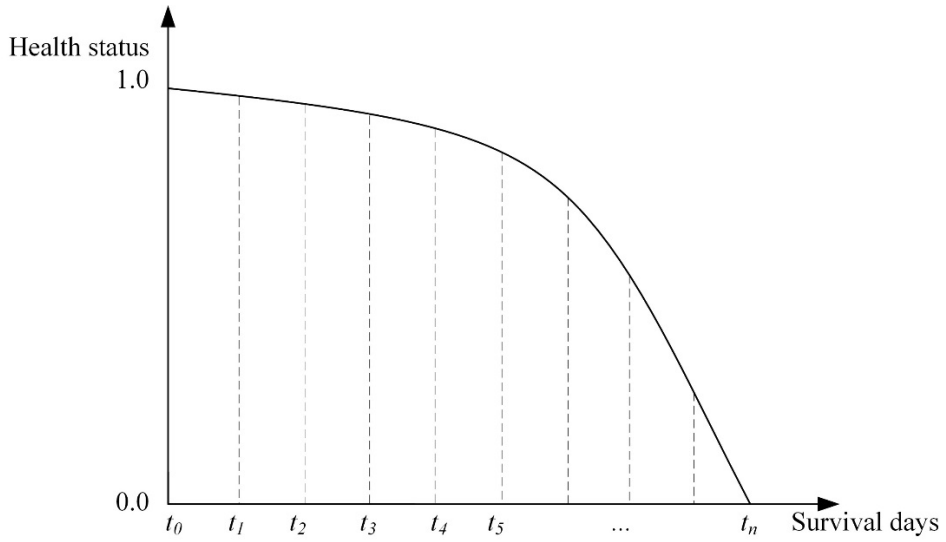


Figure 6. The hazard function of a Cox PHM

4.4. M- LSTM: A Novel Deep Learning Structure for Data Integration

Each typical deep learning structure has its expertise. For example, LSTM network is good at processing sequential data, and FCNN is good at process ordinary numeric data. In order to process both types of data mentioned above, a novel deep learning structure called M-LSTM network is designed for modelling HI based on both sequential data and ordinary numeric data. The structure of M-LSTM network takes the advantages of LSTM network and fully-connected network to handle the sequential data and ordinary numeric data simultaneously. Figure 7 indicates the structure of M-LSTM network. There are three major paths in M-LSTM network:

1. **Ordinary numeric data processing path:** Since an ordinary numeric data entry is a one-dimensional vector, the ordinary numeric data obtained from Section 4.2 is directly processed by fully-connected layers in the neural network. Hence, a two-layers fully-connected sub-network is designed in this path.

2. **Sequential data processing path:** LSTM network is expertise in learning the sequential patterns within data. The technical details of the LSTM network can be found in [39]. The sequential data obtained from Section 4.2 is further processed to three-dimensional format (i.e. features, data size and times). In order to learn the hidden patterns within the sequential data, a two-layer LSTM sub-network is deployed in this path. Moreover, a flatten layer is set to transform the output of LSTM layer for further data integration.

3. **Data integration path:** After the abstract representation is learnt by both paths mentioned above, a data integration path is needed to fuse the representation and implement the regression task to predict the automobile HI. In this path, a concatenate layer is deployed to concatenate the output from both sub-networks. Then, two fully- connected layers are

employed to fuse the concatenated representation further and learn the hidden patterns relevant to automobile HI.

There are several essential components and parameters in M-LSTM network, such as optimiser, the number of neurons and batch size, need to be determined in the actual case. Besides, to avoid overfitting, batch normalisation, l_2 regularizer and dropout techniques are deployed in M-LSTM network.

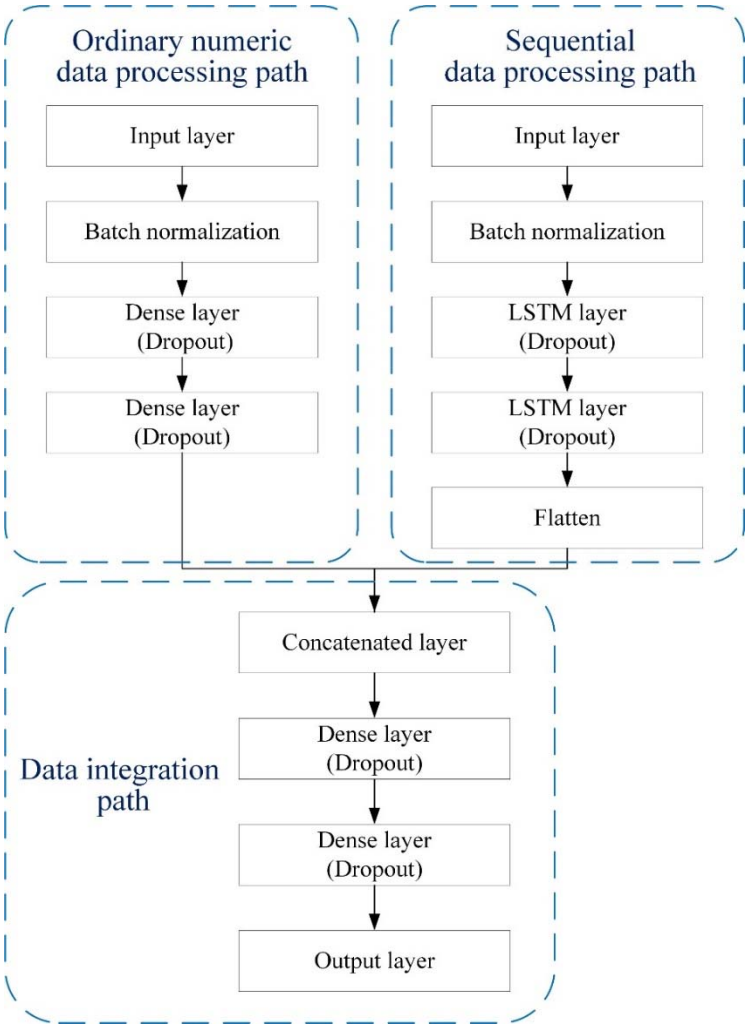


Figure 7. The structure of M-LSTM network

After the HI is predicted, it is used to estimate automobile RUL using the hazard function of Cox PHM. The survival time of an automobile when HI is 0 is denoted as

T_e, where the automobile is deemed as faulty. The True RUL at the time *i* is denoted as T_i. The predicted HI at the time *i* is denoted as R_i. The hazard function of the automobile is denoted as $f()$. The estimated RUL can be expressed by:

$$\text{Estimated RUL} = T_e - f(R_i) \tag{6}$$

The actual RUL of the automobile can be expressed by:

$$\text{Actual RUL} = T_e - T_i \tag{7}$$

In geography, since the terrain of an area is relatively stable during a period of time, the terrain is considered as constant along with time. The terrain data in RUL modelling is deemed as ordinary numeric data.

5. Experimental Setup

5.1. Data

The maintenance data used in this study records the automobile engine maintenance history. It was provided by a sizable fleet service company in the UK. The company has a strong interest in the prediction of the maintenance time of an automobile. An accurate prediction of automobile RUL can be beneficial to the company in terms of maintenance planning, job scheduling and spare parts inventory management. Maintenance data records the relevant data when maintenance is implemented. Some automobiles experienced multiple maintenances in the past. It is challenging to construct the HI for the automobile which has multiple maintenances since the health condition after maintenance cannot be recovered to 1, and therefore it is hard to estimate. Hence, only the first maintenance record of automobiles is considered in this study. The maintenance dataset was collected from 2009 to 2017, which contains 6,584 data entries. The average TBF of the automobiles in the dataset is 1201.7 days, while the standard deviation is 591.7 days. The features relevant to the automobile life cycle

were extracted from the dataset. The features in maintenance data are shown in Table 1.

Table 1. The features in maintenance data

Numeric Features		Nominal Features	
Feature	Description	Feature	Description
Vage	The age of automobile	*Model*	The model of automobile
CumM	The cumulative miles when a failure occurs	*Garage*	The garage of automobile
Model_Year	The age of the automobile model	*Area*	The area of automobile
Seq	A time index for automobile		
Regions	Four binary attributes used to identify regions		

The maintenance data was first pre-processed. The data entries with missing values and abnormal values were removed in this stage. The TBF of automobiles ranges from 0 to 2770. The early failure of an automobile is caused by multi-reasons, such as product quality and accidents. There are two types of features in the maintenance dataset, which are numeric features and nominal features. The numeric features were further classified into sequential features and ordinary numeric features. In this dataset, *Vage, CumM, Model_Year, and Seq* are the features change alone with time, and therefore these four features are deemed as sequential features, while *Regions* is considered as an ordinary numeric feature. The nominal features cannot be directly used for regression modelling. In our previous study, the nominal features were first converted to one-hot encoding format before a three-layers autoencoder was designed to transform the nominal data in one-hot encoding format to robust representation [37]. The dimensions of robust features converted from autoencoder are 16. The features mentioned above were used for RUL modelling. Besides these features, the maintenance data also contains the automobile start date and failure date, which are useful for the data transformation in the data integration stage.

The GIS data was collected according to the mobility area and time. Firstly, there are over 60 garages in the fleet management company. The garage location was set as the centre of the

mobility area for an automobile. With the consideration of the daily mileage of automobile, a circular area with a radius of 30km was set as the mobility area. The automobiles in the same garage are considered running in the same mobility area. All the garage location and mobility area were plotted in ArcGIS software. Figure 8 shows two examples of garage and mobility area. Secondly, the GIS data in this area between the automobile start date and failure date was extracted and summarised. There are three types of GIS data considered in this study, which are weather, traffic and terrain.

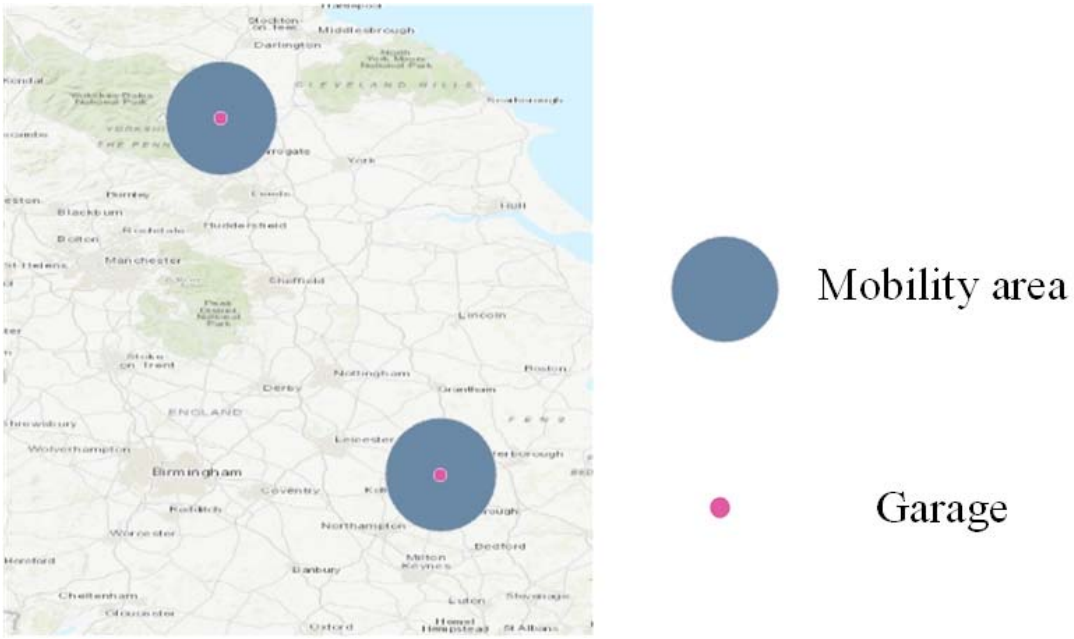


Figure 8. Examples of automobile garage and mobility area

Weather condition may affect the performance of the automobile. The weather data was collected from the website of the MET office [40], UK. There are approximately 40 weather observation stations all around the UK, with data collected from 2009 to 2017. Some mobility areas may cover multiple weather observation stations, and therefore the mean value of the data from these weather observation stations was adopted. Meanwhile, some other mobility areas do not cover any weather observation station, a linear approximation based on the data of several near weather observation stations was implemented to yield the weather data in these mobility areas. The weather data used in this study was sampled monthly and includes the

following features: *Name of observation station*, *Time*, *Rainfall*, *Max_temp*, *Min_temp*, *Days of air frost*, and *Sunshine*. With the *Name of observation station*, the location of weather observation station was identified and plotted in ArcGIS software. Then, the weather data in each mobility area was summarised.

Traffic condition is another impact that can affect automobile lifecycle. The traffic data was collected from the public dataset of the department of transport, UK. The traffic data of over 200 local authorities all around the UK, which was collected from 2009 to 2017, were available. The traffic data was sampled yearly and includes the following features: *Local authority name*, *Year*, *Link length (km)*, *Link length (miles)*, *Cars and taxis*, and *All motor vehicles*. In the above features, *Link length (km)* and *Link length miles* were converted to each other. Hence, *link length (km)* was dropped. Similar to the process of weather data, the *Local authority name* is used to identify the locations of the local authority in ArcGIS software, before the traffic data of each mobility area was summarised.

The automobile lifecycle also can be impacted by the terrain condition. The terrain data used in this study was extracted from the elevation map in ArcGIS software. The terrain features of a mobility area were directly extracted and summarised. The terrain data includes the following features: *Mean elevation*, *Maximum elevation*, *Minimum elevation*, *StD of elevation*, *Mean slope*, *Maximum slope*, *StD of slope, Mean aspect, Longitude and Latitude*. The summarisation of all the GIS data is shown in Table 2.

Some features in GIS data can be highly correlated. The redundant features may damage the algorithm performance. Hence, it is necessary to remove redundant features. Heat map is a data visualisation tool to identify the correlation between different features. Figure 9 shows the heat map of the extracted features of the GIS data. The highly correlated features (correlation coefficient close to 1 or -1) which correlation coefficient is over 0.8 were identified via heat

map and then were removed. In the heat map, there are 42 features in total. After the removal of the GIS features, only 22 features were kept for further modelling.

Table 2. The summarisation of the GIS data

Types	Sampling frequency	Feature	Description	Feature	Description
Weather	Monthly	*Rainfall*	The rainfall (mm)	*Max_temp*	The maximum temperature (°C)
		Min_temp	The minimum temperature (°C)	*Days of air frost*	The days of air frost
		Sunshine	The sunshine hours		
Traffic	Yearly	*Link length (km)*	The total length of each junction to junction link on the major road network	*Cars and taxis*	The amount of cars and taxis
		All motor vehicles	The amount of all motor vehicles		
Terrain	NIL	*Mean elevation*	The mean elevation of mobility area	*Maximum elevation*	The maximum elevation of mobility area
		Minimum elevation	The minimum elevation of mobility area	*StD of elevation*	The Standard deviation of elevation of mobility area
		Mean slope	The mean slope of mobility area	*Maximum slope*	The maximum slope of mobility area
		StD of slope	The Standard deviation of slope of mobility area	*Mean aspect*	The mean aspect of mobility area
		Longitude	The latitude of the longitude	*Latitude*	The latitude of the garage

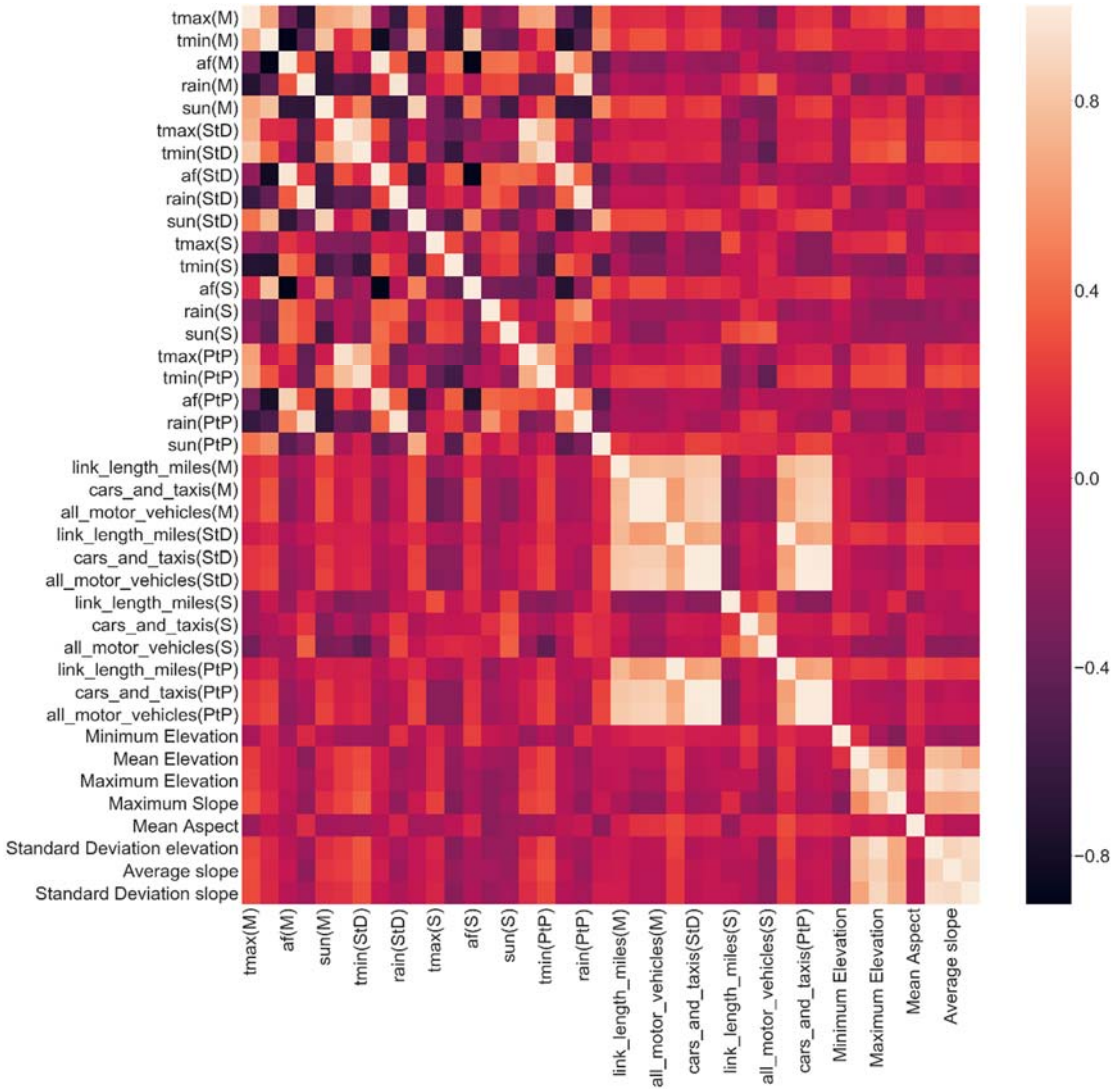


Figure 9. The heatmap of the GIS features

5.2. Experimental Setup

After both maintenance data and GIS data were collected and identified, the first step is to integrate the data. There are 6,584 data entries in the maintenance dataset. 5-fold cross-validation was adopted. In each experiment of 5-fold cross-validation, 20% of data entries in maintenance dataset were used for data integration and then RUL modelling. After the data transformation, 20% of the data will generated over 64,000 data entry, which is sufficient to train a neural network with satisfactory performance. Meanwhile, the HI construction is important. The rest 80% of data entries were used in HI construction to establish a reliable Cox PHM. It can be seen that the sampling frequency of maintenance data and three types of GIS

data are different. The maintenance data is collected only when the event of automobile failure occurs, which cannot follow a specific sampling frequency.

The weather and traffic data were collected monthly and yearly, respectively. Hence, they were deemed as sequential data in this study. The terrain data is relatively stable for a long period. Therefore, it was deemed as constant. In order to integrate both maintenance data and GIS data, the data needs to be transformed. Firstly, the sequential data in both maintenance dataset and GIS dataset was integrated. In this study, all the data was transformed to the monthly level. The idea was detailed in Section 4.2. Since the automobile start date and failure date of each maintenance data entry are known, each maintenance data entry was segmented by month. The monthly segments of all automobiles constituted a new maintenance dataset. Meanwhile, the weather data is sampled monthly, and therefore it can be directly used for data integration. The traffic data is sampled yearly. Hence, the segmentation approach was also implemented to the traffic data to generate a new traffic dataset with the monthly approximation. After all the sequential data was transformed, the sequential maintenance data, weather data and traffic data were concatenated in the feature direction. Besides, the terrain data and the ordinary numeric data in the maintenance dataset were duplicated and concatenated. After the data was integrated, a new sequential dataset and a new ordinary numeric dataset which contain over 64,000 data entries was generated. Then the sequential data further processed using time window approach. The sequence length was set as 5, which means the input data of the previous 5 step is used to predict the HI in the next step.

Cox PHM is an important model in reliability analysis. In order to establish a reliable Cox PHM, not only the features in maintenance dataset were adopted, but the terrain features and the statistical features (weather and traffic) of weather and traffic data were taken. Cox PHM is sensitive to the covariates. Hence, all the features were first used to establish a Cox PHM. By calculating the model coefficient of each feature, the redundant features were identified. Those

features which coefficient is negligible were removed from the dataset to enhance the robustness of the model. After a Cox PHM was established, it is used to approximate the HI of the automobiles in data integration set. The hazard function of each automobile in the data integration set can be produced via the established Cox PHM. The hazard function is then used to calculate the HI of the integrated dataset obtained from the data integration stage. The approximated HI was used as the new label of the integrated data. Also, the hazard function of each automobile in the data integration set is used to estimate the RUL using the predicted HI. The data in this stage was ready for HI modelling and then RUL mapping.

After the data for RUL modelling was prepared, the parameters of the merged neural network need to be determined. Firstly, the size of the neural network is an important parameter in a neural network. After several trials, the number of neurons of each layer in the ordinary numeric data processing path and sequential data processing path was set as 100. The number of lookbacks in LSTM layer was set as five, which means the variables of the previous five steps is used to predict the HI of the next stage. The number of neurons of in the first and second layers in the data integration path was set as 600 because the data dimension obtained from the two sub-networks were 600. The number of neurons in the output layer was set as one due to the neural network is designed for regression. The rectified linear unit was adopted as the activation function of all the fully-connected layers except the output layer. The activation function of the output layer was set as linear due to the modelling mission is regression. Secondly, the parameters relevant to the training process were determined. RMSprop [41] was selected as the optimiser with the learning rate set as 0.005. The batch size, training epochs and dropout rate were set as 300, 10, and 0.5, respectively.

In this study, in order to reveal the performance of M-LSTM network, four algorithms were deployed for benchmarking. The benchmarking algorithms are artificial neural network (FCNN), long-short term memory (LSTM) network, deep convolutional neural network

(DCNN) [10], and support vector machine (SVM). Two scenarios were set in the experimental study. In scenario 1, the five algorithms mentioned above were adopted for RUL modelling based on the maintenance data. The algorithm performance of M-LSTM network was revealed with the comparison to that of the benchmarking algorithms. In scenario 2, the GIS data was introduced. Whether the RUL modelling accuracy can be leveraged with the enrichment of GIS data can be revealed. Meanwhile, the identification of the impact of GIS features on automobile lifecycle needs to be revealed. Due to the training mechanism of deep learning, the results of deep learning are slightly different from time to time. Besides, the GIS features are not strongly correlated to the automobile lifecycle and therefore, it is challenging to reveal the impact of a specific GIS feature on automobile RUL. Hence, the impact of group features which are weather, terrain and traffic factors on automobile lifecycle were studied. All tests were conducted on an Intel i5-7300 2.80Ghz laptop with Nvidia GeForce GTX 1050 graphics card.

5.3. Performance Evaluation

There are two metrics adopted in this experiment to reveal the algorithm performance, which are the model correlation coefficient (MCC) and root mean square error (RMSE). These two metrics are widely used in the study of regression. MCC is used to measure the correlation between two variables and can be expressed mathematically as:

$$\text{MCC} = \frac{S_{PA}}{\sqrt{S_P S_A}}, \tag{8}$$

where,

$$S_{PA} = \frac{\sum_i (p_i - \bar{p})(a_i - \bar{a})}{n-1};\ S_P = \frac{\sum_i (p_i - \bar{p})^2}{n-1};$$

$$S_A = \frac{\sum_i (a_i - \bar{a})^2}{n-1};$$

p_i is the predicted value and $\bar{p}$ is the average of the predicted value. a_i is the actual value and

the $\bar{a}$ is the average actual value. n is the number of training data.

RMSE is a scale-dependent metric which measures the difference between the prediction value and the actual value. It is 0 if there is no difference between the prediction value and the actual values. The expression of RMSE is:

$$\text{RMSE}=\sqrt{\frac{\sum_i(p_i-a_i)^2}{n}} \tag{9}$$

Besides, the training time of each algorithm, which is the modelling time for HI, were compared to reveal the computational cost.

6. Experimental results

6.1. Scenario 1: The comparison of M-LSTM network with the Prevailing Machine Learning Algorithms

In this scenario, in order to reveal the algorithm performance of M-LSTM network, four machine learning algorithms, which are LSTM network, FCNN, DCNN and SVM were used for modelling based on maintenance data. After 5-fold cross-validation, the StD of MCC and RMSE of HI and RUL prediction were compared to reveal the algorithm performance.

The modelling results of five machine learning algorithms were shown in table 3. It can be seen that M-LSTM network achieved the highest MCC and lowest RMSE in RUL modelling, while the algorithm performance of LSTM network in terms of MCC and RMSE are more stable due to the lower StD. The algorithm performance in terms of MCC and RMSE for M-LSTM network RUL modelling is slightly better than that of FCNN. Meanwhile, SVM not only showed the worst algorithm performance in terms of MCC and RMSE, but also it took the longest time for HI modelling. Hence, the deep learning algorithms show merits in RUL modelling.

Table 3. The results of RUL modelling based on maintenance data

	M-LSTM	LSTM	FCNN	DCNN	SVM
MCC_Mean	**0.9778**	0.9713	0.9743	0.9711	0.9578
MCC_StD	0.0071	**0.0041**	0.0103	0.0059	0.0045
RMSE_Mean (days)	**349.4**	368.7	356.5	393.1	396.7
_RMSE_StD (days)	28.4	**19.5**	31.7	48.2	29.8
Training time (s)	66.5	109.4	**11.7**	21.9	3120.5

6.2. Scenario 2: RUL Modelling with the Enrichment of GIS Data

In this scenario, the GIS data was introduced into HI and RUL modelling. The pre-processed weather and traffic data were concatenated with the sequential part of the maintenance data, while the pre-processed traffic data were concatenated with the ordinary numeric part of the maintenance data. The algorithms were used for HI and RUL modelling based on the enriched dataset.

Figure 10 and 11 reveals the results of RUL modelling in terms of MCC and RMSE based on maintenance data in conjunction with/ without GIS data. It can be seen that the MCC of all the algorithms were promoted with the enrichment of GIS data, except DCNN. The algorithm performance of M-LSTM network in terms of MCC was promoted by 0.13%, while the StD shrunk by 63.2%. With the enrichment of GIS data, M-LSTM network still achieved the highest MCC. LSTM network got the most considerable improvement with the help of GIS data, while its StD enlarged dramatically. Meanwhile, with the help of GIS data, the algorithm performance in terms of RMSE of all the algorithms was decreased except LSTM network. M-LSTM network witnessed the largest decrease in terms of RMSE, which is about 3.9%. The lowest RMSE achieved in this case is 335.8 days. What is also apparent is that the StD of RMSE for M-LSTM network is far lower than other algorithms with the help of GIS data.

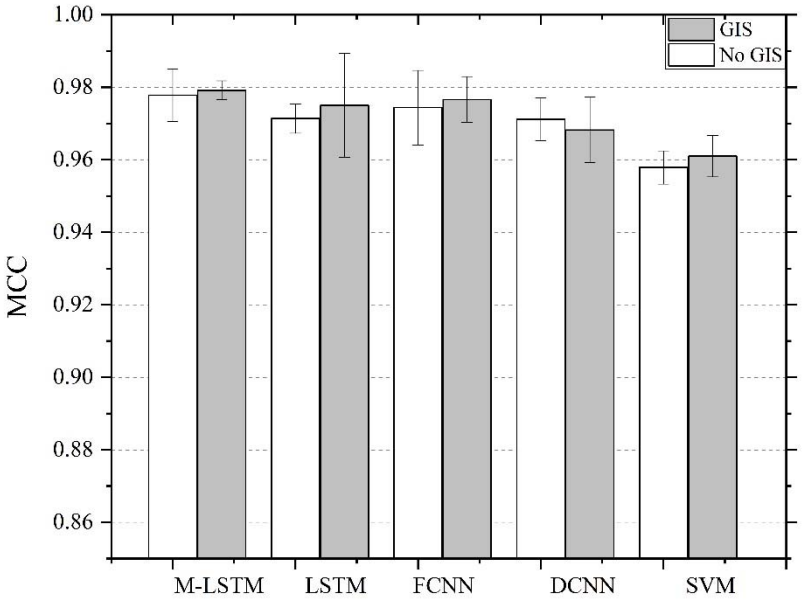


Figure 10. The MCC of RUL prediction based on maintenance data with/ without GIS data

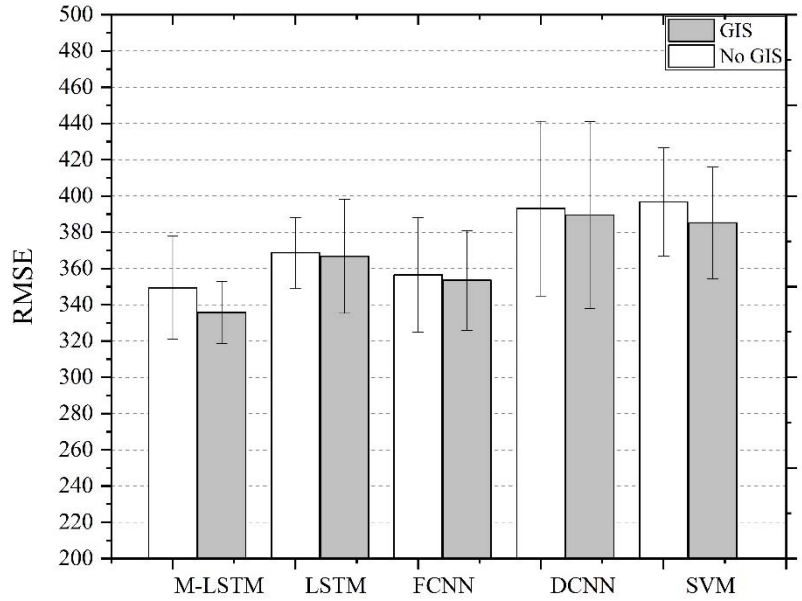


Figure 11. The RMSE of RUL prediction based on maintenance data with/ without GIS data

M-LSTM network achieved the best performance in comparison with the benchmarking algorithms. In order to further reveal how the GIS features impact the automobile lifecycle, it is worthwhile the plot the actual RUL and the predicted RUL with and without GIS features of some automobiles. Figure 12 shows the predicted RUL for automobile #21, #23, #28 and #15, which are deemed representative for most automobiles. The RUL prediction of these four automobiles shows different patterns. Figure 12 (a) shows the error of predicted RUL of both models trained with and without GIS features for automobile #21 was large in the beginning and ending stage, while the middle stage was relatively accurate. Figure 12 (b) indicates that the RUL prediction error of automobile #23 was large in the beginning, and then it shrank in the next ten months. With GIS features, the prediction errors in the first three months are lowered, while the errors are enlarged in the ending stage. The predicted RUL in Figure 12 (c) reveals that the predicted error of automobile #28 was stable and not dramatic. With GIS features, the prediction errors of automobile #23 and #28 in the beginning stage are lowered, while the errors are enlarged in the ending stage. The predicted error of Figure 12 (d) was large in all the stages without convergence.

The merits of the RUL prediction of automobile #15 with GIS features are obvious. It can be seen from the four figures that the trend of the predicted RUL was highly correlated to the actual RUL, while the predicted error varies from each other. The comparison of the RUL prediction from the M-LSTM network trained with and without GIS features is further revealed in Table 4. It can be seen that with the help of GIS features, the MCC of M-LSTM network can be promoted. The RMSE of #21 and #15 obviously decreased with the help of GIS features, while the RMSE of # 23 and #28 slightly increased.

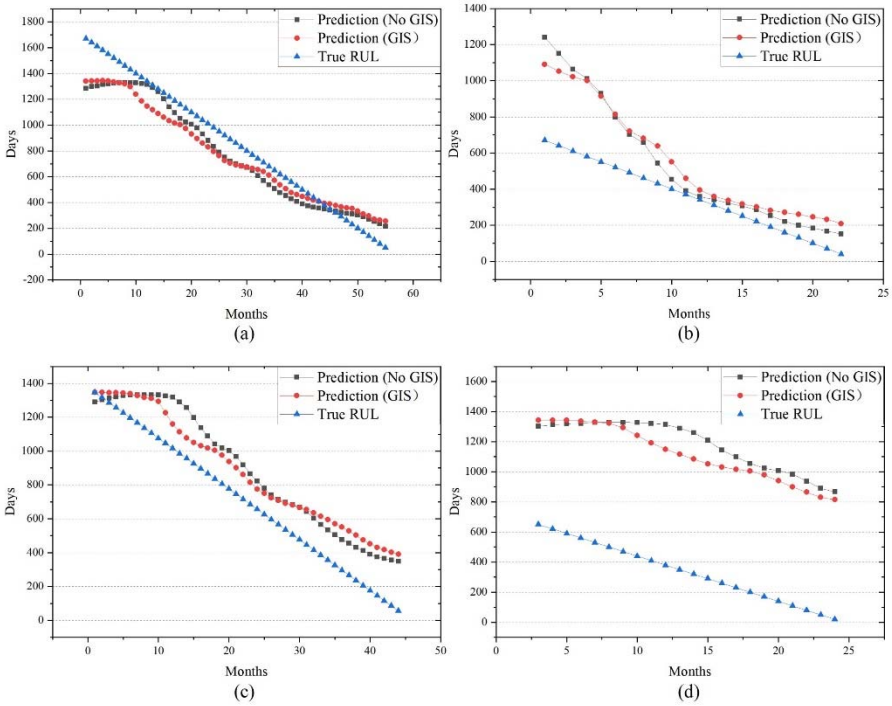


Figure 12. The predicted and actual RUL for four automobiles

Table 4. The results of RUL modelling for four automobiles

Automobile		21	23	28	15
MCC	No GIS	0.9799	0.9574	0.9801	0.8698
	GIS	**0.9896**	**0.9672**	**0.9919**	**0.9764**
RMSE	No GIS	144.1	**235.0**	**192.7**	830.5
	GIS	**110.8**	247.4	204.3	**769.9**

Since M-LSTM network showed merits in RUL modelling with the enrichment of GIS data, it was used to identify the impact of the group GIS features on automobile lifecycle. Figure 13 shows the MCC and RMSE of RUL modelling using M-LSTM network based on the maintenance data with different types of GIS data. With the help of different types of GIS data, the MCC of M-LSTM network improved and the RMSE reduced, which indicate that all three types of GIS features under investigation have a positive impact on automobile RUL modelling. Among these three types of features, M-LSTM network achieved higher MCC with the help of traffic features and achieve lower RMSE with the enrichment of weather features. Furthermore, with the enrichment of weather and traffic data, the Std of RMSE for LSTM network decreased

considerably. In contrast, the impact of terrain data on automobile lifecycle modelling is relatively low.

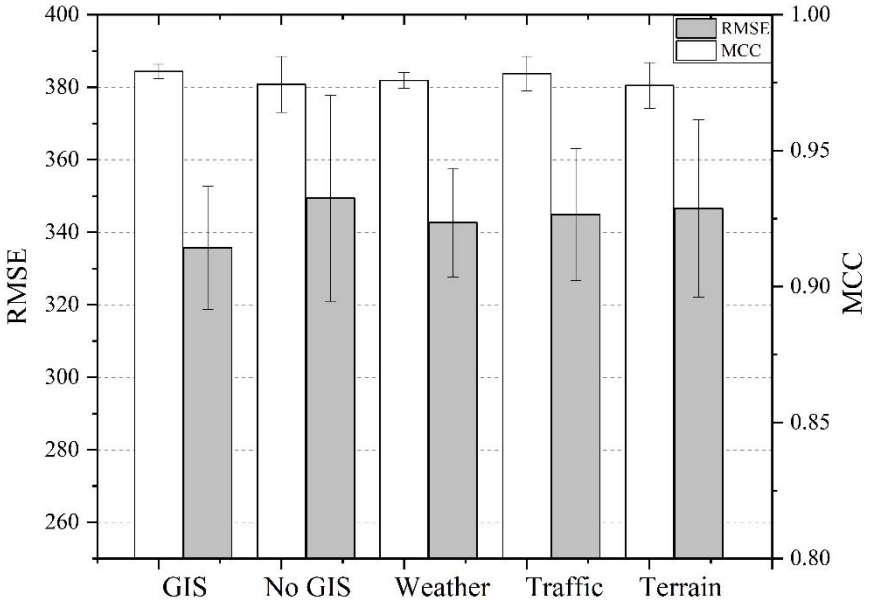


Figure 13. The MCC and RMSE of RUL modelling using M-LSTM network based on the maintenance data with different types of GIS data

7. Discussions

From the results of scenario 1, what obvious is that the M-LSTM network shows merits in modelling based on heterogeneous data, which contains sequential data and ordinary numeric data. It promoted the algorithm performance in terms of MCC and RMSE, while it requires less computational load than the LSTM network. Hence, it can be useful in heterogeneous data modelling. Meanwhile, the results of scenario 2 indicated the introduction of individual groups of GIS features promoted the algorithm performance of most algorithms used in this study in terms of MCC and RMSE. The introduction of the weather data has the largest positive impact on the algorithm performance in terms of MCC and RMSE of M-LSTM network, which indicate that weather factors have a higher impact on the automobile lifecycle. The algorithm performance of M-LSTM network in terms of MCC and RMSE is better than that of the

benchmarking algorithms, which indicates that M-LSTM network is more suitable in modelling based on multi-sources data. In the results, it can be seen that the predicted error was large when the overall life is lower than 1000 days. In the historical maintenance dataset, the feature values of data entry that TBF is lower than 1000 are similar to that of the data entry which TBF is higher than 1000. Since the automobile which TBF is lower than 1000 is the minority, it poses a challenge to the algorithm to correctly estimate its health status.

The introduction of GIS data promoted the algorithm performance of all the algorithms except DCNN witnessed the performance improvement in terms of MCC and RMSE with the enrichment of three types of GIS data. Among these algorithms, M-LSTM network achieved the highest MCC and lowest RMSE, which demonstrated that M-LSTM network shows merits in modelling with GIS data. However, the improvement of the algorithm performance in terms of MCC and RMSE are still limited. The main reason can be the GIS data collected in this study does not match the actual surrounding factors of the automobile mobility area of each automobile. For example, if there are a mountainous area and a plain area in the data collection area, the maximum elevation and the slope tends to be dramatic. However, there is not likely too many roads in such a mountainous area. Hence, the collected terrain data did show a positive impact on automobile RUL modelling. Its actual impact needs to be further investigated.

In the future, firstly, with the development of IIoT, the real-time telematics data of automobile can be collected during the in-use period. With the GPS information, the real-time GIS data can be collected from external data sources and then be transmitted to the cloud. How to integrate real-time telematics data, real-time GIS data, and maintenance data in the study of PdM will be further explored. The prediction accuracy of the automobile RUL prediction model can be further leveraged with the introduction of real-time telematics and GIS data and a more suitable data integration scheme. Secondly, the architecture of M-LSTM network can be further

explored. The M-LSTM designed in this paper adopted two sub-networks, which are FCNN and LSTM network, for feature extraction. Then the extracted features are sent to concatenated layer for integration before an FCNN is deployed for the final RUL prediction.

With the development of machine learning, there are various emerging algorithms proposed recently. It is worthwhile to explore how to improve the feature extraction in deep learning so to promote the RUL prediction accuracy. Furthermore, the actual impact of GIS features on automobile lifecycle needs to be further investigated. Since deep learning is a type of black-box algorithm, which cannot indicate how the features affect the final output, it cannot offer too many insights into the fleet management company. The GIS features are highly correlated, which leads to the difficulty of the identification of feature importance. If the relationship between the correlated GIS features and the automobile RUL can be identified, it can be beneficial to the fleet management company to optimise maintenance management. Hence, future works also target on developing a new approach to identify how correlated GIS features impact automobile lifecycle.

8. Conclusion

In automobile PdM, existing studies only focus on modelling using sensor data or maintenance data which is directly relevant to automobile failure. Other relevant factors, such as weather, traffic and terrain, which can also impact automobile lifecycle, have not been considered in automobile PdM. To fill this research gap, in this paper, we have initiated the research on an M-LSTM network-based predictive maintenance enriched by GIS data. In this approach, the data collection and pre-processing schemes for GIS data were studied, which enable the integration of GIS data and historical maintenance data. Furthermore, a new deep learning structure called the M-LSTM network was proposed for the modelling based on multi-source data. The experimental study based on the real-world data obtained from a fleet management company validated the effectiveness of the proposed approach. Experimental results have

revealed that the M-LSTM network has shown its merits in multi-source data modelling, while the integration of GIS data is beneficial in improving the prediction accuracy of RUL. Based on the promising results achieved, it encourages the fleet management company to further exploit GIS data jointly with archived maintenance data for fleet health and maintenance management, better job scheduling and garage operational management, etc.

References

[1] Zonta T, da Costa CA, da Rosa Righi R, de Lima MJ, da Trindade ES, Li GP. Predictive maintenance in the Industry 4.0: A systematic literature review. Computers & Industrial Engineering. 2020:106889.
[2] Han X, Wang Z, Xie M, He Y, Li Y, Wang W. Remaining useful life prediction and predictive maintenance strategies for multi-state manufacturing systems considering functional dependence. Reliability Engineering & System Safety. 2021;210:107560.
[3] Zhang W, Li X, Ma H, Luo Z, Li X. Transfer learning using deep representation regularization in remaining useful life prediction across operating conditions. Reliability Engineering & System Safety. 2021;211:107556.
[4] Hu J, Chen P. Predictive maintenance of systems subject to hard failure based on proportional hazards model. Reliability Engineering & System Safety. 2020;196:106707.
[5] Pham BT, Bui DT, Prakash I, Dholakia M. Hybrid integration of Multilayer Perceptron Neural Networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. Catena. 2017;149:52-63.
[6] Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM. Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities. Information Fusion. 2019;50:71-91.
[7] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436-44.
[8] Aria A, Lopez Droguett E, Azarm S, Modarres M. Estimating damage size and remaining useful life in degraded structures using deep learning-based multi-source data fusion. Structural Health Monitoring. 2020;19:1542-59.
[9] Zhang W, Zhang Y, Zhai J, Zhao D, Xu L, Zhou J, et al. Multi-source data fusion using deep learning for smart refrigerators. Computers in Industry. 2018;95:15-21.
[10] Li X, Ding Q, Sun J-Q. Remaining useful life estimation in prognostics using deep convolution neural networks. Reliability Engineering & System Safety. 2018;172:1-11.
[11] Xu X, Wu Q, Li X, Huang B. Dilated Convolution Neural Network for Remaining Useful Life Prediction. Journal of Computing and Information Science in Engineering. 2020;20.
[12] Li X, Zhang W, Ma H, Luo Z, Li X. Data alignments in machinery remaining useful life prediction using deep adversarial neural networks. Knowledge-Based Systems. 2020:105843.
[13] Li H, Zhao W, Zhang Y, Zio E. Remaining useful life prediction using multi-scale deep convolutional neural network. Applied Soft Computing. 2020;89:106113.
[14] Al-Dulaimi A, Zabihi S, Asif A, Mohammed A. NBLSTM: Noisy and Hybrid CNN and BLSTM-based Deep Architecture for Remaining Useful Life Estimation. Journal of Computing and Information Science in Engineering. 2019:1-14.
[15] Zhang Y, Hutchinson P, Lieven NA, Nunez-Yanez J. Remaining Useful Life Estimation Using Long Short-Term Memory Neural Networks and Deep Fusion. IEEE Access. 2020;8:19033-45.
[16] Mo Y, Wu Q, Li X, Huang B. Remaining useful life estimation via transformer encoder enhanced by a gated convolutional unit. Journal of Intelligent Manufacturing. 2021:1-10.
[17] Shi Z, Chehade A. A dual-LSTM framework combining change point detection and

remaining useful life prediction. Reliability Engineering & System Safety. 2021;205:107257.
[18] Yan T, Lei Y, Li N, Wang B, Wang W. Degradation Modeling and Remaining Useful Life Prediction for Dependent Competing Failure Processes. Reliability Engineering & System Safety. 2021:107638.
[19] Zhang S, Liu C, Su S, Han Y, Li X. A feature extraction method for predictive maintenance with time-lagged correlation–based curve-registration model. International Journal of Network Management. 2018;28:e2025.
[20] Chen Z, Li Y, Xia T, Pan E. Hidden Markov model with auto-correlated observations for remaining useful life prediction and optimal maintenance policy. Reliability Engineering & System Safety. 2017.
[21] Aremu OO, Hyland-Wood D, McAree PR. A Relative Entropy Weibull-SAX framework for health indices construction and health stage division in degradation modeling of multivariate time series asset data. Advanced Engineering Informatics. 2019;40:121-34.
[22] Li N, Gebraeel N, Lei Y, Fang X, Cai X, Yan T. Remaining useful life prediction based on a multi-sensor data fusion model. Reliability Engineering & System Safety. 2021;208:107249.
[23] Patil RB, Patil MA, Ravi V, Naik S. Predictive modeling for corrective maintenance of imaging devices from machine logs. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC): IEEE; 2017. p. 1676-9.
[24] Dangut MD, Skaf Z, Jennions IK. An integrated machine learning model for aircraft components rare failure prognostics with log-based dataset. ISA transactions. 2020.
[25] Calabrese M, Cimmino M, Fiume F, Manfrin M, Romeo L, Ceccacci S, et al. SOPHIA: An event-based IoT and machine learning architecture for predictive maintenance in industry 4.0. Information. 2020;11:202.
[26] Kobayashi S, Fukuda K, Esaki H. Mining causes of network events in log data with causal inference. 2017 IFIP/IEEE Symposium on Integrated Network and Service Management (IM): IEEE; 2017. p. 45-53.
[27] Tehrany MS, Pradhan B, Jebur MN. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. Journal of hydrology. 2014;512:332-43.
[28] Naghibi SA, Pourghasemi HR, Dixon B. GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran. Environmental monitoring and assessment. 2016;188:44.
[29] Rahmati O, Pourghasemi HR, Melesse AM. Application of GIS-based data driven random forest and maximum entropy models for groundwater potential mapping: a case study at Mehran Region, Iran. Catena. 2016;137:360-72.
[30] Massawe BHJ, Subburayalu SK, Kaaya AK, Winowiecki L, Slater BK. Mapping numerically classified soil taxa in Kilombero Valley, Tanzania using machine learning. Geoderma. 2018;311:143-8.
[31] Zhao W, Du S. Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach. IEEE Transactions on Geoscience and Remote Sensing. 2016;54:4544-54.
[32] Zhang Y, Zhang Q, Farnoosh A, Chen S, Li Y. GIS-Based Multi-Objective Particle Swarm Optimization of charging stations for electric vehicles. Energy. 2019;169:844-53.
[33] Allah Bukhsh Z, Saeed A, Stipanovic I, Doree AG. Predictive maintenance using tree-based classification techniques: A case of railway switches. Transportation Research Part C: Emerging Technologies. 2019;101:35-54.
[34] Chen C, Liu Y, Sun X, Di Cairano-Gilfedder C, Titmus S. Automobile Maintenance Modelling Using gcForest. 2020 IEEE 16th International Conference on Automation Science and Engineering (CASE): IEEE; 2020. p. 600-5.
[35] Wang S, Liu Y, Cairano-Gilfedder CD, Titmus S, Naim MM, Syntetos AA. Reliability Analysis for Automobile Engines: Conditional Inference Trees. Procedia CIRP. 2018;72:1392-7.
[36] Chen C, Liu Y, Wang S, Sun X, Di Cairano-Gilfedder C, Titmus S, et al. Predictive

maintenance using cox proportional hazard deep learning. Advanced Engineering Informatics. 2020;44:101054.
[37] Chen C, Liu Y, Sun X, Cairano-Gilfedder C, Titmus S. Automobile maintenance prediction using deep learning with GIS data. Procedia CIRP. 2019;81:-.
[38] Breslow NE. Analysis of survival data under the proportional hazards model. International Statistical Review/Revue Internationale de Statistique. 1975:45-57.
[39] Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation. 1997;9:1735-80.
[40] Weather and Climate Change-Met Office. 2018.
[41] Tieleman T, Hinton G. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural networks for machine learning. 2012;4:26-31.