Transfer Learning Architecture Approach for Smart Transportation System

Sujatha Krishnamoorthy(✉)

Assistant Professor School of Science and Technology, Wenzhou Kean University, Zhejiang, China
sujatha@wku.edu.cn

Abstract. An intelligent and smart transportation system aims at effective transportation and mobility usage in smart cities. In recent years, modern transportation networks have undergone a rapid transformation. This has resulted in a variety of automotive technology advances, including connected vehicles, hybrid vehicles, Hyperloop, self-driving cars and even flying cars, as well as major improvements in global transportation networks. Because of the open existence of smart transportation system as a wireless networking technology, it poses a number of security and privacy challenges. Information and communication technology has long aided transportation productivity and safety in advanced economies. These implementations, on the other hand, have tended to be high-cost, customized infrastructure systems. To address these challenges, a novel machine learning method developed for a transportation system is reused for making it more generic and smart for intelligent carriage. This type of transfer learning enables rapid progress on the task with enhanced results. In this work, together with domain adaptation, a novel weighted average approach is used to build models related to the smart transportation system. A smart system comprising of interconnected sensors along with the gateway devices can lead the way to a more efficient, viable and robust city centers. Finally, in this paper also provides a view of current research in smart transportation system along with future directions.

Keywords: Smart transportation · Transfer learning · Spatial and temporal characteristics · Connected world · Machine learning models · Homogeneous and heterogeneous transfer learning

1 Introduction

An intelligent transportation system refers to advanced application that helps to provide novel services related to various modes of transport along with traffic management, allowing consumers to be better educated and making the use of transport networks safer, more coordinated and smarter. In the recent years, many deep learning models are developed and have become an integral part of realizing the intelligent transportation [1]. This includes transportation traffic, the complex interactions and environmental elements and so on as shown in Fig. 1 below. Transfer learning helps in this smart transportation system through model features that are learned in diverse tasks by making them general:

A. K. Luhach et al. (Eds.): ICAICR 2021, CCIS 1575, pp. 162–181, 2022.
https://doi.org/10.1007/978-3-031-09469-9_15

Intelligent transportation system (ITS) includes a variety of services like traveller information systems, road traffic management, public transit system management and autonomous vehicles [2].

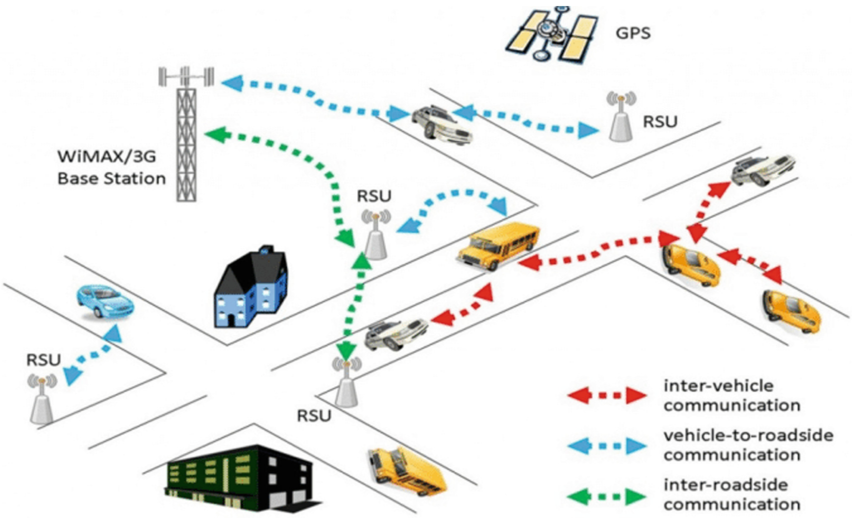

Fig. 1. A typical intelligent transportation system scenario

While smart transportation applications have been empowered by unprecedented progresses in computing, sensing and wireless technology, they will also have a wide variety of problems due to their scalability and varied quality-of-service requirements, as well as the enormous volumes of data they will generate. During the recent times, Machine Learning (ML) methods have gained significant attention in this field enabled by different technologies, including cloud- and edge computing [16]. ML has been used by a varied applications set, similar to ITS services, having a wide range of requirements.

In specific, ML models such as reinforcement and deep learning have been beneficial tools to explore underlying structures and configurations in big datasets for forecast and precise decision-making [3]. The different benefits of smart transportation system and ML approach to it are detailed below.

A. Communication

Intelligent transportation system helps create interconnected transport systems with the help of open communication that happens between different devices and vehicles.

B. Vehicle management

Managing traffic systems that in turn helps to keep the public transportation on time without any delay.

C. Real time information

Citizens have access to different real time information about the vehicular movement, traffic and other public transportation conditions.

Different sensors are used for collecting the data from the vehicles and are processed using machine learning algorithms for prediction and feedback purposes. Supervised learning methods works by inferring a regression or classification from a labeled database [4, 17]. Unsupervised learning approach on the other hand infer data without using any labels. Reinforcement Learning works towards learning to make a sequence of actions for maximizing the rewards in a given environment. Deep learning models uses artificial neural networks which consists of interconnected nodes offering non- linearity, high flexibility and data-driven model building.

Data is one of the important commodity that is extracted from the smart systems. ML try to further discover knowledge from this data. Classification, Regression, clustering, prediction and decision-making are the different features provisioned by ML that is capable of enhancing ITS and being foundations for the ITS application's. Data preprocessing, feature extraction and modeling are the main stages in the ML pipeline. Smart transportation system consists of four main components. It starts with the traffic data collection which uses devices like road cameras, GPS devices and vehicle identifiers for gathering the data in real time. The collected information provides details on speed of the vehicle, location of the vehicle and traffic conditions. Data transmission is the second stage in the pipeline which helps to transmit the data collected by the sensors to the network or the processing center where it is further treated and forwarded to applications. Data is further analyzed and the feedback is provided for the end users.

The different internet of things use cases related to this work include connected cards, vehicle tracking system, public transport management and traffic management. All of them involves usage of machine learning algorithms either at the source place where the data is collected or in the cloud. Since each use case involves multiple modules and most of them can be shared across applications with minor changes, hence use a new transfer learning approach for smart transportation system is proposed.

Transfer Learning (TL) is an artificial intelligence approach to the problem of learning where a prototype is reclaimed as the initial point for the new task that is already created [5, 18]. Developing a prototypical approach and a pre-trained model are the two common methods used for TL. In first approach need of selecting the source task is important to develop the source model, reuse the model and tune it as per the application. In the latter approach, select the source model and reuse the model and tune the model. This is very commonly used in the deep learning field.

The rest of the paper is organized as follows: Sect. 2 discusses the methods relevant to this work in the literature, Sect. 3 describes the data-based modelling of the intelligent transportation system, Sect. 4 deals with proposed transfer learning architecture, Sect. 5 shows the experimental outcomes and as a final point Sect. 6 concludes the work with further scope of study related to this topic.

2 Related Work

Emerging technology and the processing of big data have enabled the collection, analysis, storage and processing of multi-source data by systems. Cars, pedestrians, and even utilities will collect and exchange information in this area using a peer-to-peer procedure or a telecommunication network. Vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), pedestrian-to-infrastructure (P2I) or vehicle-to-pedestrian (V2P) interactions or knowledge transfer are all possible in this model. The authors [6] have reviewed the present trends in smart transportation applications along with insights in to connected vehicle environment for these systems.

When using the double loop detectors now used in traffic control centers, traffic practitioners prefer to calculate the time mean value rather than the space mean speed. The authors [7] feel that the relationship between the two speeds is important in smart transportation systems and hence have developed a probabilistic technique for appraising the space mean value from the time mean speed. They have also experimented the idea near Barcelona in real time and able to prove that the methodology can estimate with comparative fault as low as 0.5% through the proposed model.

With deep neural networks, TL begins with the process of preparing the base system on a given data set and then moving the structures to a target data set on the second network. The generalization capabilities of the deep neural network get improved with this approach. They are also useful in the time series classification problems. The researchers [8] investigated on how to handover the deep neural networks for the time series task. UCR archive that is publicly available and containing 85 datasets is used as the TSC benchmark for evaluating the potential of the transfer learning in this work. They have pre-trained a model for every dataset that is present in the archive, which is later fine-tuned on other datasets to construct different deep neural networks. At the end of the experimentation, the authors could prove that the TL can improve the predictions of the prototype depending on the transfer dataset used.

In various real-world applications, data mining along with machine learning techniques are used. Most approaches to machine learning demand that data from training and testing come from the same domain, which makes the space of the input function and the characteristics of the distribution of data the same. This assumption does not hold well when the training data is expensive to collect or unavailable. Hence high performance learners that can get trained with the data collected from different domains is needed. The authors [9] have surveyed different transfer learning process available in the literature along with their applications. The survey is made independent of the size of the data and can be useful for big data processing.

The majority of previous heterogeneous models of TL methods investigate a cross-domain feature plotting based on a small cross-domain instance-correspondence between different feature spaces, with these instances assumed to be characteristic in the source and target domains. The bias issues in this approach makes the assumptions to not hold good. As a result, the researchers [10] developed a new transfer learning method called Hybrid Heterogeneous Transfer Learning (HHTL), which allows for bias in either the source or target field in the resultant events across domains. The relationship between the source and the target has a significant impact on the effectiveness of transfer learning. The source side brute force leveraging will decrease the performance of the classifier. The authors [11] have hence devised an approach that makes use of extending the boosting framework for knowledge transfer from different sources rather than just one. Task-TrAdaBoost and MultiSource-TrAdaBoost are the two different algorithms discussed in this work which on experimentation proved that the performance is greatly improved by reducing the negative transfer. This is a fast algorithm enabling rapid retraining over new targets.

Utpal Paul et al. have conducted a measurement analysis with the help of large-scale data set that is collected through 3G cellular data network [22]. Individual subscriber behaviour is analysed and a significant variation in network is studied by the authors. The different implications with respect to protocol design, pricing parameters, resource and spectrum management are described in detail in this work. Different mobile networking networks have been developed in recent times in such a way that they have highly complex infrastructure and advanced range of related devices along with resources, as well as more diverse network formations, due to the firm development of current industries focused on mobile and Internet technology. As a result, Xiaofei Wang and others have discussed artificial intelligence-based methods for developing heterogeneous networks [23], as well as the current state of the art, prospects, and challenges.

The wireless communication techniques advancements along with mobile cloud computing, intelligent terminal technology and automotive domains are driving the evolution of automobile networks into the Internet of Vehicles paradigm. The vehicle routing problem hence gets changed based on the static data that is towards the real time traffic prediction. In this research, the authors Jiafu Wan et al. first address the classification of cloud-assisted IoV from the perspective of the service association between IoV and cloud computing [24]. After that, they assess traditional traffic prediction, which is used in both V2V and V2I communications.

The numerous traffic-related accidents that occur on expressways in a developed world are calculated to be closely related to previous traffic conditions, which are actually time-varying. To predict the likelihood of crashes, volume, speed, and occupancy-related parameters are used. These parameters are invalid for roads where traffic conditions are estimated using speed data. A dynamic Bayesian network model is designed by Jie Sun and others which models the time sequence traffic data and they have also inves tigated the relationship between dynamic speed condition data and the crash occurrence itself [25]. The authors have collected and used 551 different crashes data along with their corresponding speed related information from the expressways present in Shanghai, China. They developed the DBN models using time series speed condition data as well as different state combinations. The experimentation results from the authors show that the proposed DBN model offers a prediction exactness of 76.4% with a failure rate of 23.7% with only speed condition related data along with the nine traffic state blends. The results of the transferability verification show that the DBN models discussed are suitable for other related expressways, with a crash prediction accuracy of 67.0%.

With the ever-increasing global number of road traffic accidents, street traffic safety has become a serious issue for smart transportation systems. The identification of high-probability locations where major traffic incidents occur, so that precautions can be implemented efficiently, is a crucial step toward improving road traffic safety. The major limitations in this solutioning includes location accuracy and data availability.

To address these issues, researchers Lanyu Shang, Yang Zhang, Daniel Zhang, Yiwen Lu and Dong Wang created RiskSens, a multi-view learning method that uses social data and remote sensing data to identify dangerous traffic locations [26]. This is shown in Fig. 2 above. The authors' experimentation findings show that the RiskSens approach proposed in this paper significantly outperforms other state-of-the-art baselines for identifying different risky traffic locations in a region.

In developing countries, a lack of reliable data is a significant impediment to sustainable growth, disaster relief, and food security. Poverty data, for example, is typically scarce, labor-intensive to obtain, and has limited scope. Remote sensing data is becoming gradually available and low-priced as well. However, such data is highly unstructured, and there are currently no tools for extracting valuable insights to inform policy decisions and guide charitable efforts. A TL method [27] has been discussed by Marshall Burke, Neal Jean, David Lobell, Michael Xie and Stefano Ermon, where night-time light power is used as a data-rich substitute. The authors trained a completely convolutional neural network (CNN) model to predict night time lights using daytime imagery while studying characteristics that can be used to predict poverty. The researchers show that these learned features are extremely useful for poverty plotting, even precluding the prognostic presentation of field-collected survey data.

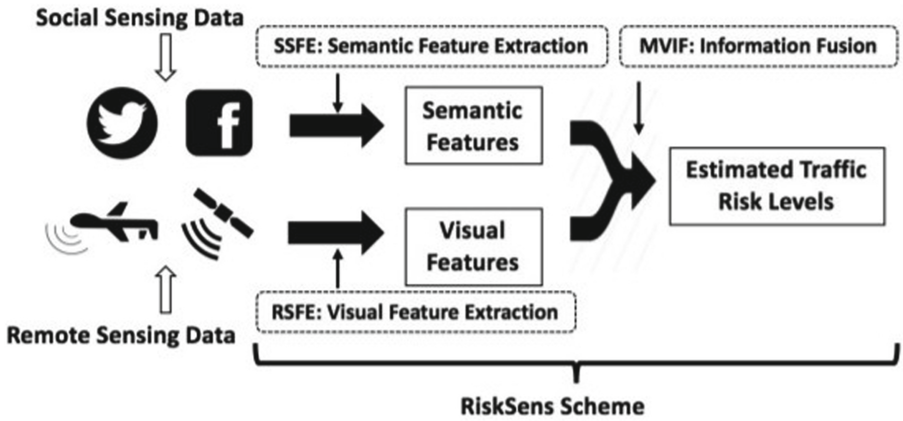

Fig. 2. RiskSens scheme

3 Data Based Modeling for Smart Transportation System

For implementing an intelligent transportation application, five different data processing stages have to be considered. Figure 3 below shows the different functional requirements for a smart transportation system. It is a list of possibilities that can help to make a model actionable and it includes usability, self-sustainability, application context, traffic theory and transferability.

1) Data collection through devices and sensors

Together with state-of-the-art and sophisticated microchip, RFID (Radio Frequency Identification) and low-cost smart beacon sensing technologies [12, 19], technical advances in information technology and telecommunications have strengthened procedural capabilities that will simplify motor safety assistance for smart transportation systems worldwide. Such sensing systems for smart transportation include infrastructure and vehicle based networked systems. These infrastructure sensors are long-lasting instruments that are fixed or installed in the driving road or surrounding path as required. They can be manually distributed by sensor injection machinery for fast positioning or by preventive road erection maintenance. Vehicle-sensing strategies include the placement of electronic communications beacons for vehicle-to-infrastructure and infrastructure-to-vehicle and can also deploy video-based automatic number plate segmentation or vehicle magnetic signature recognition technologies at preferred intervals to improve continuous monitoring of vehicles operating in sensitive world zones. The data collected through such sensing systems needs to be pre-processed and stored before further analysis.

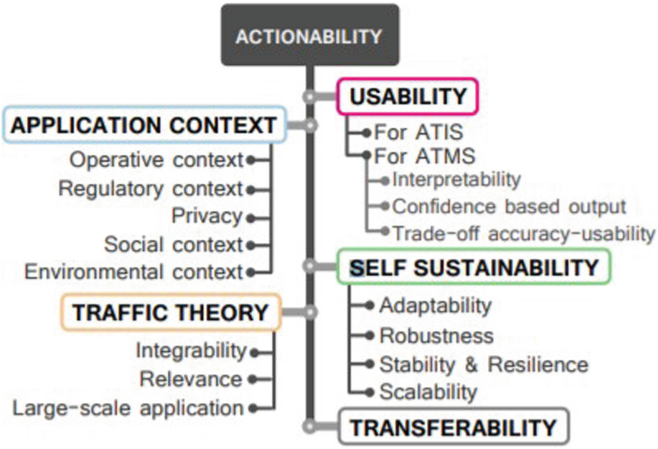

Fig. 3. Functional requirements for a smart transportation system

2) Data Pre-processing

The data comes from the sensors and devices in different forms and formats which needs pre-processing before modelling [13, 20]. Also, the corrupted instances present in the captured data can distort the databased model outputs and hence an actionable data pre-processing must emphasis not only on refining the captured data quality in terms of regularity and completeness, but also on giving valuable insights about the essential phenomena yielding corrupted, missing and/or outlying data, along with their effects on modelling. The study of possible pre-processing stages of data using experiential knowledge of domain features can be used as a method to increase the performance of the target learner, in addition to resolving inconsistencies in the domain varying step for dissemination. The heuristic knowledge will be characterized by a set of complex rules or connections that traditional transfer learning methods cannot explain. In certain instances, this heuristic knowledge will be exhaustive for each domain, resulting in no standard solution. If such a pre-processing phase, however, may lead to better target learner performance, then the effort is probably worth it.

3) Modelling

Modelling phase works towards extracting knowledge from the pre-processed data by constructing a model [14] that characterizes the distribution of the data. If this modelling can be generalized through transfer learning, then it can be used across domain rather than limiting to a particular domain. Machine Learning algorithms are hence put in to use in this stage which allow for modelling automation for instance, to understand patterns relating to the input data to a set of supervised outputs directing to automatically label unseen new data, to forecast future values based on the previous inputs, or to examine the output formed by a model when processing the data that is provided as input. In our use case, where the goal is to model data communications within complex systems such as transportation grids, the modelling choice routes to groups of diverse learner types.

A key feature of this modelling is the generalization of the established model to new unforeseen data. So the design goal would be to find the trade-off between the generalization and the model performance. For smart transportation, the accuracy in prediction can be improved by analysing different models and picking the right architecture or by combining different models for the new system. Optimization needs to be taken care when different models are combined together thereby increasing the system computational complexity.

4) Adaptation

A real time ITS environment is provided for the trained model to see how it adapts and behaves. Actions taken from the outcome result of a data-based model can aid for tactical, strategical or in operational decision making [15]. The output of the prior data-based modelling stage can be used to quantitatively evaluate the fitness or quality of the system. Finally, the suggested actionable data handling workflow contemplates model adaptation as a new processing level that can be applied over various modelling stages in the pipeline. Adaptations can be observed from two standpoints: automatic versions that the system is prepared to do when some circumstances happen, or the adaptations that are derived due to user changes.

Different models for following the vehicle patterns in the same lane can be studies through Gazis-Herman-Rothery (GHR) model. In this model, a vehicle fast-tracks in reaction to the velocity and front vehicle distance. The GHR model in its most general form is represented as:

$$a_n(t+T) = \alpha \frac{v_n^m(t)}{(X_{n-1}(t) - X_n(t))^I} (V_{n-1}(t) - V_n(t)) + K_i a_{n-1}(t) + K_2 a_n(t) \qquad (1)$$

where $a_n(t + T)$ refers to the speeding up of next car at any given time denoted by t + T,

T = Actual Time

V_n = Following car velocity

V_{n-1} = Leading car velocity

X_n = Follower car longitudinal position

X_{n-1} = Leading car longitudinal position

A = Acceleration at any given time t for both cars

Alpha, m, l, k1 and k2 are the other related parameters.

This model provides a general stimulus response and it is observed that the follower's acceleration is relational to the comparative speed between the trailing vehicle and the leading vehicle, whereas the actual data analysis is again inversely proportional. This model is derived from different basic prototypes such as G-H-P model, C-H-M system and basic GHR model.

The purpose of the calibration technique is to reduce the discrepancies between the driving behaviour simulated and the driving behaviour measured. This measured difference is referred to as the relative error, which is defined in Eq. 1 below, with variable 'a' designated as the error dimension.

When comparing the calibrated variables, it is important to evaluate the compassion of various constraints of the objective function, which is how much the value of the objective deviates when adding a slight adjustment to the original parameter. Data-driven modelling along with simulation technique is a noteworthy research focus. Compared to traditional knowledge-based modelling and mechanism modelling approaches, it demonstrates numerous advantages in operability and accuracy. Nevertheless, when implementing such modelling methods in practice, it is still doubtful since data-driven display is occasionally bad in description and data noises can also cause additional errors during modelling.

4 Proposed Transfer Learning Architecture

The case of smart transport domain usage that is considered defined by the letter 'D' and has two parts: a function vector space denoted by X and a marginal distribution of probability of the same by P (X). Here X = {x1,…,xn} X. If the machine learning algorithm can classify defects and each software metric in the algorithm is considered a function, then xi is the i-th feature instance conforming to the i-th module in the software, n is the total number of feature vectors present in X, Y is the space of all likely feature vectors, and X is a particular learning sample in this environment.

Once the source area is trained for a certain application, now that can extend the same for a different transportation related application through transfer learning instead of training it from scratch. This new transport domain is represented by DS with the resulting source task TS, along with the target area DT with the analogous TT task. TL refers to the procedure of refining the target projecting function fT() with the help of related information from TS and DS, where TS = TT or DS = DT. This symmetric transformation mapping is represented in Fig. 4 below.

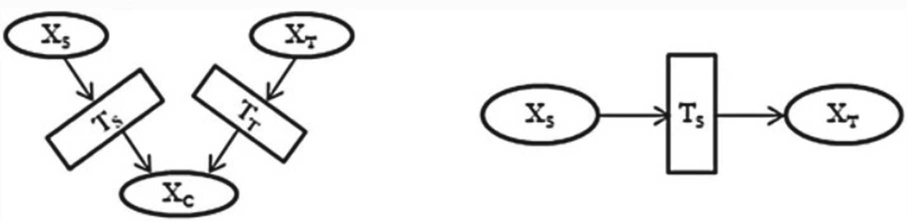

Fig. 4. Symmetric transformation mapping between source and target domains

If the distributions of the transport area are not found to be identical, then additional steps for domain revision are required. The type of data transfer is another essential feature of this transfer learning technique. This include four categories namely transfer learning through instances, learning through features, learning through shared parameters and finally transfer information created on some distinct association among the basis and target areas. This is presented in Fig. 5 below:

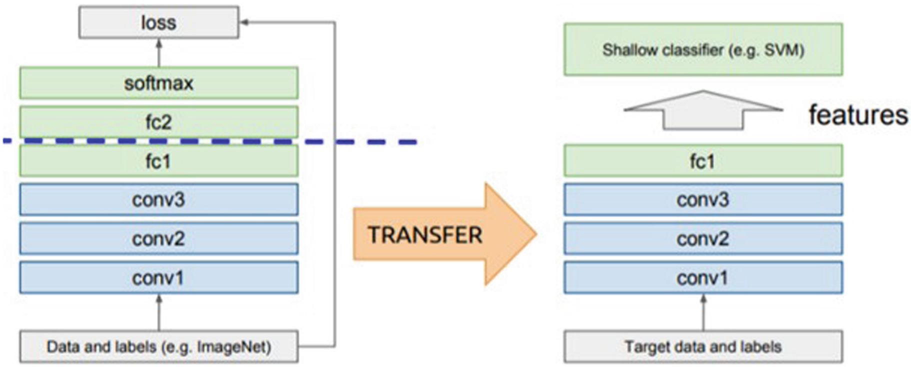

Fig. 5. System architecture for transfer learning

The proposed architecture of transfer learning addresses the following: What kind of data should be transferred between tasks, when to transfer this data and how to transfer it effectively without increasing the complexity of the system. The proposed algorithm takes advantage of the source domain's inductive biases to improve the target task effectively. Depending on the application, the learning could either be self-teaching or multitasking. It is suggested the use of four different forms of transfer, including the transfer of instances, the transfer of feature representations, the transfer of parameters and the transfer of relational information. The algorithm infers a mapping from the trained examples set. The inductive bias or conventions can be classified based on a variety of variables, including the hypothesis space within which it confines the method and the search procedure within that space. Hypothesis biases have an effect on how the algorithm has learned from the model on a particular mission.

D. Homogeneous transfer learning

The homogeneous TL categories can be built on parameters, instances, asymmetric features, symmetric features, relational or hybrid (both instance and feature based). The source data will be labeled while the target data can be labeled, unlabeled or partially labeled. So, the requirement here is to bridge the gap among the basis and the target domains. Proposed Strategies for this problem is

(a) Correct the marginal distribution differences in such a way that (P (Xt) = P (Xs)).
(b) Correct the conditional distribution differences in such a way that (P (Yt|Xt) = P (Ys|Xs)).
(c) A hybrid method of correcting both the above mentioned distribution differences.

In this method, reweight the samples collected in the source domain that helps to precise the marginal distribution changes. It uses these reweighted occurrences for training in the target domain. The conditional distribution should remain same in both the domains for better performance. The weights are adjusted based on the statistical values such as the mean and variance of the data that is under test. Maximum Mean Discrepancy is a distance metric that is applied on the probability measures space which has found different applications in nonparametric testing and machine learning, as shown in the following Eq. (2).

$$Disat(P(x_s), P(X_T)) = \left| \frac{1}{n_s} \sum_{i=1}^{n_s} \varphi(xs_i) - \frac{1}{n_T} \sum_{j=1}^{n_r} \phi(xT_i) \right|_H \tag{2}$$

Based on the statistics and their contributions to the classification accuracy in the target domain, part of the source domain labeled data can therefore be reused in the target domain as well after re-weighting.

E. Heterogeneous transfer learning

Different feature spaces are used to describe the basis and target areas in the case of heterogeneous transfer learning. The characteristics are used for learning by reducing the difference in the latent space between the various distributions. Data label obtainability is one of the functions of the primary application. Heterogeneous transfer learning solutions aims towards bridging the gap between the different feature spaces and change the problem to a homogeneous transfer learning one where additional distribution that could be either conditional or marginal differences will need to be modified. Machine learning algorithms have already shown promising results in the transportation industry, where it has demonstrated better performance compared to the traditional solutions. Nevertheless, the transportation issues are still rich in relating and leveraging machine learning techniques and need more attention. The fundamental goals for these models are to decrease congestion, increase safety and reduce human errors, optimize energy performance, lessen unfavorable environmental influences, and progress the productivity along with surface transportation efficiency. The machine learning pipeline and the interaction between ITS and ML is shown in Fig. 6:

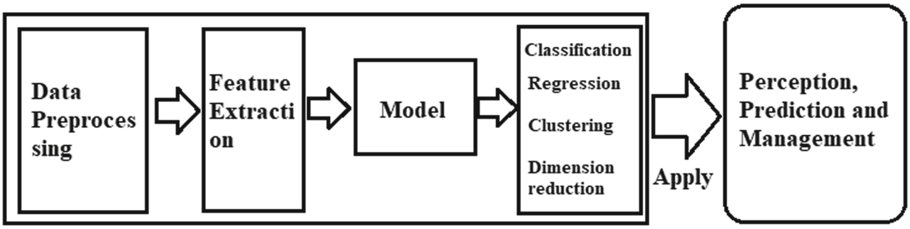

Fig. 6. ML pipeline and interaction with ITS system

The raw data could also be obtained, apart from the sensors, by triangulation process, vehicle re-identification, GPS-based methods and rich monitoring based on smartphones. Mobile operators across the world are also are becoming an important player in these value chains as they provide dedicated apps which can be used for making mobile payments, provide insights in to data and navigation tools, offer discounts and incentives to the end customer, and act as a digital e-commerce medium as well.

The proposed system has two models namely the convolutional base and the classifier. The convolutional base consists of a pooling stack and convolution layers. Development of image features is the main objective of this convolution base. Typically, the classifier used in this method is generated by completely linked layers. Using the detected features and classifying the image is the main objective of the classifier. A completely related layer is a layer whose neurons have a total effect on all previous activation of the layer. Once the system is trained with the model, it can re purpose a pre-trained model by removing the actual classifier and then introduce the new classifier that fits the ITS purpose and fine tune it through one of the following strategies:

(a) Train the full model: the pre-trained model's design is used and trained as per the ITS dataset in this case. The model is learned from scratch and thus needs a considerable amount of dataset.
(b) Train some layers while leaving the rest frozen: The general features are referred in the lower layers which are problem independent and the higher layers refer to precise features which are problem dependent. The weight of the network is adjusted and the frozen layer present in the model does not change during the training stage. If there are large number of parameters and the dataset is small, then there will be more frozen layers to avoid the problem of overfitting. On the other hand, if the dataset is huge with lowered parameters, then by training more layers to the new task, that can increase the model output as the issue of overfitting is not a problem.
(c) Freeze the complete convolutional base. This situation relates to the serious state of the freeze and train trade-off. In this method, the key concept is to maintain the convolutionary base as such and then re-use only the output as input for the classifier. Feature extraction happens through the pre-trained model and it is useful when the computational power is low, the provided dataset is small, or when the pre-trained model can solve multiple problems.

Since it is a hyper-parameter that is dependent on the weight change in the network, the learning rate associated with the convolutional component must be carefully chosen. In general, the high learning rate can make the system to lose the previous knowledge while the small learning rate is good to use. This will also make sure that the weights are not adjusted too often in the system.

The system model is presented in Fig. 7 where a set of vehicles on the road are equipped with different sensors to collect the raw data and transmitted through a mobile network to the edge computing device. Data accumulation provides the strength for analyses to capture some data insights that would not be conceivable from single sensors. The sensed data is transmitted frequently and the vehicles interacts with the gateway devices on every transmission. Based on the intended service, the gateway device will collect and handle the data accordingly. Transportation efficiency, vehicle security, travel safety, environment monitoring, are just few samples of types of services that can be offered. Before the feedback is provided to the vehicles or the end users, the machine learning algorithms can be run either on the edge computing system or on the cloud. The concept of transfer learning is also realized on the gateway device or on the cloud depending upon the application.

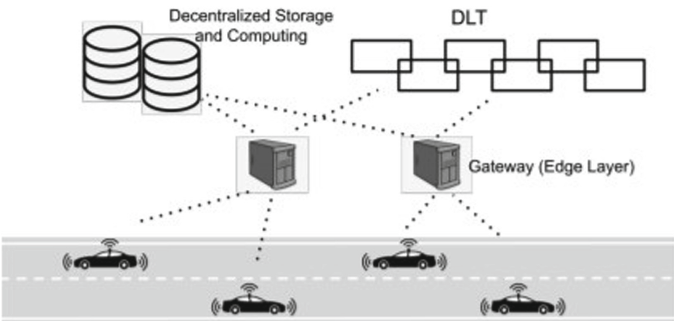

Fig. 7. System model for smart transportation system

For the purpose of smart transportation system, select the pre-trained model of convolutional neural network (CNN) which has a 4-layer architecture and then deep-belief network (DBN) model was employed to distinguish between the various associated activities. The dataset used for experimentation in this system is divided into three diverse groups such as speed limit overrun, immediate line overrun and yellow-line driving. This method can be used for different category features for different kinds of vehicles without training them independently.

Size similarity matrix is one which control the different choices in the system. Based on the size of the dataset, this matrix helps to classify the computer vision problem. This matrix is also useful for fine-tuning the model and repurposing the previously trained method. We have also performed weight transfusion based experiments, in which only a pre-trained weights subset of the system is transferred, with the remainder of them being just initialized randomly. While comparing the convergence speeds of these weight transfused models with full transfer learning, it is observed that the reuse of the extracted features are happening only at the lowest layers of the system.

The transfer learning scenario can be considered as a set of segments of the road “n” that are built with speed sensors. Each of the sensors “i” can provide the traffic speed at any given point of time “t” and is represented as vi[t]. Transfer learning is applied to this use case in order to predict the future speed of traffic at a given time. Historical data can be generally used for future value prediction but when such values are not available, then the concept of transfer learning helps.

The model proposed in this work will exploit this dataset for some source areas and then will build a prediction model for target location where there is only little or no available data. The data format consists of the road network represented through links and nodes. Each node in the network characterizes the latitude and longitude properties. Every link in the network will help to connect nodes and most streets consists of multiple links. The average traffic speed is given by the row values of a particular link at any given point of time. Different spatial and temporal features are extracted from the given data which will act as the essential constituents of the proposed approach. Once the features are extracted, different machine learning methods such as support vector machines, linear regression, convolutional neural networks etc. are used for training the system followed by testing on a new dataset not related to the training as such.

For smart transportation networks, various types of wireless communications technologies have been proposed. Radio mobile communication on VHF and UHF frequencies are extensively used for long and short range communication within ITS. This proposed intelligent transportation system based on transfer learning can be applied for various use cases, including Controlling traffic flow (traffic lights, measuring traffic flow, analysing traffic flow, and controlling guidance equipment), and managing public transportation (highway and tunnels, parking lots, expressway, railway and subway, bus, taxi and truck), Publication of Traffic Data (LED plate release information, SMS, radio station and television, terminal and website for public inquiry), Control of Traffic Offenses (over speed, red light running, wrong direction, occupied lane), Management of the Vehicle and Driver (vehicle information, driver information, driving route tracing, violation record and penalty), Statisticians and research (record demand, log, user management and simulation), daily tasks and emergency management (command centre, resource dispatch, pre-plan and daily task management) as well as real-time traffic status monitoring (accident, traffic jam and abnormal status).

5 Experimentation and Results Discussion

The availability of models developed for the source task and also tested is one of the significant criteria for the successful use of transfer learning. There are several advanced deep learning frameworks for TL and research purposes available across domains. Different pre-trained prototypes are typically shared in the parameters/weights form which is attained while being qualified to a stable state. For smart transportation system, the popular computer vision models include VGG-19, VGG-16, XCeption, Inception V3 and ResNet-50. One of the such model for training and testing is used in this method.

The identified Dataset is first divided into two categories — namely the training set and the testing set. Three different scenarios namely the no transfer task, cross transfer task and the local transfer task are considered for testing. The first one will predict the future speed based on the history of the data, second one will build a model first and then test on another target region that is not completely related while the last one will also have model training with the exception being testing from the same set.

Freight management, arterial and freeway management, transit management systems, regional multi modal and traveller information systems, incident and emergency management systems, and information management systems are all major areas of intelligent transportation systems in metropolitan deployments. Its applications are not limited to highway traffic alone, with electronic toll collection, highway data collection, traffic management systems, vehicle data collection, and transit signal priority being among them.

Open source library for machine learning purposes which includes grid based search tests and helps us to find the best performing model along with the appropriate hyper parameters. Our system has: 6 Hidden layers, 6 Neurons per layer, learning rate is 0.01and minimum error is 0.01.

Wireless communications, inductive loop detection, sensing technologies, bluetooth detection, computing technologies and video vehicle detection are some of the enabling technologies used in this research. During the initial stages, the larger networks needs more number of epochs to fit the data. Conversely after some number of epochs, it exceeds the smaller ones and achieves the best score.

When it comes to transfer learning, the proposed network is first pre-trained using simulation and then applied on the real data. When compared to the control network, these approaches work well. The convergence time is also much faster with this proposed approach. Some iterations are required for the random initialization network in order to fit the new weights in it. The complexity of the neural network, the predictive capabilities and the size of the gap between the simulation and real data will decide on the success of the different approaches discussed. An empirical study was conducted in South Indian districts, using three types of traffic datasets: floating cars, annual average daily traffic and public transportation routes.

Table 1. Comparison of transfer performance

Location	RMSE	MAE
Bellary	10.63	7.62
Parbhara	5.74	4.33
Hampi	4.91	3.04
Guntur	8.11	5.65
Jaina	11.32	8.92

There are two metrics used in this work to calculate the precision of continuous variables, namely the mean absolute error and the root-mean-square error. The first one deals with the error average magnitude in the prediction set without direction consideration.

$$MAE = \frac{1}{n}\sum_{j=1}^{n}\left|y_j - \underset{j}{\widehat{y}}\right| \tag{3}$$

The latter one called as the Root Mean Square Error is a quadratic scoring rule. It also helps to measure the average magnitude of the error. This is calculated as the square root of the average of prediction and actual values squared differences.

$$RSME = \frac{1}{n}\sum_{j=1}^{n}\left|\left(y_j - \underset{j}{\widehat{y}}\right)^2\right| \tag{4}$$

The average model prediction error is expressed through these two metrics MAE and the RMSE. They can range between 0 to infinity. They are both indifferent to the error direction. The lower the value, the better the output, and so on, as both scores are negative. RMSE will give more importance to large errors due to squaring of errors when compared to MAE. Table 1 gives the transfer performance across different locations in Karnataka, India in terms of MAE and RMSE. From this table, we observed that larger regions that has different kinds of links gave us better performance when compared to the other locations nearby. This is also represented through a graph in Fig. 8 below.

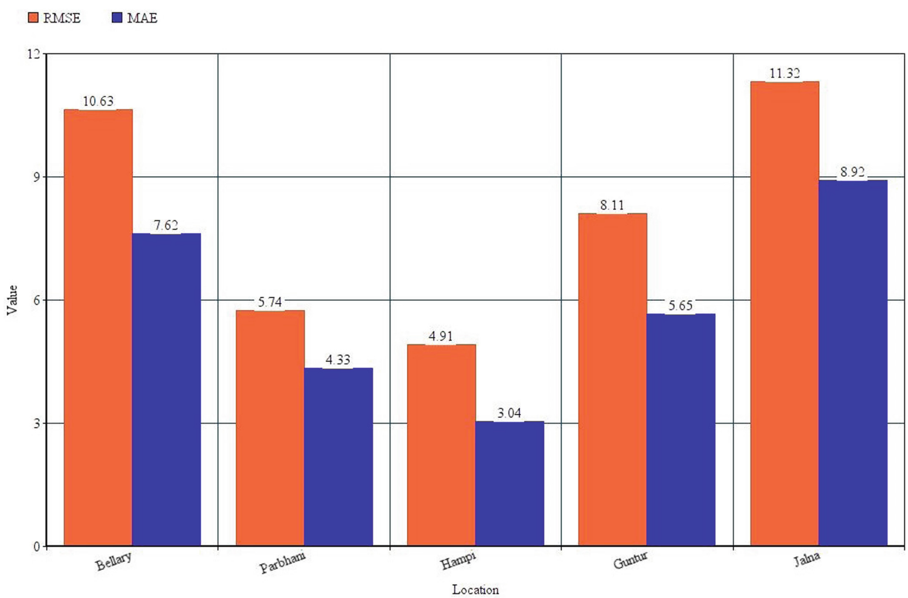

Fig. 8. Performance comparison across locations

Below shows the results and summarized statistics along with the error rates for the proposed method vs. other architecture for ITS in the literature that resorts to DLT features. The proposed transfer learning based smart transportation system outperforms in terms of performance when compared to simple pre-train CNN system that uses PCA in the same selected dataset (Table 2).

Table 2. Results for 60, 120 and 240 vehicles based on latency and error rates

Vehicles	Method	Average latency	Error rate
65	Fixed Random	72.68 s	14.36 %
	Dynamic Random	57.1 s	18.26%
	Proposed Method	20.40 s	0.70%
120	Fixed Random	86.85 s	24.48%
	Dynamic Random	67.5 s	18.99%
	Proposed Method	24.90 s	1.0%
240	Fixed Random	187.62 s	42.80%
	Dynamic Random	128.19 s	44.85%
	Proposed Method	71.37 s	6.45%

It is clear from the above table that both the error ate as well as the average latency is very less as compared to the existing approaches. While the average rate for the proposed method is ~ 1%, the other two approaches have well above 15% which is not acceptable in case of a smart transportation system and hence unusable. The empirical cumulative distribution function is shown in Fig. 9 below.

Similarly, the error rate across methods is shown in Fig. 10 below. It is clear that through an appropriate selection of full nodes, it is plausible to achieve consistent ledger updates or in other terms low errors, thus making feasible the use of IOTA to provision intelligent transportation system.

During our experimental assessment, all the full nodes had typically a low computational load. However, results indorse that the node selection is quite relevant. As an additional validation of this claim, in our initial tests we tried to feat a heuristic, substitute to those offered in the previous section. The idea here was to find the best N full nodes, in terms of existing resources, and use them to provide or validate the transactions. A gateway can also be used to employ an edge computing model as an alternative solution which would be an interesting future work.

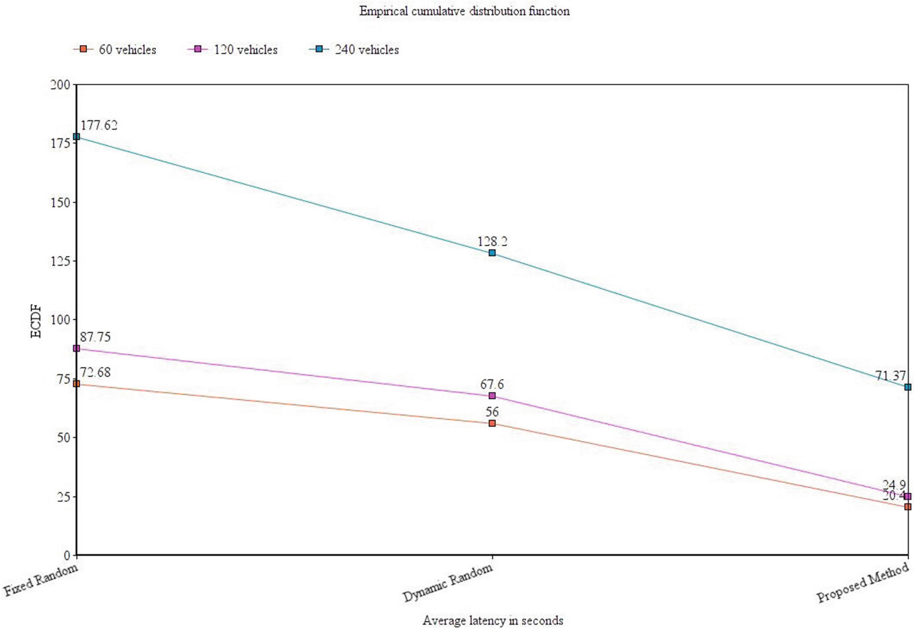

Fig. 9. Empirical cumulative distribution function

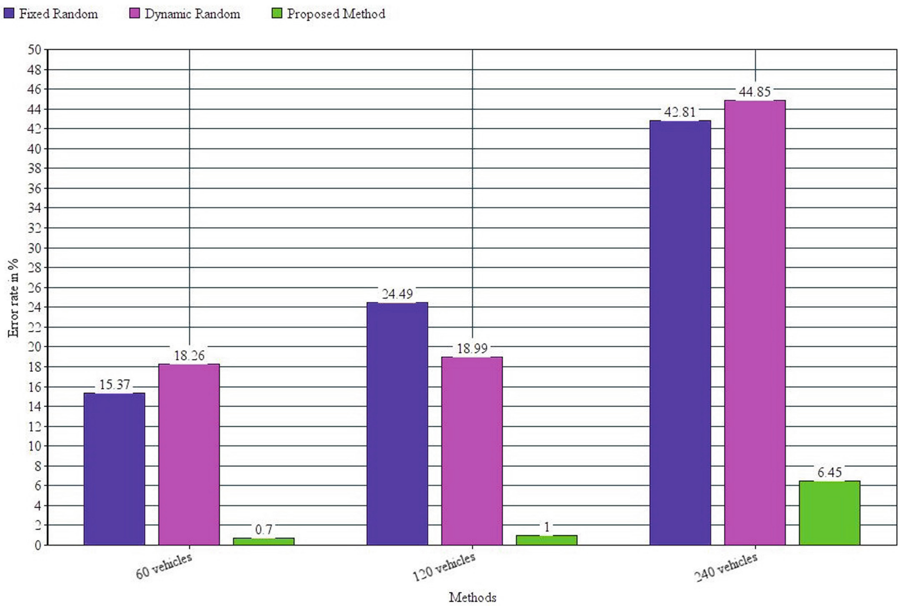

Fig. 10. Error rate analysis across methods

6 Conclusion and Future Directions

Intelligent transportation applications include warning systems for emergency vehicles, automated road compliance, variable speed limits and systems for crash avoidance. It involves the collection and processing of collected data for the purpose of providing information, control the actions of drivers, fleet operators, travellers and network managers. It provides a better understating of the transport network, providing new methods to manage the network and services to the public as well. ITS can be beneficial on its own or supporting other measures. It is not so easy to train the system for each of these applications related to transportation. Transfer learning helps simplify this task through pre-trained models used for other tasks. In this research, the proposed uses a new transfer learning architecture which is optimized for smart transportation system without compromising on the performance. ITS provides speed control devices that are not aimed at prosecutions like speed activated signs and displaying registration of speeding vehicle. In several transfers learning-based applications, the domain adaptation process focuses on either changing the conditional distribution differences or the marginal distribution differences between the source and target domains. Due to the lack of target data labels, modifying the conditional distribution differences is a difficult task. Drivers always wanted more information and more reliable journeys. We have addressed these expectations and issues in this work and moving forward, we would like to address the other issues associated with the marginal distribution differences as well.

References

1. Veres, M., Moussa, M.: Deep learning for intelligent transportation systems: a survey of emerging trends. IEEE Trans. Intell. Transp. Syst. **21**, 1–17 (2019)
2. Dabiri, S.: Application of deep learning in intelligent transportation systems. In: proceedings of Dabiri, Corpus ID: 86693644, February 2019
3. Haydari, A., Yasin, Y.: Deep reinforcement learning for intelligent transportation systems: a survey. IEEE Trans. Intell. Transp. Syst. 1–22 (2020)
4. Zantalis, F., Koulouras, G., Karabetsos, S., Kandris, D.: A review of machine learning and IoT in smart transportation. J. Future Internet, **11**(4), 94 (2019)
5. Krishnakumari, P., Perotti, A., Pinto, V., Cats, O., Lint, J.W.C.: Understanding network traffic states using transfer learning. In: 21st International Conference on Intelligent Transportation Systems, December 2018
6. Ho, H.W., Agachai, S.: Smarter and more connected: future intelligent transportation system. IATSS Res. **42**(2), 67-71 (2018)
7. Robuste, F., Soriguera, F.: Estimation of traffic stream space mean speed from time aggregations of double loop detector data. Transp. Res. C **19**(1), 115–129 (2011)
8. Forestier, G., Fawaz, H.I., Idoumghar, L., Weber, J., Muller, P.A.: Transfer learning for time series classification. In: IEEE International Conference on Big Data (Big Data), December 2018
9. Khoshgoftaar, T.M., Weiss, K., Wang, D.: A survey of transfer learning. J. Big Data, **3**(1), 9 (2016)
10. Tsang, I.W., Zhou, J.T., Yan, Y., Pan, S.: Hybrid heterogeneous transfer learning through deep learning. In: Proceedings of the National Conference on Artificial Intelligence, vol. 3, pp. 2213–20 (2014)

11. Doretto, G., Yao, Y.: Boosting for transfer learning with multiple sources. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1855–62 (2010)
12. Fong, S.L., Bakar, A.A.A.B.A., Ahmed, F.Y., Jamal, A.: Smart transportation system using RFID. In: 8th International Conference on Software and Computer Applications (2019)
13. Zelaya, C.V.G.: Towards explaining the effects of data preprocessing on machine learning. In: 35th International Conference on Data Engineering (ICDE), April 2019
14. Islam, M.S., Okita, T., Inoue, S.: Evaluation of transfer learning for human activity recognition among different datasets. In: IEEE Intl Conf on Dependable, Autonomic and Secure Computing, August 2019
15. Yu, C., Wang, J., Chen, Y., Huang, M.: Transfer learning with dynamic adversarial adaptation network. In: IEEE International Conference on Data Mining (ICDM), November 2019
16. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**, 1345–59 (2009)
17. Lai, T.M., Bui, T., Lipka, N., Li, S.: Supervised transfer learning for product information question answering. In: IEEE International Conference on Machine Learning and Applications (ICMLA), December 2018
18. Zhang, Y., Hongxiao, W., Daniel, Z., Dong, W.: DeepRisk: a deep transfer learning approach to migratable traffic risk estimation in intelligent transportation using social sensing. In: International Conference on Distributed Computing in Sensor Systems (DCOSS), May 2019
19. Shah Singh, M., Pondenkandath, V., Zhou, B., Lukowicz, P., Liwicki, M.: Transforming sensor data to the image domain for deep learning -an application to footstep detection. arXiv Computer Vision and Pattern Recognition, July 2017
20. Jelena, F., Maksims, F., Jörg, M.: Big data processing and mining for next generation intelligent transportation systems. Jurnal Teknologi **63**(3), 21–38 (2013)
21. Walker, S.J.: Big data: a revolution that will transform how we live work and think. In Int. J. Advertising **33**(1), 181–183 (2014)
22. Subramanian, A.P., Das, S.R., Paul, U., Buddhikot, M.M.: Understanding traffic dynamics in cellular data networks. In: Proceeding of IEEE INFOCOM, pp. 882–890, April 2011
23. Leung, V.C.M., Li, X., Wang, X.: Artificial intelligence-based techniques for emerging heterogeneous network: state of the arts opportunities and challenges. IEEE Access, **3**, 1379–1391 (2015)
24. Liu, J., Wan, J., Shao, Z., Imran, M., Zhou, K., Vasilakos, A.V.: Mobile crowd sensing for traffic prediction in internet of vehicles. Sensors **16**(1), 88 (2016)
25. Sun and Sun: A dynamic Bayesian network model for real-time crash prediction using traffic speed conditions data. Transp. Res. Part C: Emerg. Technol. **54**, 176–186 (2015)
26. Wang, D., Zhang, Y., Shang, L., Lu, Y., Zhang, D.: Risksens: a multiview learning approach to identifying risky traffic locations in intelligent transportation systems using social and remote sensing. In: IEEE International Conference on Big Data (Big Data), pp. 1544–1553 (2018)
27. Lobell, D., Xie, M., Burke, M., Ermon, S., Jean, S.: Transfer learning from deep features for remote sensing and poverty mapping. In: Thirtieth AAAI Conference on Artificial Intelligence (2016)
28. Bellone, M., Caltagirone, L., Wahde, M., Svensson, L.: Lidar-based driving path generation using fully convolutional neural networks. In: IEEE International Conference on Intelligent Transportation Systems (ITSC) (2017)
29. Liu, J., et al.: Resume Parsing based on Multi-label Classification using Neural Network model. In: 2021 6th International Conference on Big Data and Computing (2021)
30. Wang, X., Wang, L., Zhang, Y., Wang, J.: 3D-lidar based branch estimation and intersection location for autonomous vehicles. In: IEEE Intelligent Vehicles Symposium (IV), pp. 1440–1445, June 2017