Archives of Computational Methods in Engineering
https://doi.org/10.1007/s11831-021-09590-x

REVIEW ARTICLE

Big Data Processing and Analysis in Internet of Vehicles: Architecture, Taxonomy, and Open Research Challenges

Ansif Arooj[1] · Muhammad Shoaib Farooq[2] · Aftab Akram[1] · Razi Iqbal[3] · Ashutosh Sharma[4] · Gaurav Dhiman[5]

Received: 23 October 2020 / Accepted: 9 April 2021

Abstract

The extensive progression in the Internet of Vehicles (IoV) and the exponential upsurge in data consumption reflect the importance of big data in IoV. In general, big data has gained a significant attraction in academia and industry to provide valuable business intelligence and evidence-based decisions. This has been a key enabler for the advancement of the Internet of Vehicles (IoV) in which big data can be leveraged for efficient processing and valuable decisions. Moreover, data acquired from connected vehicles, traffic monitoring, social media feeds, and, crowd-sourcing can strengthen urban development and management. The purpose of this study is to synthesize a systematic review of all related research articles from January 2014 to September 2020 in well-alleged venues. We have rigorously surveyed the research papers to understand potential opportunities, methodologies, and challenges of using big data in IoV. This review shows that big data can play a key role in providing sound and valuable predictions and also provide a comprehensive analysis of several methods, tools, and techniques for the use of big data in IoV. Apart from reviewing the state-of-the-art studies of using big data in IoV, a taxonomy of the said also has been proposed. Furthermore, the article outlined and discussed several key challenges in IoV with notable recommendations and open research dimensions for using big data in IoV.

1 Introduction

Big data research has potential to apply evidence-based methodologies for decision making in several domains and the importance of using big data insights in Internet of Vehicles (IoV) processes for the security and sustainability cannot be denied. Most significantly, in smart environment several observational and monitoring equipment such as on-board and road sensors, roadside units, cameras, etc., are installed to observe patterns and conditions to promote safety and resource management. These electronic components gather massive amount of observational data. Managing and obtaining insights from this massive produced data are the key challenges of big data. But several computational and analytical approaches have enabled IoV to gain better understanding of traffic trends and telemetries [1–3].

Moreover, the advancement in information systems, sensing and communicating capabilities and, smart physical infrastructure creates new opportunities to diminish real-world problems such as traffic congestion, responsive and effective government processes for traffic monitoring, controlling, route management and urban planning etc., and many more. Real-world data streams acquire high-volume of heterogeneous data and, are pushing the traditional

✉ Muhammad Shoaib Farooq
shoaib.farooq@umt.edu.pk

Ansif Arooj
ansif.arooj@ue.edu.pk

Aftab Akram
aftab@ue.edu.pk

Razi Iqbal
razi.iqbal@ieee.org

Ashutosh Sharma
sharmaashutosh1326@gmail.com

Gaurav Dhiman
gdhiman0001@gmail.com

1 University of Education, Lahore, Lahore, Pakistan

2 University of Management and Technology, Lahore, Lahore, Pakistan

3 Department of Computer Information Systems, University of the Fraser Valley, Abbotsford, Canada

4 School of Electronics and Electrical Engineering, Lovely Professional University, Jalandhar, Punjab, India

5 Department of Computer Science, Government Bikram College of Commerce, Patiala, India

Published online: 07 May 2021

processing units to green field monitoring solutions with low-latency and lesser energy consumption based applications. The foremost requirement of processing and analytical engine is to guarantee the predictable outcomes with higher accuracy [4].

Researchers, Industries and organizations have been acquainted with the notion of big data even before the 1950s, decades earlier than the existence of the term “big data”; basic analytics were used to discover the hidden knowledge and trends using spreadsheets or manual interpretation. According to an estimate, 50% of the world’s population lives in urban areas that is expected to surge to 70% by 2050 [5]. Therefore, it is predicted that authorities will face numerous challenges from maintainability to sustainability and from safety to energy use and effective quality of services. Thus, evidence-based techniques for decision-making will be required to strengthen the idea to use big data analytics in various domains.

It is also noticeable that the idea of smart and connected cities (SCC) has also encouraged many researchers to come up with innovative concepts for better policies and standards [6]. Traffic observations and related data consist of several hidden contents that can be used to support and improve the sustainability and maintainability of components in IoV. For instance, traffic congestion can be identified by using roadside units and on-board vehicle speed sensory data. Once the congestion is observed by the data for a specific area, an alert message can be generated for driver assistance and can provide appropriate alternative routes, hence causing a reduction in traffic congestion [7, 8]. At traffic signals, a vehicle’s wait duration and status can provide insightful interpretation and can lead to the management of traffic light policies in order to improve the traffic movement and flow [9]. Trajectory data and, vehicular imagery data containing live simulation and observation have the ability to detect hindrance and classify infrastructure objects for controlling vehicular activities, for instance, braking, turning, speed and detection of collisions [3]. Thus, the application of big data processing and analysis in IoV has the potential for extraordinary real-time applications and solutions.

The influx of IoV data and technologies involved in IoV can fit every aspect of the big data characteristics, as 5Vs are the dimensions to handle the diversification of data [10]. When it comes to big data a large “volume” of continuous live streams is gathering and analyzing simultaneously, for the required insights and required rapid ingestion for instant response. Connected vehicles (V2V) and everything (V2X) generate and disseminate massive amounts of data. The data generation capability of a connected and smart vehicle is 30 gigabytes per day approximately [11] and will keep on increasing up to zeta bytes over time.

Secondly, a “variety” of data can be collected from IoV, for instance, audio, video, images, sensory observation, social media feeds, and GPS, etc., [1]. Moreover, “velocity” states the speed of data generation and dissemination. For exemplification let’s consider one situation, an accident at a corner of the road is captured and spread this information within edges using social media in seconds. This is also the ability of big data to analyze real-time data streams even before storage in databases. Some of the key abilities in IoV are response time, transaction processing time, data analysis and instant responses. And it can be managed by big data analysis [12]. “Veracity” is another feature of big data where the trustworthiness is the decisive goal of the big data world [10]. Because the massive volume and variety of data may contain misleading tags, observations, unique instances those don’t happen before, and typos, etc., that will lead towards the less reliability of data [1, 9–11]. Thus, the idea of providing fairness and truthfulness has significant importance in big data processing and analysis. Lastly, the “value” of IoV data is the ultimate goal of big data processing and analysis because the data without any benefits is just like a stegosaurus. So, it is necessary to understand the clear goals of big data with financial and rational benefits.

In big data domain, the most important ability is the instant and efficient processing of data, but with cloud computing, it is normally difficult to process such huge amounts of data to meet the required services offered by the domain. As the pressure of parallel tasks and responsibilities is already a burden on servers, thus the required proceeding and results cannot be achieved. However, edge computing is the solution to these problems where the processing power is distributed and a push generator (source) is used to process the data at the point of origins [16]. Considering the data generation by IoV components, in edge computing processing engine is responsible for data computation in the closed proximity of data sources e.g., sensors, cellular devices, roadside units, etc. and ability to reduce the bandwidth and latency of networks as compared to the cloud computing [16].

The prime intent of this review is to:

- Identify the big data characteristics and opportunities in the domain of IoV and to provide a consolidated understanding of big data processing and analytical techniques in IoV.
- Enable support for decision making in real world applications i.e., urban mobility, route monitoring and management, travel planning, road capacity management, traffic flow monitoring and management etc.
- Propose the framework to discover the most frequently used data processing and analytical techniques in the literature.

Although applications of big data in IoV have significant potential and also have many research challenges that need to be properly addressed. To best of our knowledge, a systematic literature review of big data processing and analytical techniques from data acquisition to management and storage to processing in IoV domain has not been done before. However, we have seen research in railway and transportation management using big data reviews been conducted. However, the applications of big data in IoV has not been considered and the authors of the previous researches have not also presented any taxonomy to help the researchers and industry [17–19]. In this review, we have initially explored the sources of big data acquisition and then the framework of big data IoV has been presented and discussed. We also have summarized the data processing and analysis tools, methods and platforms in IoV. Moreover, we have developed the taxonomy of using big data in the domain of IoV. This taxonomy will help the future researches in order to understand important aspects of using big data specifically in IoV and intelligent transportation.

The organization of this review is as follows: Sect. 2 provides the evolution of IoV from a standalone vehicle to IoV, Sect. 3 describes the research methodology of the study; Sect. 4 presents the results of the selection process of systematic mapping. In Sect. 5, the main findings of the systematic review are presented with the quality assessment criteria and summary of selected papers with evaluation results. These results have outlined the major contribution in the field. Finally, limitation, discussion, conclusion and future work present in Sects. 6, 7 and 8.

2 Evolution from Vehicles to Internet of Vehicle (IoV)

The history of automobile evolution has passed through many innovation stages. The development has progressed through the different alterations in size, style, features, and décor, while technological enhancement has pushed automobiles to adopt the latest trends. The timeline of IoV is illustrated in Fig. 1. Till the end of the 1990s, the automobile had started adding different safety parameters but this whole new area of innovation was initiated with the first “connected car” by General Motors (GM) in 1996 with features for emergency control and management [20]. Proceeding with this innovation Mercedes Benz launched key-less go and BMW assistance solutions for telemetries services in 1997. But the extraordinary feature of today’s IoVs, ‘navigation’ in cars was introduced in 2001 which helps people with stolen vehicle management by using smartphones or web connections [21].

Telemetric management and network access devices are also key features of this era. Network access devices have the ability to store and transmit vehicular data and have encouraged researchers and industries to identify more

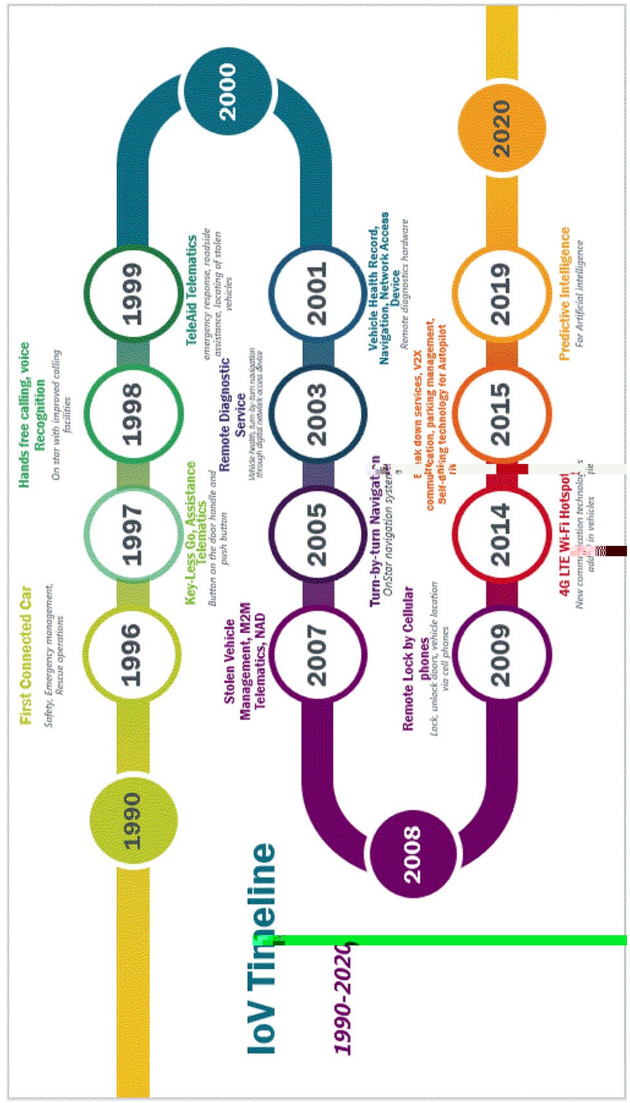


Fig. 1 Timeline of Internet of Vehicles (IoV) from 1990 to 2020

hidden capabilities of vehicles. In 2003, remote diagnostic service by GM was introduced which has the ability to observe and disseminate vehicular health data to the telematic service providers.

OnStar started turn-by-turn on-road navigation in 2005. Moreover, machine-to-machine (M2M) telematics opened the new paradigm for IoV, where the concept of V2V and V2X became a reality. These capabilities and enhancement in communication technologies also have strengthened the domain and the result has boasted in 2014 when Audi vehicles added 4 G LTE Wi-Fi hotspots, breakdown, rescue services, and parking management in vehicles. Moreover, V2X communication was presented in 2015 and the major contributor to this revolution was AA. Traffic light data was announced by Audi in 2017.

Lastly, predictive intelligence by Stratio automotive was the boost to create cognitive intelligence and automation [22]. It is expected that the global market for connected cars will be increased by 270% by the end of 2022 and approximately 125 million cars with extensive embedded connectivity will be available between 2018 and 2022. Moreover, fully autonomous cars will be commercially available from 2025 to 2045 [23]. Therefore, the idea of using big data applications in IoV and ITS will have a collaborative advantage in order to facilitate users.

3 Research Methodology

3.1 Systematic Literature Review Protocol

In this section, we define the protocol for a systematic literature review of the domain. Proceeding with the idea of [24], we have developed a protocol to identify, plan, synthesize and screen the available and published work to pursue the holistic related evidence for specific research questions. We describe the systematic review to discover the body of knowledge related to big data processing and analytics in the field of IoV. The objective of the review is designated to find answers to the questions by searching for digital libraries using literary search terms.

3.2 Research Objectives

The main objectives of this study are:

RO1 Identification of state of the art and focused research contribution of big data in IoV.
RO2 To prepare a year-wise summary of major research contributions from January 2014 to September 2020.
RO3 Consolidate the proposed architecture, tools, techniques, and methods for IoV using big data.
RO4 To propose a taxonomy of using big data in IoV.
RO5 To produce a comprehensive list for multidimensional and heterogeneous data types those gathered from IoV and what are the available tools for data acquisition?
RO6 To develop a comprehensive study for IoV predictive techniques for decision-making in the real world.
RO7 To identify the challenges and opportunities of the domain.

3.3 Research Questions

The proposed research questions for the present review are:

RQ1: What are the key research contributions, publications, timeline and, taxonomy of big data in IoV?
RQ2: What sort of data types have been acquired and managed in IoV for big data?
RQ3: What kind of big data processing and analysis system architecture, tools, and methodologies are used in IoV?
RQ4: How IoV is using predictive insights for real-world applications by big data?
RQ5: What are the substantial challenges and opportunities of big data in IoV?

Each of the above mentioned research question has a specific motivation: for example, the first research question is aimed at exploring the big data frameworks, architecture, tools, technologies and methodologies in IoV and ITS and, to develop a timeline of published papers and research contributions. The second question aims at exploring different types of data acquired from different sensory and observational units of IoV, to store and manage for the optimal decision-making. in the third research question, we intend to identify the system architecture in order to address the critical issues for big data processing and analysis. The fourth research explores the big data analytical results for optimal decision-making and predictions in the real world. And finally, the fifth research question aims to identify current challenges and future prospects of big data analysis and processing in IoV.

3.4 Search Strategy

There is a need for a well-organized research plan for extracting the relevant information and narrow down the retrieved results for focused answers and solutions. Planning is a major part of an effective search to get appropriate and meaningful results. For the effective exploration of the domain, we have used automatic and manual searching and exploration techniques. Initially, we have executed an automatic search technique by using search string over the search engines at digital and online repositories. Followed by the manual search method to collect more and more relevant literature for big data in IoV. The strategy for manual search techniques was the reference papers of primary retrieved studies. For better precision and recall, we have limited the search terms for expanding the coverage of search by conditions described as:

- Exploration of relevant keywords related to the research questions of the study.
- Identification of synonyms and alternative terms
- Formulation of search string by concatenating the key and alternative terms with Boolean operators as "AND", "OR".

3.5 Search Terms & Search String

Development of search terms requires deep observation of all related keywords that can precisely retrieve the relevant literature from the huge number of repositories. For ensuring the reliability of the term, we examined the primary concepts and lexicon in the domain of big data for IoV and identified the keywords to address the research questions. The meta-analysis includes the term *"Big Data Processing and analysis"* in the field of *"Internet of Vehicles"*. Table 1 describes the initial keywords for relevant research retrieval.

The resulting list describes (Table 1) the keywords and related terms and, it has been further discussed with our research group to eliminate the irrelevant terms. For forming a better search string with the Meta string, we have concatenated Boolean operators and wildcard special characters. The search string is sub-categorized into three parts. The first part of the search stream is fully focused on IoV, ITS or, connected vehicles. Secondly, big data processing and analysis is a sub-string of searching techniques and thirdly both strings are integrated for identifying the IoV applications in smart cities thus it is a part of the additional search string. Table 2 presents search strings to search from digital libraries.

3.6 Literature Resources

IEEE Xplore, ACM, Springer, Wiley and, Elsevier are primary search repositories for this paper. Moreover, journal and conference publications from other reputed journals are also part of our study. Some of the results were gathered from the Google search engine and rest was from Google Scholar due to in-depth content study and coverage of the topic.

3.7 Inclusion and Exclusion

Initially, the search criterion was just to extract the maximum number of publications in IoV domain. To ensure the most relevant results, we have included literature from the past six years, from January 2015 until September 2020 the term Big Data was focused on the emerging technology for the real-world application in 2012 [17]. This is a matter of fact that in 2015, big data started its emergence in many organizations, globally [18]. The repositories are searched by the literary resources as defined in Sect. 3.6. We have considered the following evaluation parameters to include the research papers for review.

- January 2014 until September 2020, related literature must be included.

Table 1 Keywords for the search string

Terms	Keywords, and synonyms
Data*	Data Analytics, Data analysis and processing, data processing, data science, data lakes, etc
Big*	Big Data, Big data Analysis, Big data analysis and processing, big data processing, big city, Big Data Applications etc
Internet*	Internet of Things, Internet of vehicle, Internet of everything, Internet of medical things, Internet of recording things, Internet of retail, etc
Smart*	Smart Vehicle, Smart phone, smart city, smart transportation, etc
Connected*	Connected vehicle, connected devices, connected networks. etc

Table 2 Search string strategy for database

Database	Search strategy
IEEE Xplore	((“Internet of Vehicles (IoV)” OR (“Intelligent Transportation System (ITS)” OR (“Connected Vehicles”)) AND (“Big Data” OR “Big Data Analysis” OR “Big Data Processing” OR “Big Data Processing and Analysis” OR “Smart Cities”)
ACM digital library	((Internet of Vehicles (IoV) OR (Intelligent Transportation System (ITS) OR (Connected Vehicles)) AND (“Big Data” OR “Big Data Analysis” OR “Big Data Processing” OR “Big Data Processing and Analysis” OR “Smart Cities”)
Science Direct	Title, abstract, keywords: ((Internet of Vehicles (IoV) OR (Intelligent Transportation System (ITS) OR (Connected Vehicles)) AND (“Big Data” OR “Big Data Analysis” OR “Big Data Processing” OR “Big Data Processing and Analysis” OR “Smart Cities”)
Taylor & Francis Online	[All: ((“Internet of Vehicles (IoV)” OR “(Intelligent Transportation System (ITS)” OR (“Connected Vehicles”)] AND [[All: “Big Data”] OR [All: “Big Data Analysis”] OR [All: “Big Data Processing”] OR [All: “Big Data Processing and Analysis”] OR [All: “Smart Cities”]]

- The research article must include credible sources in references and literary resources.
- The research paper must have full-length and published in reputed journal or conference.
- The research paper must satisfy the search string.
- Selected paper should be in English manuscript.
- Selected paper must have to satisfy at least one research question.

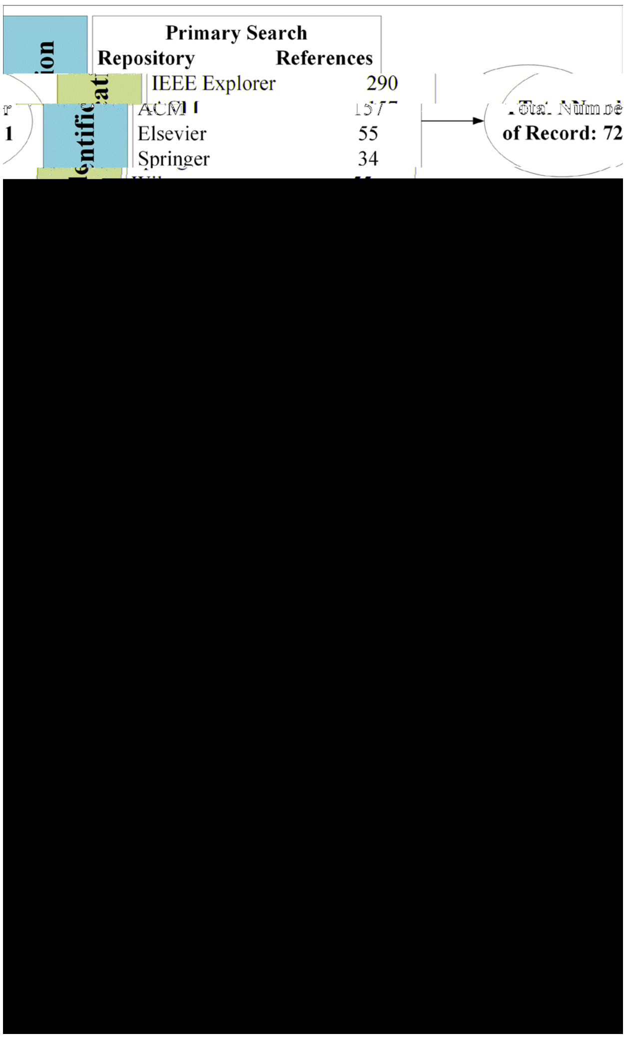


Fig. 2 Selection process of the study

4 Selection Process

Research strategy was proposed by [19], and have encouraged the process of relevancy for review with the research area. The selection process of this study is illustrated in Fig. 2 and Table 3 describes the stages of the study. Initially, the search string was applied at the identification phase of the study and after the primary search, a total of 721 papers have been retrieved. For better results, we have screened the papers based on secondary and tertiary search strategy; including the title and abstract based search of all retrieved papers.

After applying the inclusion strategy, we have excluded the papers from our studies which do not fulfil the criteria as described in section 3.7. In eligibility phase of the study, we have studied full-text of research papers for individual assessment and in-depth study. Then have excluded the studies which do not have any significant relevance with the domain of this study. Table 2 shows all related keywords that have been included in this review. Moreover, to satisfy our research questions we also have included big data processing and analysis, cloud and edge computing as well for exploring the available techniques. Thus, the resultant papers are 90 in total as enlist in Appendix as Table 12 along with the individual research contribution.

4.1 Step 1: Identification

The first step of this review was the identification of all primary resources, for this purpose the search string was passed through the Google search engine and Google scholar. The major resources were IEEE Xplore, ACM, Elsevier, Springer, and, Wiley etc. but there were other repositories, for instance, European Journal of Transport and Infrastructure Research, Hindawi, IET Intelligent Transport Systems and Taylor & Francis etc., have been also included in this study. The total number of retrieved papers were 721 and majority retrieved research papers were journal, conferences, and book chapters. Most significantly, we have included only those research papers having keywords of the search string. Table 7 describes the selected attributes of the retrieved results.

Table 3 Selection process

Stages	Phase	Description
S1	Identification	Primary search based on the search string from all repositories described in Table 2
S2	Screening	Subdivided into secondary and tertiary search based on title and abstract based search. After screening, we have excluded the papers that do not satisfy the inclusion criteria of the study
S3	Eligibility	Full-length papers have been reviewed and excluded the irrelevant studies after consensus and snowball tracking
S4	Inclusion	Selected papers included for the qualitative synthesis of this study

4.2 Step 2: Screening

This stage of review is sub-categorized in two search techniques; secondary and tertiary search. Secondary search was applied to the total retrieved results in the primary phase of the study, by reviewing the titles of all research papers. In this case, the duplicates, irrelevant and unauthentic resources have been identified and excluded from the primary results and passed to the next screening level as tertiary search. All abstracts of selected papers have been studied and further discussed with the peers of research team and classified the proposed research approaches of studies as; case studies, reviews, surveys, experiments and simulation etc. At this stage, without judging the quality of empirical studies total of 190 papers have been selected for the next phase of review.

4.3 Step 3: Eligibility

This phase addresses the empirical quality analysis of all selected papers and subdivided into full-text search and snowball tracking. Analysis has been executed on the full-length study of selected papers and quality assessment strategy is used to calculate the quality of selected papers, as described in Sect. 5.1. Moreover, the references of selected papers were also explored and reviewed by snowball tracking methodology. The papers were excluded with reasoning and consensus between authors. After closed review and assessing the eligibility criteria, 132 papers were selected for further processing.

4.4 Step 4: Inclusion

Lastly, quantitative and qualitative synthesis was applied to eligible research papers and 90 papers were finalized for the systematic review. The details of selected papers are enlisted in Appendix as Table 12 published during January 2014 to September 2019 and only those publications were included those have fulfilled the research criteria (Table 4). Included papers have been categorized with respect to ratio as IEEE Xplore 56.7%, ACM 6.7%, Elsevier 11.1%, Springer 12.2%, Wiley 1.1% and, miscellaneous 12.72%, respectively. All research papers were rigorously evaluated by qualitative and quantitative evaluation and assessment. Summary of the total selection process is represented in Table 3, with the detail of final selected papers is enlisted in Table 5. Most of the papers 56.7% are selected from IEEE Xplore and almost 56.7% of the publications were from finalized from journals. Detail of finalized papers is presented in Appendix as Table 12 with attributes as the channel of publications (books, journals, conferences), publication year, authors name, publisher, research type, domain, and title of the research. Research categories are classified as the following:

Table 4 Summary of quality assessment criteria

Sources	Ranking	Score
Journal	Q1	4
	Q2	3
	Q3 and Q4	2
	Not in JCR and web of science	0
Conference	CORE A	1.5
	CORE B	1
	CORE C	0.5
	Non-CORE	0
Book	Citation and indexed in web of science	2
	Indexed in web of science	1
	With citation	0.5
	No citation and ranking	0

- Description as DESC
- Simulation as SMU
- Case Study as CS
- Framework as FW
- Prototype as PRT

Moreover, the domain of the study is defined as the set of search string keywords as:

- Internet of Vehicles: IoV
- Big Data: BD
- Connected Vehicle: CV
- Intelligent Transportation System: ITS
- Smart Cities: SC

5 Data Analysis and Results

Quality assessment (QA) criteria of this study are carried out in the form of a questionnaire, inspired by [20]. The questionnaire was prepared to identify the quantitative and qualitative parameters of the study and filled by the authors of this study. Tabular descriptive evaluation is presented in Table 4, and we have drawn our conclusion and recommendations based on research results.

5.1 Quality Assessment Criteria

Cohen's Kappa Coefficient is used for finalizing the judgment of evaluation criteria between two authors describing the same concept of domain. The Kappa coefficient is selected as 0.95 as perfect agreement assessment criteria described by [29]. The evaluation criteria of papers have been mapped as:

- (QAC-1)- c to the field of big data analysis in the field of IoV. The possible responses were “Yes” and marked as (+ 1), partially as (+ 0.5) “No” indicated as (+ 0)
- (QAC-2)- The research paper presents the architectural solution for big data analysis and processing in IoV. The possible responses for the outcomes were “Yes”, “No” as (+ 1) and (+ 0) respectively.
- (QAC-3)-The study presents an empirical solution in the domain of IoV. The responses were “Yes” and “No” as (+ 1) and (+ 0).
- (QAC-4)- Lastly, the study discussed the current challenges and opportunities in the field of IoV in terms of big data processing and analysis. The possible answers were “Yes” (1), partially (+ 0.5), and No as (+ 0).

For journal full-length quality assessment and evaluation, the criteria as been marked as:

- Q1 ranked journal as (+ 4)
- Q2 ranked journal as (+ 3)
- Q3 & Q4 ranked journal as (+ 2)
- Not JCR and web of Science indexed (+ 0)

Moreover, the evaluation criteria for the full-length study was ranked as described by [22]:

- CORE A conference is marked as (+ 1.5)
- CORE B conference is ranked as (+ 1)
- CORE C conference is ranked as (+ 0.5)
- If not satisfy the CORE ranking is marked as (+ 0)

For book ranking, we used the number of citations and web of science indexes as here-under:

- Book with a citation is ranked as (+ 0.5)
- Book that is indexed is the web of science repository is ranked as (+ 1)
- Book with indexed in web of science and also have citation is ranked as (+ 2)
- Book with no citation and indexing is ranked as (+ 0)

The quality assessment parameters for review were questioned as:

- The review criteria as having been described in section Table no 5 have been met?
- Are all related literature search have covered all aspects of the study?

5.2 Search Results

We have finalized 90 research papers out of 721 papers at the inclusion phase of this study published from January 2014 to September 2020. These selected papers are retrieved from reputed journal websites and google scholar and, categorized as books, book chapters, journal papers and, conference papers. The publication results are presented in Table 6. 56.7% from IEEE Xplore, 6.7% from ACM, 11.1% from Elsevier, 12.2 % from Springer, 1.1% from Wiley and, 12.72% from miscellaneous research resources have added for review. Figure 3 illustrates the publication status and contribution. Figure 4 represents the year wise publication status and after observation of graphical representation of results, it is noted that 41% selected publications are published in the year 2015 including journal publications, conference proceedings and, book chapters.

Furthermore, in the inclusion phase of this review, all selected papers are rigorously being evaluated through quality assessment standards and it is also noted that 56.7% of selected papers are from journals, 37.8% are from conference proceedings and, 5.5% are from books. After determining the quality, assessment criteria QAC-1 to QAC-4 on selected papers for systematic review scores are presented in Table 9. The overall results of this systematic review are 68% with respect of domain-relevant research questions, which is relatively aligned and relevant to the research domain.

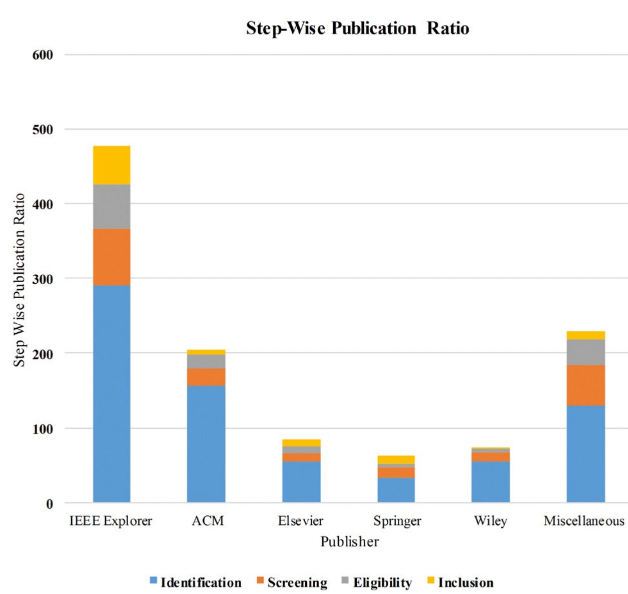


Fig. 3 Illustration of publication status with respect to process wise selection process from identification to inclusion

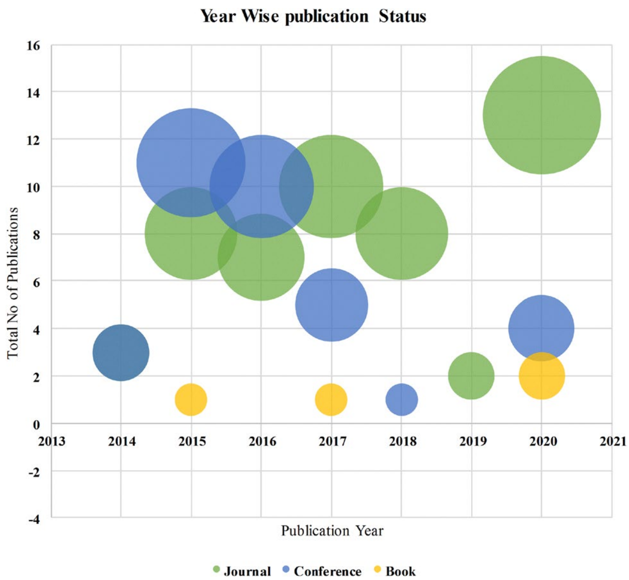


Fig. 4 Year wise publication Status of all selected papers

Table 5 Selected publication for the study

Sources	Total	Journal	Conferences	Book chapters
IEEE Xplore	51	27	24	0
ACM	6	1	5	0
Elsevier	10	9	1	0
Springer	11	3	4	4
Wiley	1	1	0	0
Misc.	11	10	0	1
Total	90	51	34	5
Ratio %		56.7	37.8	5.6

Table 6 Publication results

Sources	S1	S2	S3	S4	Ratio%
IEEE Xplore	290	76	60	51	56.7
ACM	157	23	18	6	6.7
Elsevier	55	11	9	10	11.1
Springer	34	13	5	11	12.2
Wiley	55	12	6	1	1.1
Misc.	130	55	34	11	12.72
Total	721	190	132	90	100

5.3 Assessment of Research Questions

5.3.1 Research Question 1: What are the Key Research Contributions, Publications, Timeline and, Taxonomy of Big Data in IoV?

Figures 3 and 4 shows the publication status and year-wise publication distribution of selected papers. Interestingly all the major publications are after 2014, this represents the initial generation of the big data applications and is the new paradigm of emerging technologies. 6 papers (6.7%) were published in 2014, 21 papers (23.3%) in 2015, 17 papers (18.9%) in 2016, 16 papers (17.8%) in 2017, 9 papers (10.0%) in 2018, 2 papers (2.2%) are published in 2019 and 19 papers (21.2%) are published in 2020 until September. Thus most of the selected papers have been published in 2015 (22.2%) and individual research contribution of all selected research papers along with research contribution is presented in Appendix as Table 12. Figure 4 shows the sources of publication from academia and industry. Our results represent that 56.7% were published in reputed journals, 37.8% in conferences and 5.6% are, as book chapters and sections. This is highly possible that some of the industrial published papers are not included in the scientific repositories. This is also being observed that the industrial papers were found as white papers, blogs and websites and, are not published in peer-reviewed repositories, therefore, they are not included in the selected papers. Moreover, IEEE Xplore has the major publication in domain with 56.7% and other journals contributions are ACM 6.7%, Elsevier 11.1%, Springer 12.2%, Wiley 1.1% and 12.72% from miscellaneous such methods of the IoV agents, classification, format, tools and related examples. The data sources are categorized as sensory observation, ITS, GIS and map-related sources, social media feeds of drivers or passengers, linked data, archive, and legacy data.

Vehicular observation is possible due to embedded sensory objects like accelerometer, gyroscope, airflow and pressure sensor, etc. These sensors help to monitor the status of processes happening in vehicles such as engine characteristics, brake status, velocity, and movement. On one hand, on-board sensors observe real-time data values and on the other hand, GPS and in-built imagery devices not only capture videos and pictures of vehicle surrounding but also share positioning and navigational data points of vehicles [1, 3, 7, 9, 35–38]. More Importantly, these data sets are valuable assets for government and agencies to manage traffic and other events. For instance, traffic congestion, any protest, and, the catastrophic situation can be monitor remotely. So data set of many vehicles form crowd-sourcing network [39–42]. Data collected from the sensory observation units are raw in nature and thus require pre-processing and integration in order to understand the actual meaning of observation. Text, CSV, images, videos, GIS-related data is the usual format of data in observational and ITS units.

Secondly, people on-board, bus or, passing by the surroundings of a vehicle such as a bus stop, train, and metro station are also sources of IoV data. For example, GPS data of people's mobile phones and social media feeds are continuously being shared and present the complex association

Table 7 Selected attributes for systematic literature review

Attribute	Sub-attribute	Description
General observation	Paper number, type of publication (journal/conference), title, publication year, institute, research domain, proposed research, name of authors	General overview of research papers in the domain of IoV
Research papers and proceedings	Paper number, title, type, publisher, year, research domain, focus of study	describes the list of complete information related to the selected research papers and proceeding
Assessment criteria	Paper number, problem area, proposed solution, research contribution, weaknesses and, future work	The respected papers are analyzed and assessed by reviewing the problem domain, solutions, research contribution and, weaknesses
Selected papers repository	Paper number, publisher, country, reference papers including journal, conference, workshops, book chapters	Provide complete information related to the selected papers
Status of papers	Paper number, status of selection (after meeting the selection criteria), Total papers	Describe the list and the total number of papers after the evaluation of the selection procedure
Selected research	Selected paper count, journal publication, conference proceedings, book chapters	Provide the number of total publications in the research domain

as linked data in RDF, XML or, SPARQL format. People on board and vehicular observation both can share information about the same event happening nearby. Thus, this idea has encouraged to apply data fusion techniques in different data sources and form Cyber-Physical and Social Networks (CPSN) for real-world applications [39, 43]. Thirdly, fixed sensors in smart cities, for instance, smart surveillance camera, smart meters, air pollution sensors, smart bins, irrigation sensors, Road Side Units (RSU), infrared detectors and, ultrasonic sensors are rich sources of data in IoV [44–46].

As defined previously in this section, data generated in IoV are gathering from heterogeneous sources in heterogeneous types, and form big data thus eligible for big data computation due to 5Vs properties. The prediction of 200 sensors on a vehicle by 2020 will also cause an increase in data volume. It is predicted that about 4000 GBs of heterogeneous data will be generated by smart vehicles every day [47]. It is also evident that traditional sensors, for instance, GPS and ECU, etc., generate less data than the modern sensors such as Lidar and Camera and thus encourage a large number of applications in the real world like autonomous vehicles [23, 48, 49].

Moreover, sensory data can be acquired from other vehicles as well (V2V), by using DSRC channels. This data may include alert messages or any updated information. Platoons of vehicles are the formation of vehicle channels nearby and at same geolocation can be used for the dissemination of control messages. Besides the use of data for V2V communication management, this data can also provide links between vehicles and smart infrastructure (V2X) by using WiFi specifically described as Intelligent Transportation System. V2X data can be used for traffic monitoring and navigational route management [1, 2, 50–52, 52–57]. Apart from vehicular, cellular and other sensory data acquisition from ground objects, IoV uses space and other aerial platforms such as drones, unmanned aerial vehicles (UAVs), satellites, and High Altitude Platforms (HAP) as data acquisition sources [46, 58]. The detail of sensory units is presented in Appendix (13)

Proposed Taxonomy In IoV, big data processing and analysis have been used to fortify the decision-making process in real-time applications. Integrating the state-of-the-art big data algorithms with IoV is the first step of getting meaning out of stegosaurus datasets. The review conducted on research studies for understanding the means of functionalities, processing, and capabilities of big data in IoV. The resultant equation of the process includes several key classification phases as data acquisition, data storage, data processing, and, data analysis.

These categorizations can help to understand the taxonomy of big data in IoV. The major concerns from physical structure to data transformation and, to normalization are dealt with the phase of data acquisition. From there the story of big data starts, these data points collecting from heterogeneous data sources and then transform and normalize depending upon the cross infrastructure and cross-process platforms followed by the data management and storage at the cloud or, forwarded to data process and analysis engines.

Data processing engines used to batch and real-time processing, tensor decomposition, semantic derivation and, data fusion to process and consolidate these multi-dimensional points to add cognitive understanding of the context. This understanding supports analyzing data through descriptive, predictive, and, prescriptive levels of analysis. Data integration, mining, neural networks, machine learning and, heuristic algorithms are used to develop short and long time prediction models in IoV. Urban planning, scheduling, resource optimization and, statistical and economical econometric are some of the key decision venues of IoV. The taxonomy of

Big data in IoV has been proposed in Fig. 5 to help IoV researchers in selection possibilities.

5.3.2 Research Question 2: What Sort of Data Types Have Been Acquired and Managed in IoV for Big Data?

In IoV, data is collected from heterogeneous devices and has inherent sensory observation capabilities. The major agents of data acquisition in IoV are categorized into three forms: i) vehicle, ii) people on board and, iii) fixed sensors in smart cities [31–34]. Table 8, represents the summary of data sources and acquisition methods of the IoV agents, classification, format, tools and related examples. The data sources are categorized as sensory observation, ITS, GIS and map-related sources, social media feeds of drivers or passengers, linked data, archive, and legacy data.

Vehicular observation is possible due to embedded sensory objects like accelerometer, gyroscope, airflow and pressure sensor, etc. These sensors help to monitor the status of processes happening in vehicles such as engine characteristics, brake status, velocity, and movement. On one hand, on-board sensors observe real-time data values and on the other hand, GPS and in-built imagery devices not only capture videos and pictures of vehicle surrounding but also share positioning and navigational data points of vehicles [1, 3, 7, 9, 35–38]. More Importantly, these data sets are valuable assets for government and agencies to manage traffic and other events. For instance, traffic

Fig. 5 Proposed taxonomy

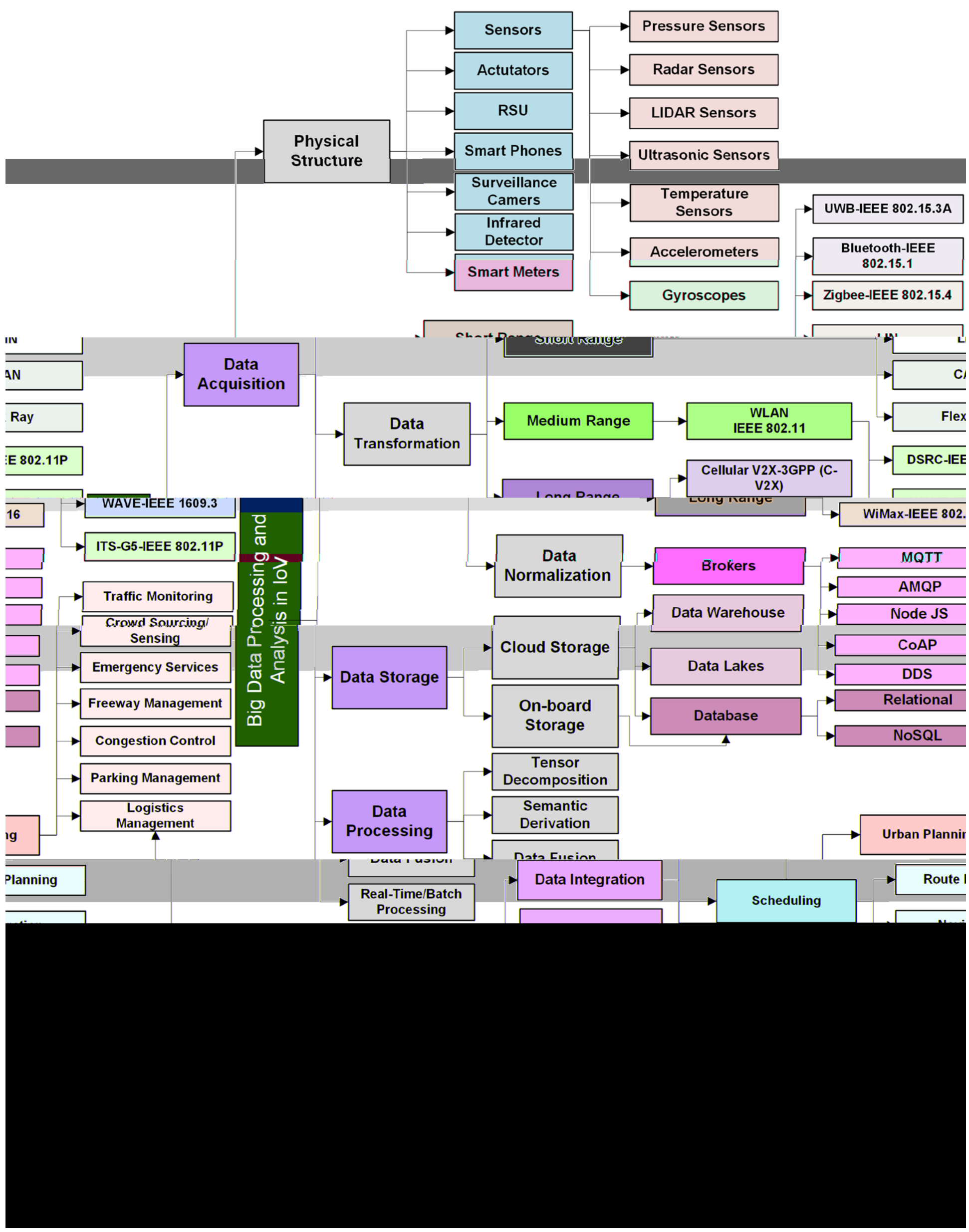


Table 8 Types of data for IoV

Data sources	Classification	Format	Storage tool	Example
Sensory observation	Raw data	Text, CSV, sensor formats, images, videos	RDBMS, MongoDB [59], NoSQL etc	Vehicular real time road side sensors, fixed sensors, move-able sensors, smart phone sensors, in-vehicle HD video streaming, magnetic loops, microwave radar, laser-based systems, infrared detectors, and ultrasonic detectors, vehicle speed, vehicular imagery and videos data sensing [1, 2, 6, 31–34, 37–42, 44, 50–57, 63, 75–79, 132, 137, 138]
ITS	Structured	CSV, XML etc	RDBMS, MySQL, PostgreSQL	Car navigation [34, 56, 80, 81] Power consumption [23, 37, 37, 76, 82, 83] Traffic signal control data [84], Traffic monitoring [50, 50, 53, 57, 85, 86, 137, 141, 145, 148]
GIS and map related data	Semi-structured	GIS	Esri open data hub, natural earth data, OpenStreetMap, NASA's socioeconomic data and applications center (SEDAC), Influx DB [33]	GPS coordinates, time, velocity and direction Process [34, 42, 50, 52, 54, 54, 56, 57, 75, 80, 87, 88, 136, 138]
Social media	Semi-structured or unstructured	Text, CSV, RDF, XML, JSON etc.	MongoDB [59], Couch DB, Cassandra [60]	crowd sourcing [3, 89] Social media analysis [45, 77, 90–92], sentimental analysis of drivers and passengers [33, 46, 93, 136]
Linked data	Semi-structured	RDF, XML, SPARQL	Neo4J, Cassandra [60], Hadoop HDFS [69]	Urban mobility [38, 53, 94] route monitoring and management, travel planning, road capacity management, traffic flow monitoring and management [38, 50, 54, 85, 95, 142]
Archive and legacy data	Structured	RDBMS, XML, JSON, SQL, ERM	RDBMS, MySQL, PostgreSQL	Toll taxes data, driver license information [33, 37, 46, 72, 96, 137]

Table 9 Quality assessment results

Sr.	Ref.	QAC 1	QAC 2	QAC 3	QAC 4	Rank	Rank score	Total score	QAC%
1	[32]	1	1	1	1	Indexed in web of science with 3 citation	2	6	75.00
2	[97]	1	1	1	0.5	3 citation	0.5	4	50.00
3	[53]	1	1	1	0.5	Not ranked	0	3.5	43.75
4	[98]	1	1	0	1	CORE A	1.5	4.5	56.25
5	[1]	0.5	1	1	0.5	Not ranked	0	3	37.50
6	[26]	1	0	0	0.5	Not ranked	0	1.5	18.75
7	[56]	1	1	1	1	Not ranked	0	4	50.00
8	[57]	1	1	0	0.5	CORE B	1	3.5	43.75
9	[75]	1	1	1	1	Not ranked	0	4	50.00
10	[91]	1	1	1	0.5	Not ranked	0	3.5	43.75
11	[99]	0	1	1	0	Not ranked	0	2	25.00
12	[94]	1	1	1	0.5	Not ranked	0	3.5	43.75
13	[34]	1	1	1	0.5	Not ranked	0	3.5	43.75
14	[52]	1	0	1	0	Not ranked	0	2	25.00
15	[50]	1	1	1	0.5	Not ranked	0	3.5	43.75
16	[51]	0.5	1	0	0	CORE B	1	2.5	31.25
17	[54]	1	1	1	0	Not ranked	0	3	37.50
18	[87]	0.5	0	0	1	CORE A	1.5	3	37.50
19	[40]	0.5	0	1	0.5	Not ranked	0	2	25.00
20	[82]	1	0	1	0.5	Not ranked	0	2.5	31.25
21	[100]	1	0	1	0.5	Not ranked	0	2.5	31.25
22	[41]	1	0	1	0	Not ranked	0	2	25.00
23	[101]	0.5	0	0	0.5	Not ranked	0	1	12.50
24	[55]	1	1	0	0.5	Not ranked	0	2.5	31.25
25	[42]	1	1	1	0.5	CORE B	1	4.5	56.25
26	[68]	1	1	1	0	Not ranked	0	3	37.50
27	[83]	0.5	0	0	0.5	Not ranked	0	1	12.50
28	[44]	1	0	1	0	Not ranked	0	2	25.00
29	[45]	1	1	1	0	Not ranked	0	3	37.50
30	[80]	1	0	1	0	Not ranked	0	2	25.00
31	[88]	1	0	1	0.5	Q1	4	6.5	81.25
32	[2]	0.5	0	1	0	Q1	4	5.5	68.75
33	[102]	0.5	0	1	0.5	Q1	4	6	75.00
34	[76]	0.5	1	1	0	Q2	3	5.5	68.75
35	[79]	1	1	1	0.5	Q3	2	5.5	68.75
36	[10]	1	1	0	0	Q2	3	5	62.50
37	[11]	1	1	1	0.5	Q1	4	7.5	93.75
38	[23]	1	1	1	0.5	Q4	2	5.5	68.75
39	[103]	1	1	0	1	Q1	4	7	87.50
40	[58]	1	1	0	1	Emerging sources citation index	0	3	37.50
41	[90]	1	1	0	0.5	Q1	4	6.5	81.25
42	[104]	1	1	0	0.5	Q1	4	6.5	81.25
43	[46]	1	1	0	0.5	Q2	3	5.5	68.75
44	[89]	1	1	0.5	0	Emerging sources citation index	0	2.5	31.25
45	[37]	1	1	1	1	Emerging sources citation index	0	4	50.00
46	[62]	1	1	1	0.5	Q1	4	7.5	93.75
47	[63]	1	1	1	0	Q1	4	7	87.50
48	[105]	1	1	1	0.5	Q2	3	6.5	81.25
49	[84]	1	1	1	0.5	Q2	3	6.5	81.25
50	[33]	1	1	1	1	Q2	3	7	87.50

Table 9 (continued)

Sr.	Ref.	QAC 1	QAC 2	QAC 3	QAC 4	Rank	Rank score	Total score	QAC%
51	[86]	1	1	1	0.5	Q1	4	7.5	93.75
52	[106]	1	1	1	0.5	Q1	4	7.5	93.75
53	[107]	1	1	1	0.5	Q1	4	7.5	93.75
54	[108]	0.5	1	0	0.5	Emerging sources citation index	0	2	25.00
55	[93]	1	1	0	0.5	Q3	2	4.5	56.25
56	[108]	0.5	0	1	0	Q1	4	5.5	68.75
57	[72]	1	1	1	0	Q1	4	7	87.50
58	[85]	1	1	0	0	Not ranked	0	2	25.00
59	[48]	1	1	1	0.5	Q2	3	6.5	81.25
60	[96]	1	1	1	1	Q1	4	8	100.00
61	[109]	1	1	0	1	Q1	4	7	87.50
62	[92]	1	1	0	1	Q3	2	5	62.50
63	[110]	1	1	1	0.5	Q1	4	7.5	93.75
64	[111]	1	1	0	0	Q1	4	6	75.00
65	[77]	1	1	0	0	Q2	3	5	62.50
66	[78]	1	1	1	1	Q1	4	8	100.00
67	[112]	1	1	1	0	Q1	4	7	87.50
68	[113]	0.5	0	0	0.5	Not ranked	0	1	12.50
69	[3]	1	1	1	0.5	Q4	2	5.5	68.75
70	[38]	0.5	0.5	0	0	Not ranked	0	1	12.50
71	[114]	1	0.5	0	2	Not ranked	0	3.5	43.75
72	[130]	1	1	1	1	Q1	4	8	100
73	[131]	1	1	1	0	Not ranked	0	3	37.5
74	[132]	1	0	1	0.5	Book with 1 citation	0.5	3	37.5
75	[133]	1	1	1	0	Q1	4	7	87.5
76	[134]	1	1	1	0	Q2	3	7	75
77	[135]	1	1	1	0	Not ranked	0	3	37.5
78	[136]	1	1	0	0	Q1	4	6	75
79	[137]	1	1	1	0	Q1	4	7	87.5
80	[138]	1	1	1	1	Q1	4	8	100
81	[139]	1	1	0	1	Q1	4	7	87.5
82	[140]	1	0	0	0	CORE A	1.5	2.5	31.25
83	[141]	0	1	1	0	Q1	4	6	75
84	[142]	1	1	1	0	Q1	4	7	87.5
85	[143]	1	1	0	0	Q1	4	6	75
86	[144]	1	1	0	0	Q1	4	6	75
87	[145]	1	0	0	1	Book with 2 citation	2	4	50
88	[146]	1	1	0	0	Q1	4	6	75
89	[147]	1	0	0	0	Book with no citation	0	1	12.5
90	[148]	0	1	0	0	Not ranked	0	1	12.5
Summary									
					Q1	Q2	Q3	Q4	Sum
Total score					80.5	69	56.5	37.5	243.5
Average score					0.89	0.77	0.63	0.42	0.6763
Percentage					89%	77%	63%	42%	68%

congestion, any protest, and, the catastrophic situation can be monitor remotely. So data set of many vehicles form crowd-sourcing network [39–42]. Data collected from the sensory observation units are raw in nature and thus require pre-processing and integration in order to understand the actual meaning of observation. Text, CSV, images, videos, GIS-related data is the usual format of data in observational and ITS units.

GPS data of people's mobile phones and social media feeds are continuously being shared and present the complex association as linked data in RDF, XML or, SPARQL format. People on board and vehicular observation both can share information about the same event happening nearby. Thus, this idea has encouraged to apply data fusion techniques in different data sources and form Cyber-Physical and Social Networks (CPSN) for real-world applications [39, 43]. Thirdly, fixed sensors in smart cities, for instance, smart surveillance camera, smart meters, air pollution sensors, smart bins, irrigation sensors, Road Side Units (RSU), infrared detectors and, ultrasonic sensors are rich sources of data in IoV [44–46].

As defined previously in this section, data generated in IoV are gathering from heterogeneous sources in heterogeneous types, and form big data thus eligible for big data computation due to 5Vs properties. The prediction of 200 sensors on a vehicle by 2020 will also cause an increase in data volume. It is predicted that about 4000 GBs of heterogeneous data will be generated by smart vehicles every day [47]. It is also evident that traditional sensors, for instance, GPS and ECU, etc., generate less data than the modern sensors such as Lidar and Camera and thus encourage a large number of applications in the real world like autonomous vehicles [23, 48, 49].

Moreover, sensory data can be acquired from other vehicles as well (V2V), by using DSRC channels. This data may include alert messages or any updated information. Platoons of vehicles are the formation of vehicle channels nearby and at same geolocation can be used for the dissemination of control messages. Besides the use of data for V2V communication management, this data can also provide links between vehicles and smart infrastructure (V2X) by using WiFi specifically described as Intelligent Transportation System. V2X data can be used for traffic monitoring and navigational route management [1, 2, 50–55, 57]. Apart from vehicular, cellular and other sensory data acquisition from ground objects, IoV uses space and other aerial platforms such as drones, unmanned aerial vehicles (UAVs), satellites, and High Altitude Platforms (HAP) as data acquisition sources [46, 58]. The detail of sensory units is presented in Appendix.

5.3.3 Research Question 3: What Kind of Big Data Processing and Analysis System Architecture, Tools, and Methodologies are used in IoV?

Implementation of big data architecture and platforms in IoV is heterogeneous in nature. As a matter of fact, the IoV system integrates several different information technologies (sensory devices, information systems, electronics, and other physical objects) for efficient and secure information exchange to control and manage IoV operations.

The general big data architecture in IoV consists of six layers: (1) Data acquisition, (2) Data transformation (broking) and normalization (3) Data Storage, (4) Data processing (real-time, batch), (5) Data analysis and, (6) decision making. The data is gathered from different sensory and observer units and then prepared for transmission and data storage. Normalization, data fusion and, data brokering are some of the famous methods for data integration, data reduction, and compression. These methods are helpful to deal with the heterogeneity data transformation issues, for instance, naming conventions, formats, etc. Data Storage in IoV manages heterogeneous data for supporting the data analytics and decision-making processes. Repositories, caches, data warehouse, data lakes, unstructured data, etc. are instances of data storage. MongoDB, Cassandra and HBase are market solutions for databases in IoV [59–61] Data processing in IoV is subdivided into two categories as real-time (stream) and batch data processing. Real-time processing is used to preprocess the live data prior to transferring to the database to reduce the processing load. As the noisy data may bring processing load to database and it is also evident that updating operations in database are already an inevitable load because many different applications may be acquiring it to record and can halt the processing operation. Data-driven methodologies are used to handle such conflicts on the subset of stream data that can be extracted and managed at distributed locations [62, 63].

This process can reduce the processing load for a database and can increase the efficiency and scalability as well. The basic real-time operations in IoV are data collection, acquisition, resource allocation, management and, automatic handling of events and generate alert messages. For IoV several market solutions are available for real-time processing for instance Flume, Flink, Samza, Storm and spark etc., [64–68]. Furthermore, batch processing is based on archive data and used to manage the complex processes by using artificial intelligence for further exploration and management. Hadoop, Pig, and hive are some of the different applications for batch processing in IoV [69–71]. Analysis Layer of IoV performs analysis on data gathered from the real world, which is actually a representation of the virtual cloud environment to process, sort and analyze retrieved and gathered data from the first layer. This layer works as a Centre for information management where Vehicular Cloud Computing techniques and big data analytics (BDA) work together for operational management. Importantly, service management is also a major concern of this layer where dedicated services are provided to manage the applications for users to make better decisions. Lastly, the expert systems and all analytical solutions and interpretations lead to intelligent decision-making [72].

Big data platforms in IoV are available to facilitate IoV applications. For instance, Traffic Telco architecture is proposed to gather and analyze taxi GPS data to analyze the route density for rerouting and reducing traffic congestion [73]. Moreover, social media feeds are used to create the heat map of crowdsourcing data using Hadoop, Apache Spark and, Hive [69, 71, 74]. Data analytical techniques are employed to identify the best route. In addition, an interface for simulating the real-world traffic flow data is also part of this architecture and is facilitated to reduce traffic congestion. In smart cities, another application in IoV named as Hut was used to gather and process transportation data [91]. Node-RED is used for data acquisition from heterogeneous objects. The processing engine was based on machine learning approaches and provide cloud-based storage by Amazon S3. Before storage Apache Kafka broker was used to transform the data streams into a concrete format.

MLib by Apache Spark is deployed as the data analytics engine to process more than three thousand sensory observational data from Madrid, Spain [35]. Lamda is another complete solution for distributed traffic analysis and can process stream and batch based data, simultaneously [52]. The working of this architecture is divided into three layers; speed, batch, and serving. Each layer has its own responsibility from data gathering to analysis and visualization. The data is acquired from heterogeneous devices by using common web service interfaces. Speed layer processes real-time processing solution by Spark whereas; batch layer processes distributed and archive data for data analytics. Moreover, the serving layer enables real-time and archived datasets for consolidated processing and insights. HBase and Hadoop were used for archive data processing and have the ability to process real-time and batch processing separately as well as consolidated as required [97]. LEGIS was used as a visual interface for mapping the results of real-time traffic patterns. Lastly, video analytics is also an important feature of this layer to analyze traffic behavior, number of vehicles, road intersections, traffic signals, and speed limits, etc. This data then proceeds to the service layer of the architecture for processing [90].

Table 10 provides a summary of big data processing and analysis methodologies, data acquisition sources, tools and, classification of currently available solutions. Additionally, the purpose of the analysis is classified in descriptive, predictive and prescriptive analysis to demonstrate the level of significance in real-world applications of big data in IoV.

5.3.4 Research Question 4: How IoV is Using Predictive Insights for by Big Data?

In this section, we present how IoV is using data processing and predictive analytical techniques in real-world applications. Data processing either batch or real-time requires the results that can be mapped to resolve real-world operations. We have categorized these predictive insights for IoV in five major categories: Urban Planning, real-time observation, and dissemination of information, Collision detection and prevention, economic benefits, safety, and security. And the summary of the key applications of these insights is presented in Table 11.

(i) Urban Planning: IoV archive data can support urban planning activities by using evidence-oriented approaches because heterogeneous and multidimensional data collection can strengthen the decision-making processes rather than relying on unquantifiable assumptions. Firstly, IoV addresses problems for traffic management by replacing the fragmented and self-reporting system by an extensive system for accurate results. These results are used to observe traffic dispersion and flow thus making it easier to guide drivers or autonomous cars for rerouting [1, 34, 65, 66, 68, 71, 91], Secondly, parking automation solution is also developing area. People on board can book, navigate and pay parking tickets using smartphones or even displaying units in cars. As a result, parking management can reduce the blocking states and lead to an increase in parking capability. Moreover, the traffic situations can help to reconstruct infrastructure for the elimination of problems that cause road congestion [73], and simulation software can be helpful to test prototypes before the building which reduces the overall cost of construction. Smart traffic lights are the next area of interest in urban planning. Vehicles have to wait many hours at traffic signals which causes fatigue and fuel wastage. Automatic controlling of the signals may lead to better resource management [9]. Lastly, public transportation can be organized by pattern identification and correlation analysis for scheduling and route optimization [54, 76, 88].

(ii) Real-time observation and dissemination of information: Although smart vehicles ensure the ease and comfort of drivers, the vehicle itself is the source of data collection. A smart vehicle can capture and disseminate sheer amount of data such as GPS location, fleet and trajectory moreover people on board either drivers or passengers are also sharing valuable information regarding road condition and social media feeds [80, 90, 91]. This crowdsourcing data provision behavior of people thus can recommend better services. Informed decisions are better than the individual decision and IoV can assure the potential benefits for end-users [51, 76, 79].

Moreover, traffic congestion in cities can have many causes, for instance, poor weather, catastrophic events or social events or accidents; if people will receive the instant information they can plan or commute accordingly [50, 51]. Additionally, road accidents can be detected in advance such as a sharp road turn may cause 40% of the accidents and if this information is exchanged with driver promptly before reaching this point can save human lives. Emergency systems and assistance can provide for recovery and evacuation, if necessary by using mobile applications and billboards

[50]. Predictive analysis of this extensive multidimensional and multivariate data will have extraordinary decision support for instance; reasons for traffic congestion can consolidate for external and government agencies to plan accordingly. Data from IoV will create a rounded picture for many real-life aspects and will not only increase the efficiency rather provide the financial and environmental benefits as well [77, 84, 92, 113].

(iii) Collision detection and prevention: Safety is the first most concern of every person while he is traveling or walking on the road and advance analysis to detect collisions will not only improve the traveling experience rather it will reduce the probability of road accidents as well. Numerous solutions for collision detection are based on visual understanding by deep learning techniques [55, 102]. Initially, the identification of spots at roads that have higher accidental ratios have been identified and then rigorous learning algorithms are applied for understanding the collision nature. This data can be helpful in the creation of learning algorithms by using the data aggregation properties for instance; causes, frequencies, and categories of collisions. Moreover, this learning can provide consolidated information by visualizing techniques and thus this information is an asset for government authorities for policymaking, infrastructure changes and, campaigning. Video analysis of real-time and archive data ensure the proactive safety procedures for potential risk assessment and can improve traffic behaviour and safety standards [3, 3, 49, 99].

(iv) Economic Benefits: It is evident that smart vehicles will not be cheaper as compared to the current market price of models available but will lead to saving huge financial benefits in the future. According to the research [116], in Germany only, the drivers have to tolerate traffic congestion for 21 years, in 2011 and which is equivalent to 450,000 kilometers for drivers to stay on the road approximately. Furthermore, in the US this wastage is about 200 million hours per year with $115 billion wastage in 2009. If we estimate the amount of time spent, fuel wastage and, destroying other environmental factors it is equal to 100 million euro which is too much in aggregate. With IoV, it is easier to prior identify the patterns for traffic congestion, road accidents, social or political events and, roadwork. This identification will analyze and interpret information using real-time data processing and provide an advantage to reroute in advance and will lead to a decrease in the fuel amounts and environmental issues.

(v) Safety and Security: Technological advancement is encouraging scientists and researchers to build security algorithms and standards to ensure the security of devices. In addition, market contributors are adapting these security standards to improve security for people on board or on the road. As the ability of connection with servers has a significant impact on our daily lives as every user can receive the update reports on weather, road condition, potential accidental spots, forthcoming collisions or, any event that is causing road diversion in advance [54]. Therefore, people can plan their routes accordingly. Sensory observation and connection have not only developed the IoV, rather they have fortified the idea of autonomous cars. These cars can recognize objects, disseminate information and create a link with servers so it is not possible to lose or suffer any damage. Adapting a change in a dynamic environment, early warning system, safety recommendation, rule-based predictions, and service management in critical situations are prominent benefits of IoV [114] (Fig. 6).

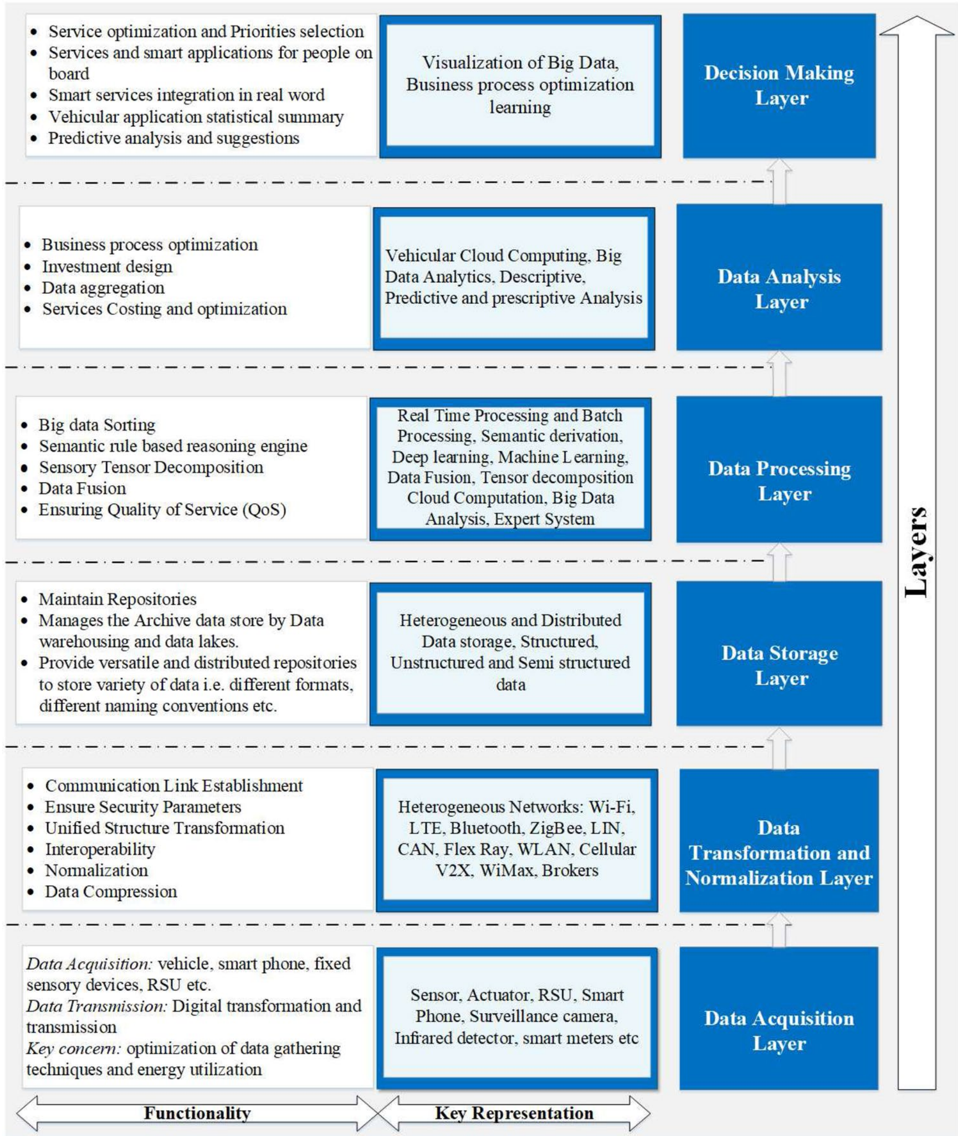
5.3.5 Research Question 5: What are the Substantial Challenges and Opportunities of Big Data in IoV?

After the literature review of the domain, it has been noticed that authors have discussed several challenges because of the critical nature of IoV. In this section, we will summarize data collection and storage, privacy and security, quality and fairness, and lack of standardization and legislation. Figure 7 represents the illustration of the key challenges of big data in IoV.

(i) Data collection and storage: The multidimensional data generation from heterogeneous devices (i.e. in-vehicle sensors, RSU, camera, road sensors, and smartphones, etc.) requires intensive analytical interpretation for extracting knowledge from it [84]. Thus, there is a need for real-time analytical engines to process data efficiently. Moreover, it is also evident that these devices are in third party ownership, for instance, telecommunication authorities and thus collected data is not directly accessible for government agencies to provide the potential benefits for users in the future. Although the archival data has opened the opportunities to develop new algorithms for artificial intelligence. There is also a need to research governance, security, and ethics in order to manage positive solutions [92, 109, 117].

The challenges of data gathering are not limited to direct access rather there is also a need to pre-process data to violate noisy data points, but unfortunately due to large heterogeneous data, it is quite difficult to even pre-process the data. In addition, the formats of data may also cause an issue of interoperability and thus become a headache for data scientists to handle these problems. High cost of devices and difficult maintenance are also some of the challenging issues of governments because the up-gradation of these solutions is very advance [106]. Mobile mounted sensory devices are although costly but more useful than fixed devices and can be a rich source of information. The results of these devices are subjected to weather conditions [50]. Moreover, diverse data types, properties, and, formats are other challenges. Because the repositories for IoV data are MongoDB, Cassandra, Neo4J and no SQL, etc. and to store

Fig. 6 Architecture of IoV with respect to big data processing and analysis



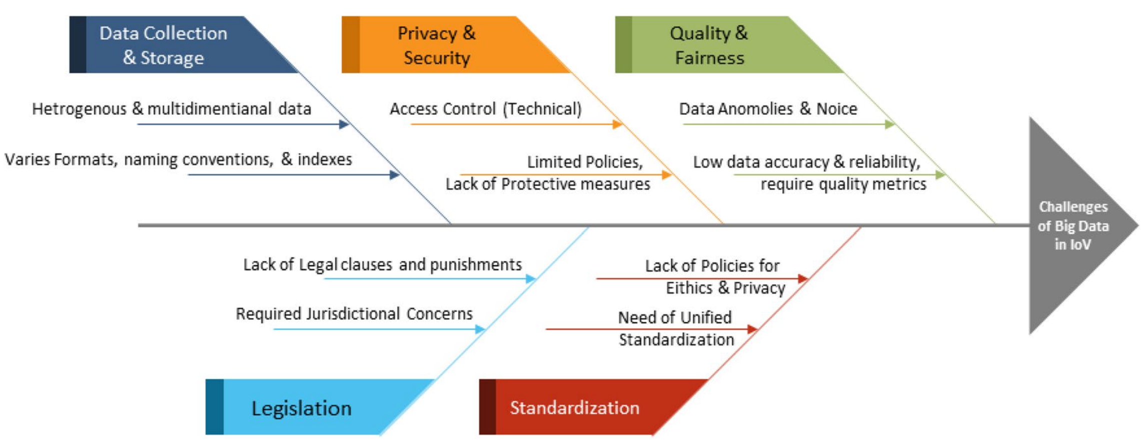


Fig. 7 Illustration of key challenges of big data in IoV domain

Table 10 Summary of methodologies, tool, technologies of big data in IoV

Purpose	Method	Data sources	Big data analysis	Classification of processing	Tools
Descriptive analysis	Data integration	ITS	Big Data fusion [24, 30, 32, 37, 40, 48] Big Data integration [38, 46, 72, 79, 91, 98] Record linkage, schema mapping [23, 24, 31, 32, 108, 115]	Real time, batch processing	Apache Tika, Talend, Alteryx, pentaho
	Data mining	ITS, SAG (space, air, and ground) data	Classification, clustering, association rule mining, correlation analysis [31, 38, 79, 82, 83, 86, 89, 90, 133, 135, 136, 138]	Real-time, batch processing	Hadoop [62, 69], Hive [64, 71], Strom
	Trajectory analysis	Urban mobility and trajectory data	Resource optimization [75, 81, 82, 88, 98, 104], Route management, route planning [26, 34, 46, 48, 48, 49, 54, 56, 56, 57, 73, 74, 80, 81, 105, 111, 134, 138], Energy efficient and intelligent vehicle (EEIV) [15, 23, 30, 37, 69, 75, 76, 76, 82, 83, 130, 136, 141] Vehicle operations and manufacturing, vehicle distribution, after sale services, insurance, fleet management [71, 76, 78, 83]	Real time	ADAS (Advanced Driver Assistance Systems), Spark [86, 93], Flink [58, 65], Charged coupled device (CCD) sensory system [104, 110]
	Scheduling	Logistics data	Resource management [26, 31, 46, 48, 49, 54, 56, 57, 73–75, 80, 81, 81, 82, 88, 98, 104, 105, 111], Energy management [15, 23, 30, 37, 69, 75, 75, 76, 76, 82, 83, 83, 130, 132, 140, 141], Optimization [75, 81, 82, 88, 98, 104]	Real time, batch processing	Garmin, StreetLine, esri, Yarn, Apache Spark etc.
	Statistical and econometric	Weather data, drivers and road data	Route planning and management [26, 34, 48, 48, 49, 54, 56, 57, 73–75, 80, 81, 81, 82, 88, 98, 104, 105, 111]	Real time	Spire, esri, Factual, Scala etc.

Table 10 (continued)

Purpose	Method	Data sources	Big data analysis	Classification of processing	Tools
Predictive analysis	Neural networks	Time series	Collaborative applications, resource optimization [75, 81, 82, 88, 98, 104, 134], deep sensing [2, 50, 81, 88] Semantic learning [51, 58], Route planning and management [26, 34, 46, 48, 49, 54, 56, 57, 73–75, 80, 81, 81, 82, 88, 98, 104, 105, 111],	Real time	Keras, Neuroph, NeuroSolutions, ConvNetJS
	Short-term prediction models	ITS	Object detection [51, 58, 98, 100], Lane detection [38, 46], Route planning [26, 34, 46, 48, 49, 54, 56, 57, 73, 74, 80, 81, 105, 111, 138, 141, 142], Freeway management [98, 104]	Real-time	Microsoft Azure, gamelan, Deep-Sense.ai, Amazon SagaMaker, etc.
	Long-term prediction models	Urban roads, bus and transit, bus transport	Congestion management [26, 105], Traffic jam prediction [26, 34, 79, 86, 87, 94] Transit management, freight management, city-wide transportation management [26, 77, 79, 84, 86, 86, 93, 99, 101, 105, 106, 108]	Real time, batch processing	Google Analytics, Microsoft Azure, Amazon Rekognition, Penthao, Datameer etc.
Prescriptive analysis	Heuristic algorithm	Vehicular routing	Optimizing routing [75, 81, 82, 88, 98, 104, 104, 110, 142], Emergency detection and earlier Recovery [72, 79], Object and hindrance detection [38, 46]	Batch processing	Python, Anaconda, Orange, R
	Deep learning	Vehicular routing, Time series	Vehicle license plate recognition [25, 30, 33, 37, 38, 46, 65, 72, 89, 96], traffic accidents detection [72, 79], deep learning [2, 50, 81, 88] and social media data sources [37, 45, 70, 77, 83–85, 90–92, 145], Traveler satisfaction [25, 33, 38, 46, 86, 93, 147]	Batch processing	Tensor flow, Keras, Caffee, Microsoft cognitive tool kit, Apache Singa etc.

Table 11 Predictive insight and applications of big data in internet of vehicle (IoV)

Predictive insight	Key features
Urban planning	Traffic guidance [41, 96] Navigation [1, 34, 54, 56, 56, 57, 80, 81, 111], Emergency services [79, 131, 132] Location based services [34, 42, 50, 52, 54, 56, 57, 75, 80, 87, 88] Crowd sensing, sourcing and management [45, 77, 90–92] Resource planning and optimization [82, 88, 98, 104, 134] Improve traffic congestion [85, 105] Air pollution [77, 78, 136, 137] Estimation of capacity changes [23, 38, 139, 140], Parking management [89, 102]
Real-time observation and dissemination of information	Capture and disseminate traffic videos [3, 76, 110, 133] Tracking vehicular movements [37, 57, 111] Identification of driving patterns to prevent probable accidents [79], Crowd sourcing [45, 77, 90–92, 136]
Collision detection and prevention	Identification of near-collisions of cars with pedestrians and bikes [38, 55, 110] Object detection and segmentation [77, 87, 110] Provide emergency solutions [79], Management of Internet of Battlefield Things (IoBT) [96]
Economic benefits	Induction coil usage [45, 108, 130] remote vehicle condition diagnosis [97], economy and governance [32, 36, 57] efficient, sustainable, competitive, productive, road safety/infotainment [68, 80] intelligent transportation systems [48, 113], and self-driving systems [48, 82, 108]
Safety and security	Identification of IoV security attacks as GPS deception, masquerading attack, wormhole attack, development of security models as Microsoft's STRIDE, SAODV, Ariadne and SRP, TESLA authentication scheme, SLPD, ALAR, and STAP, provide evidence forensic [33, 62, 78, 85, 96, 108, 112–114, 132, 144, 145, 147]

a variety of formats in a single data store is not possible [60, 118]. For instance, document-based databases and Moving Object Databases (MODs) are better for spatio-temporal and GPS related data storage (MongoDB, Cassandra, etc.) with parallel and distributed processing and, social media feeds are stored well in graph and association oriented databases. Due to distributed, heterogeneous and variety of datasets the main challenges are efficient query processing, need base data processing, transformation and emergence of multivariate data into the concrete form [59, 60]. The issue of multivariate data processing is well demonstrated in [37], where three different database systems are used for data storage. (1) PostgreSQL, a relational database was used to store relational data; (2) HBase was used for non-relational storage, and (3) Memcached were used for caches data that was stored in memory for performance optimization. This description demonstrates the headache of data storage for different formats and this effort leads to meet the time-related data processing [112].

(ii) Privacy and Security: In IoV, the main interconnected objects are vehicles, RSU, sensors, actuators and, servers, etc. These devices transmit data using untrusted networks in order to complete different tasks. Thus, security threats in IoV can be categorized into five categorizes as (i) connected vehicle security, (ii) smart device security, (iii) security related to services provided on-board, (iv) V2X security and, (v) Data security. People on-board or on the road are unaware of how much of their data is collected, dissemination and storing at distributed servers. Unconsented parties such as Google Maps, Apple Maps, and android-based location generators are continuously observing and storing data [112]. As a result, people have to bear the data vulnerability that can cause hacking, data leakage and, information stealing. Moreover, credit cards and other payments during travel such as parking, bus and toll ticket, etc. prepare a log file of travel from source to destination [54, 102]. This data can be used to identify the patterns of spending and movement and indeed are susceptible points. DoS, DDoS attacks, Botnets, injected codes, timings attacks, firmware analysis, and sinkhole attacks, etc. are some of the well-known attacks in IoV. Data from smart vehicles and devices, however, is not available for government agencies, authorities and companies, directly but can be accessible at any point of movement. So, security experts are forced to develop effective privacy regulations and standards. Moreover, it is also evident by [119], that the reported IoT attacks were increased by 600% in just one year from 2016 to 2017. The report also included a summary of these attacks as 21% from China, 11% in the US, 6% in Russia and, 7% in Brazil. These attacks will ultimately affect the communication channels for IoV and ITS. Therefore, there is a need for security and privacy regulations to prevent the vulnerabilities and harmful consequences that can cause security obstacles. General Data Protection Regulation (GDPR) and Federal Trade Commission are initiatives towards enhancing the obligation for enhancing protective measures in IoV but this is an open research area to handle future threats [112, 120]. (iii) Quality and Fairness: Quality and fairness matrix in IoV falls into three different categories; (i) data quality (DQ) that have been acquired from multidimensional and heterogeneous space, (ii) information quality (IQ) relies on the level of satisfactory decisions, and (iii) Service Level Provisioning (SLA) for creating a channel between user and authorities or agencies in order to provide extensive application-level quality standards. These three considerations are important in terms of IoV application quality but the main challenge is to measure the quality of data and information, because the acquired data has noise, uncertain values, anomalies and erroneous and can thus cause inaccurate decisions [84]. Therefore, the rigorous data preprocessing and collection methods are needed because the traditional preprocessing methods are not enough in this regard. Lastly, the SLA needs to update the process of issuing digital certification due to the nature of the domain with changing technological parameters. Interoperability, security, limited connection, performance, privacy, and legislation are also important issues in the IoV domain and are addressed thoroughly [26, 112]

(iv) Standardization and legislation: All functions in IoV are based on heterogeneous interaction channels and thus have added vulnerability in communication. Rather SLA, SSL, and digital certificates are maintained during interaction but due to the intelligent hacking software, these are not enough. In fact, in these links, human interaction is supposed to be minimal but due to the primary user and beneficiary of IoV, there is a need to encompass and creation of policies and standards in order to maintain confidentiality, integrity, and authentication [121]. Until today, the challenging question is that who will develop the standards and policies in IoV and who has the authority to enforce them? From the last decade, it also has been observed that the attacks on data and communication channels are increasing day by day thus this is the open dimension to study the security aspect for smart and connected vehicles. In response, some of the well-known legislation plans are US ITS plan 2015 [122], SAE J3061 [123], NHTSA 2016 [124], IEEE 1609.2 2013 [125] and ETS-TS-102949 [126], but the isolation of trusted boundaries are continued, therefore new standardization is required to address and enforce the performance, quality, and security policies to improve the IoV security and privacy issues.

6 Limitations

The key limitations of the review are biased publications, misclassification, inaccurate data retrieval, and wrong interpretation for selection. The reason behind this misleading information is the positive representation of studies and proposed methods because authors usually negate and hide negative and conflicted concerns. This is also observed that the frequency of publication of negative results and variance factors are very minimal [28, 127]. In order to avoid such misleading and inaccuracy, we have explored papers in well-known repositories and data stores as described in the research methodologies section of this study. The research protocol has ensured to find the maximum number of papers and broaden search techniques has also helped to identify the papers with negative concerns and challenges of studies. However, it is also possible that Big Data in IoV is a new and emerging topic in the computing field, as both industry and academia have an interest in it. So, the possibility of research conducted by agencies or organizations might be available in the form of white papers or may not be published and available for the general public. Therefore, the identified papers and studies might not have addressed and included these findings in the systematic review. However, scientific databases and reputed platforms are used for the findings of relevant papers which have resulted to only retrieve the higher quality papers.

The selection bias referred to the falsification of statistical results and analysis of the selected publication and we have rigorously addressed this issue by creating a strict protocol for the selection of papers [28]. We have used keywords, terms, and expressions for ensuring the inclusion of as many as possible relevant papers. The designed protocol has restrictively defined the inclusion and exclusion methods for selection. However, Big Data in IoV is the main search domain for the study and it is possible that all relevant research papers for emerging issues and techniques are not included in this review. Because most of the research has addressed the Big data processing and analytical techniques separately and explained only single or two methods for the experiment. The goal of this study was to identify the available architecture, techniques, methods, and tools for Big data in IoV. Based on the main plot of the study, we have made sure to retrieve only relevant papers to the search terms. Moreover, the inaccurate data extraction method and misclassification lead to the understanding of different reviewers for the same question. To solve this conflict, we have developed a group of two reviewers to retrieve relevant research papers and discussed the relevancy in order to include or exclude the research paper. In addition, the full-length study has been done for selected papers to satisfy the research questions.

7 Discussion

The main purpose of this review is to develop an understanding of research contributions, applications, and architecture in the field of Big Data in IoV. The aim of this study is to examine and explore the current practices and solutions to optimized services with efficient and sustainable resource management in IoV by using big data, rather than identifying the solutions for the problem domain. Most significantly, during the literature review, it has been noted that Big Data is an emerging field and still in the initial phase. Most of the projects of Big data in IoV are either conducted in industries or academic research institutes that have limited scope. During this review, we have also observed that the proposed frameworks and architectures in research are at an abstract level only and tested at simulation software for validation but the real-world scenarios are different from simulations due to several other factors as well [11, 97]. Consequently, these ideas and projects are infancy thus cannot provide measurable advantages to end-users in the real world. And require many stakeholders to manage and consolidate findings, for instance, service providers, government agencies and authorities, etc. Thus to implement IoV and process big data is a time taking process. Moreover, the security and privacy of people is another concern because people on board or on roads want to be secure, therefore, there is a need to develop advanced standards and legislation in this regard [109, 114]. System performance is the most important aspect of big data processing since the user requires an instant response and alert messages. And it is also evident that traffic on roads is unpredictable so as the other events [54]. Thus performance testing and quality of services need to be defined accordingly. Different available platforms, for instance, Hadoop usually manage and process a single format of data. Since IoV data is heterogeneous, multivariate, and available in different formats [97]. So, it is required to prepare a distinct data storage and processing platform to manage IoV data. Available platforms and solutions should be available as open-source for updating and testing these solutions. Lastly, the big data field is still in the maturing process and an open innovative paradigm for getting the best practices. Thus providing emergence with IoV will require an open contribution from a different perspective in the real-world.

8 Conclusion

Smart vehicles are now getting eminent due to their extensive sensory and communication abilities and thus the government agencies and companies are trying to develop sustainable and efficient policies in order to manage the quality of services in IoV. Big Data processing and analytical methods in IoV are required for management of large datasets. Because the heterogeneous and multivariate data enables decision-makers to collect and triangulate shreds of evidence from the real-time environment to make optimal decisions.

Thus in this study, we have explored the key features of big data acquisition, storage and management in the IoV domain. A systematic literature review has been made to validate these features from well-known repositories, journals and conferences. We have explored that, the conventional gathering, processing and analyzing methods are not enough to yield the optimal and relevant insights so we have summarized the big data based architecture, tools and methods in IoV. This assessment and review have been done until year 2019 to examined industrial and academic research and development contributions. And we have defined the taxonomy of using big data in IoV. Several convergences and applications of IoV technologies also have been discussed in this review paper including urban planning, real-time observation and dissemination of information, collision detection and prevention, economic benefits, safety and security. Lastly, we have presented the open research challenges relevant to the data collection, storage, standardization, and quality and fairness of services in big data IoV. This study concluded with the idea that big data will have profound impact on the future design of IoV and intelligent transportation systems and will provide the base for safety, reliability, efficiency and potential economic benefits.

Appendix

See Tables 12 and 13.

Table 12 Key research contributions

Study title	Key research contributions
Big Data analytics for connected vehicles and smart cities [32]	A comprehensive book for big data analytics applications for connected, smart vehicles, and transportation systems Specifically describe the data analytics expansion in design, development and, engineering of smart and connected vehicles. And how it can strengthen the different applications in real-world such as education, economy, advanced technology, environment, safety, and energy
A study of big data solution using Hadoop to process connected vehicle's diagnostics data [97]	A book chapter presents one experiment by using vehicle diagnostics data for useful insights into the IoV ecosystem Apache Hadoop framework with Hive, Sqoop have been deployed to test the outcomes
Short-term traffic predictions on large urban traffic networks: Applications of network-based machine learning models and dynamic traffic assignment models [53]	Implicit and explicit models provided for urban road networks for short-term traffic predictions Bayesian and Artificial Neural networks have been applied to test the Rome traffic and smart vehicle data
Integrating connected vehicles in the Internet of Things ecosystems: Challenges and solutions [98]	oneM2M architecture have been presented for efficient and more sustainable transportation system A framework of the proposed architecture has been implemented by using a prototype
Aggregating and managing big real-time data in the cloud-application to intelligent transport for smart cities [1]	The objective of the study was to manage and aggregate big data IoV data by cloud services for optimal decision making Data cleaning, storage, integration have been presented in a service-oriented cloud-based system for transportation management
Providing clarity on big data technologies: a structured literature review [26]	A systematic literature review has been prepared to consolidate classification approaches for future big data projects by using Hadoop in different applications for instance IoT, IoV, and smart infrastructure
Real-time smart traffic management system for smart cities by using the Internet of Things and big data [56]	In this study, they have presented one framework as STMS based on big data IoV Also have proposed the low-cost resource management in smart traffic management by using predictive analysis
Traffic big data analysis supporting vehicular network access recommendation [57]	In this paper, the traffic model for network recommendation has been developed by using big data analysis Secondly, presented the analytical framework for the vehicle recommendation system by several data sources such as traffic status, user preferences, service applications, and network conditions An Android application has been developed for enabling the vehicular access network
The paradigm of big data for augmenting internet of vehicle into the intelligent cloud computing systems [75]	Provided big data mining solutions, and how vehicles and roadside's sensors and moveable agents generating data, including mobile, time series and other sensor devices collect and process data in IoV
Social networking and big data analytics assisted reliable recommendation system model for Internet of Vehicles [91]	This paper has presented a comprehensive survey on SIoV Also have proposed a social recommendation model to establish links between social networking and SIoV. The future IoV system design also has been proposed in this study based on CPSN
Security of communications in connected cars modeling and safety assessment [99]	Represents the architecture of Vehicle to Internet (V2I) and V2V based on security standards and describes how the security attacks can detect and prevent in IoV environment
Big data analytics for logistics and transportation [94]	Presented a comparative review on the latest applications of big data analytics in IoV and also have prepared one experiment to detect the container code using Hadoop
Towards real-time road traffic analytics using telco big data [34]	Outlined the components for the traffic-TBD (Traffic Telco Big Data) and how to process the big data for road analytic and prediction system
Real-time traffic-based routing, based on open data and open-source software [52]	Described how cloud-computing and big-data management technologies can assist in decision making in IoV Also presented the proof-of-concept prototype and experiment by publically available data for the city of Edmonton for route planning, traffic management

Table 12 (continued)

Study title	Key research contributions
Scalable transportation monitoring using the smartphone road monitoring (SRoM) system [50]	Proposed the Smartphone road monitoring (SRoM) system to identify the sensing capabilities of road sensory networks Identified the big data processing and analysis for crowdsourcing, road safety events, and many applications
A predictive data-driven model for traffic-jams forecasting in smart Santander city-scale testbed [51]	Demonstrated the predicting traffic jams based model for QoS The cross-validation method is used for training by Neural Network and Decision Tree algorithms
Strategway: web solutions for building public transportation routes using big geodata analysis [54]	Provided the solution to collect data from heterogeneous resources, and big data processing by using geo objects and machine learning algorithms for optimal route planning
Future connected vehicles: challenges and opportunities for Spatio-temporal computing [87]	Prepared a comprehensive study to identify some of the research questions in IoV Created a summary of the vast opportunities and challenges of using big data analytics in IoV
Visual fusion of mega-city big data: an application to traffic and tweets data analysis of metro passengers [40]	Used smart card and twitter data to develop the visual integration of transportation data to demonstrate the possibilities and usefulness of time series visualization for optimal learning
Autonomous vehicles safe-optimal trajectory selection based on big data analysis and predefined user preferences [82]	Presented a model for trajectory management in autonomous vehicles Also have used Big Data mining and analysis to identify the real-life accident data and real-time connected vehicles' data. Driving preferences and trajectory management have been proposed in this study
Supporting large scale connected vehicle data analysis using HIVE [100]	Presented a case study of using HIVE for bug data processing and analysis to enable transport planning Also, have investigated big data approaches by comparing the use of HIVE with MapReduce and Spark programming
Real-world applications using parallel computing techniques in dynamic traffic assignment and shortest path search [41]	Demonstrated the real-world big data-enabled applications to address the parallel computing algorithms And how to achieve high performances in real-time management operations by preparing the first test case to examine the traffic data of the Austria region for traffic forecasting and, route guidance
A framework for smart transportation using big data [101]	Proposed a framework of big data-enabled smart transportation And how big data analytics can be used to enable smart transportation and IoV by using resource Optimization and Data exploration methods
Traffic accidents analyzer using big data [55]	Developed an application based on web services to analyze and visualize the major traffic accident information and accident prevention by using MapReduce in Hadoop
Vehicles as big data carriers: Road map space reduction and efficient data assignment [42]	Proposed an embedded algorithm to compute an offloading overlay on-road data for optimal data management and processing in a big data environment
Big data analytics architecture for internet-of-vehicles based on the spark [68]	Presented the data fusion techniques for IoV in the big data environment Data acquisition methods (GPS, camera, radar, fixed), traffic incident information, infrastructure operation status, meteorological, and environment information
Impacts of IoT and big data to the automotive industry [83]	Identified the impact of energy-efficient and intelligent vehicle (EEIV) How different applications of IoV and automative industry can benefit by using big data analytics such as vehicle operations and manufacturing, vehicle distribution, after sale services, insurance
Big data analytics architecture for real-time traffic control [44]	Represented the data acquisition techniques by periodic vs. non, periodic, descriptive vs. predictive, real-time vs. non-real-time, single vs. multiple sources And how IoV can be used for urban planning, transportation operation, safety and traffic control in a real-time environment
Taxi of the future: Big data analysis as a framework for future urban fleets in smart cities [45]	Presented the smart infrastructure based big data definition as urban data, mobility data, energy-related data for logistics management, public transportation, infrastructural, social and how mobility sector can be strengthened by using IoV based solutions

Table 12 (continued)

Study title	Key research contributions
Traffic big data-based path planning strategy in public vehicle systems [80]	Proposed architecture for public vehicle (PV) and how the traffic-management and QoS, can be optimized by using big data processing tools
Optimization of real-time traffic network assignment based on IoT data using DBN and clustering model in smart city [88]	Represented the real-time traffic network for optimal resource management by using deep learning and clustering Uses the GIS, speed, route data for cost management, resource management in smart cities, and transportation
Traffic flow prediction with big data: a deep learning approach [2]	Simulated an experiment that how deep learning on big data can be used for traffic flow prediction and rerouting
Deep learning for decentralized parking lot occupancy detection [102]	Used deep learning and big data analytics for occupancy detection and parking management
Cyclist-aware traffic lights through distributed smartphone sensing [76]	In this study, they have explored different types of data in IoV can be acquired in a smart environment as GPS coordinates, time, velocity and direction process: smartphone sensing, power consumption, time-of-arrival estimation, kiosks for big data processing and extensive future predictions
Cognitive Internet of Vehicles and disaster management: a proposed architecture and future direction [79]	Proposed a solution for using cognitive system capabilities in IoV And how these services can be applied in real-time applications for instance disaster management and recovery
Big data: a survey [10]	A comprehensive survey has been presented of all available frameworks, tools, and methods specifically in real-world applications
An unlicensed taxi identification model based on big data analysis [11]	Presented the model, framework, prototype, and architecture by an advanced traveler information systems (ATIS), advanced vehicle control systems (AVCS), business vehicle management (BVM), advanced public transportation systems (APTS), and advanced urban transportation systems (AUTS), D2ITS Simulated different algorithms as Markov chain based Bayesian decision tree algorithm, yielded learning-driven ITS, real-time Kalman filter model, SVM, feature refinement model for optimal decision making in big data and smart infrastructures
Development and transport implications of automated vehicles in the Netherlands: scenarios for 2030 and 2050 [23]	Developed the future case scenarios to prepare the SEEPT framework. For several IoV applications as innovation and business planning, forecasting and strategic policy analysis, automated vehicles, automotive human factors and biomechanics, micro-electronics, vehicle type approval, vehicle engineering, traffic management, transport modeling, geography and geo-informatics, accessibility and regional economics, estimation of capacity changes
Coalition games for spatio-temporal big data in Internet of Vehicles environment: a comparative analysis [103]	Used big data analysis and processing techniques in spatiotemporal data stores in IoV Compare the results of different algorithms as Bayesian coalition game (BCG), learning automata (LA), Nash equilibrium (NE), probabilistic belief for IoV applications
Internet of vehicles in big data era [58]	In this study the heterogeneous data sources as on-board sensors (e.g., inter-vehicle distance, blind point objectives, pilot camera video, etc.), traffic light status, neighboring vehicle brake notification, road map, driving state (Velocity, tire presser, etc.), (coolant temperature, RPM, etc.) have been presented for many IoV applications such as: vehicle management, in-vehicle HD video streaming, autonomous driving, HD map, Los A2G connections, drone communication, crowd-sourcing. Moreover, they have presented how to manage big data storage by using solid-state drive (SSD), Cohda wireless OBU, NVIDIA DRIVE PX.
Social big-data-based content dissemination in the Internet of Vehicles [90]	Used social big data for connection probability estimation by cumulative distribution, regression, classification, clustering, latent variable modeling, Bayesian nonparametric learning, for peer discovery, power control, and channel selection in IoV

Table 12 (continued)

Study title	Key research contributions
Cooperative fog computing for dealing with big data in the Internet of Vehicles: architecture and hierarchical resource management [104]	Presented the big data-enabled CFC-IoV (architecture) based on a fog-computing-based intelligent vehicular network. For many Applications in IoV mobility control, multi-source data acquisition, distributed computation and storage, and multi-path data transmission Also described how several factors such as latency, mobility, localization, and scalability can be used for resource management
Cognitive Internet of Vehicles [46]	Used spatiotemporal data as steering, braking, number of sensors, cameras, lidar, radar, lane changing, acceleration, and overtaking other vehicles. For control services applications in IoV such as content distribution, traffic monitoring, and optimization, MAC coordination, data offloading Also described how strategic services can be used for decision makings as driver's behavior and pattern analysis, emotion analysis, driver and passenger's health monitoring, network resource allocation, and optimization Different challenges also have been addressed of using Big data in IoV such as data integrity, privacy breaching by insiders, data delivery delays and modification, data interoperability, authentication, secure localization, secure vehicular communications, data security, access control
Solving vehicular ad hoc network challenges with big data solutions [89]	Used vehicular heterogeneous data to identify challenges in big data applications based on several factors as performance, enhance travel security, traffic control, highly heterogeneous sources, efficiently and cost-effectively, inconsistent, incomplete, or noisy data Described several applications of Big data in IoV for real-time traffic monitoring, congestion-detection systems, crowdsourcing Prepared a comparative analysis of using machine learning techniques, such as text mining and information retrieval as complementary tools for making decisions
CarStream: an industrial system of big data processing for internet-of-vehicles [37]	Proposed a system by using radio frequency identification (RFID), electronic control unit (ECU), UCAR, etc. for big data processing Also have categorized factors of data quality (outlier analysis, data cleaning, data quality inspection, data filling), fleet management (electronic fence, vehicle tracking, fleet distribution, gasoline anomaly detection, fleet operation report, driving behavior analysis, driving event detection) decision making (driver profit assessment, order prediction, dynamic pricing) system monitoring, multi-stream fusion, trajectory compression
Information-centric networking for connected vehicles: a survey and future perspectives [62]	Proposed and comprehensive survey on future trends in IoV and connected vehicles How to manage safety-critical data, multicast support, and in-network data caching, information-centric networking, Interoperability in IoV is the major work in this paper
Big data-driven vehicular networks [128]	Proposed one architecture of IoV based on big data Also classified several data techniques such as vehicle sensing data, GPS data, self-driving related data: vehicular mobile service data and 5G technologies (eMBB, URLLC, mMTC) for Applications such as road safety/infotainment, intelligent transportation systems, and self-driving systems, mobile channels, pedestrians, terrain, and obstacles, Machine Learning with NB and SVM
Traffic congestion detection system through connected vehicles and big data [105]	Used SUMO in OMNet++ and Veines vehicular network for identifying the reasons for traffic congestion using big data And identified the methods for reducing the total emissions of CO2 and decreasing travel time
Internet of vehicles and cost-effective traffic signal control [84]	Proposed and cost-effective traffic signal control management solution by reducing high computing overhead, high installation and maintenance cost and, high susceptibility Prepared an experiment in the veins vehicular simulation framework to test the proposed solution

Table 12 (continued)

Study title	Key research contributions
A study on big data thinking of the internet of the things-based smart-connected car in conjunction with controller area network bus and 4g-long term evolution [33]	Prepared a comprehensive study of big data in IoT and connected vehicles and take Tesla autopilot(auto steer, auto park, driver assistance visualization) as a test case example. Used Microsoft Structured Query Language (MS-SQL), OpenXC, Google uses a self-developed file system (GFS) cloud services, elastic file System (EFS), Elastic Block Store (EBS), and Simple Storage Service (S3), Amazon Compute Cloud (Amazon EC2), Microsoft Azure, HDFS, Sqoop as the available framework
Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways [86]	Identified the traffic dynamics, the study has used data mining (random forest) and Bayesian inference technique Explored several factors as Indirect (peak hour, higher volume and lower speed upstream of crash locations) and direct (higher congestion index downstream to crash locations) congestion indicators
Estimating online vacancies in real-time road traffic monitoring with traffic sensor data stream [106]	Discussed spatial and temporal dimensions, processing stream data, trajectory data, traffic-monitoring, and process big data by using Hidden Markov Model (HMM), multiple linear regression. Prepared a prototype to test the architecture using Ubunto Server and Apache Storm
Enabling smart transportation systems: a parallel Spatio-temporal database approach [107]	Proposed parallel-distributed network-constrained moving objects database (PD-NMOD) architecture for transport management by using big data processing in the Spatio-temporal database. Identified several factors i.e., efficiency, reliability, and security, safety, and environmental sustainability, congestion avoidance, reducing greenhouse gas emissions, and effective traffic accident response, etc.
The real-time city? Big data and smart urbanism [129]	Prepared a comparative study to identify the importance of big data in smart cities Identified several factors i.e., economy and governance, efficient, sustainable, competitive, productive, open and transparent cities, city development, etc.
Big wave of the intelligent connected vehicles [93]	Consolidated the information of Connected vehicles and IoVs i.e.,self-backhaul ultra-dense networks (UDNs) combined with millimeter-wave (mmWave), immune optimization algorithm (IOA), navigation, traffic information service, remote vehicle condition diagnosis, driving safety service such as forward collision warning, vehicle control loss warning, vehicle to pedestrian collision warning
Breaking the Blockage for Big Data Transmission: Gigabit Road Communication in Autonomous Vehicles [108]	Proposed an autonomous relay algorithm called ATLR for blockage big data transmission in autonomous vehicles.
Vehicular content delivery: A big data perspective [72]	Prepared a survey for vehicular content transformation the data used and collected are vehicle status, user behaviors, and environmental features location-centric, user-centric, and vehicle-centric for comfort driving experiences and content-rich multimedia bases QoS.
Application of big data in intelligent traffic system [85]	Presented the survey of applications of big data in ITS and IoV Consolidate several factors i.e., customer information of railway, road traffic, aviation industry, public transit, are recorded, and tens of billion travel records, induction coil at bayonet point, an infrared detector, microwave detector, the ultrasonic detector, laser detector, video detector, GPS vehicle location, PGIS System, traffic guidance, evidence forensic, video surveillance
Big autonomous vehicular data classifications: towards procuring intelligence in ITS [48]	Consolidate ITS and big data solutions in telematics and real-time analysis, big data analytics, data fusion or knowledge discovery, distributed data storage, vehicular cloud, data classification, mathematical modeling Test and experiment in Hadoop for the proposed solution

Table 12 (continued)

Study title	Key research contributions
Vehicular fog computing: architecture, use case, and security and forensic challenges [96]	Described the summary of using vehicular fog computing networks and also presented the summary of challenges specifically in security and forensics in IoV Also have presented different security factors, for instance, communication efficiency, latency, location awareness, and real-time response Prepared a consolidated presentation for several challenges in traffic control, driving safety applications, entertainment services, urban warfare, and battlefields on the internet of battlefield things involving military vehicles, etc.
Connected vehicles: Solutions and challenges [109]	Developed a summary of challenges and solutions of IoV and also have summarized the challenges in multimedia, monitoring, control, emergency, security, audible text messages, Siri Voice Control How the world-known brands such as BMW (connected drive, mobile office) Audi (Audi Connect, Google earth), Ford (SYNC), Toyota (Touch 2), Volvo (Apple CarPlay), GM (OnStar) are using big data insights for optimal solutions
Big Data for transportation and mobility: recent advances, trends and challenges [92]	Identified the high volume, high velocity, and/or high variety, safer, cleaner, and more efficient big data properties in multiple and heterogeneous transportations/mobility applications for instance routing, planning, infrastructure monitoring, network design Described the Social Media (unstructured), sensor data (unstructured or semi-structured) and Open Data (either unstructured—raw text—, semi-structured—JSON/XML—or semantically structured—Linked Open Data). can be used in Smart transportation management
Big data for pedestrian volume: exploring the use of Google Street View images for pedestrian counts [110]	Used Google Street View, ACF Pedestrian detection, image segmentation by using several data i.e., corresponding image ID, the IDs of adjacent shooting locations, shooting location and other shooting parameters, Charged Coupled Device (CCD) sensor
Soft computing in big data intelligent transportation systems [111]	Proposed a solution based on the several big data processing systems i.e., Fuzzy control, Genetic algorithm, in EBoxII embedded devices, with the .NET Compact Framework, implemented in C#, the XML data format in SOAP web services for minimizing driving, or waiting time, route guidance system in IoV
Leveraging big data for the development of transport sustainability indicators [77]	Prepared a summary of heterogeneous data sources in IoV as RFID, smart cards, real-time, cellular, applications as Biketastic, Creek Watch, "Noise Meter, Social Media, e GPS, GSM, Wi-Fi, accelerometers, compasses, ambient light sensors, microphones, cameras, and gyroscopes, among other sensors and services in development include indoor navigation, air quality sensors, particulate matter sensing, thermometers, gas sensors, and others the growing availability of apps and services that can relate, link, and transmit data between agents and actors
Fog based intelligent transportation big data analytics on the Internet of Vehicles environment: motivations, architecture, challenges, and critical issues [78]	Presented the survey on big data and intelligent services IoV service providers to several categories as Lambda. Data acquisition: in sensors, actuators, mobile phones, tablets, vehicles, smart devices traffic guidance: navigation, emergency services, location-based services, (saving 160 USD). Vehicle user: lower insurance rates, lower operation costs, less time spent in traffic. Society: fewer crashes, fewer congestions, Less CO2 emissions, (saving 420 USD) Automobiles manufacturers: Lower services and, warranty costs, (saving 300 USD)
A secure mechanism for big data collection in large scale internet of vehicle [112]	Presented a security mechanism for big data in IoV Simulated a project-based to test this security feature
"Intelligent transportation system (ITS): concept, challenge, and opportunity" [113]	Presented different applications of ITS and IoVs for instance, traffic congestion, environmental impact, energy consumption, accidents and safety, high maintenance cost, land consumption. Also have presented the framework for advance transportation and management system (ATMS), advance travel information system (ATIS), advanced vehicle control system (AVCS)

Table 12 (continued)

Study title	Key research contributions
Video analytics towards vision zero [3]	Used traffic videos (e.g., detecting cars, pedestrians, and bikes for tracking their movements, near-collisions of cars with pedestrians and bikes or patterns of bikers crossing a busy intersection by using data mining, and deep learning methods Also, have discussed how this traffic management can be used for crowdsourcing
Urban human mobility data mining [38]	Provided data gathering strategies and methods as pervasive GPS, cellular network location for traffic congestion or air pollution problems, traffic congestion or air pollution problems, collection and cleaning and urban planning
Security and privacy on the Internet of Vehicles [114]	Presented the summary of security and privacy in IoV How to handle different attacks happened in IoV? Such as a Sybil attack. GPS deception, Masquerading attack, Wormhole attack Also represents the IoV security Models: Microsoft's STRIDE, SAODV, Ariadne and SRP, TESLA authentication scheme, SLPD, ALAR, and STAP
Big data analysis technology for electric vehicle networks in smart cities [130]	Used the K-means and fuzzy theory to prepare a simulation and model Described the impact of simulation, different factors on traffic conditions and under route guidance
Vehicle refueling behavior model based on spatio-temporal big data monitoring platform [131]	Identified the mobile gas refueling behavior of driver Prepared the spatio-temporal analysis for ecological environment management
Heterogeneous-Internet of Vehicles (Het-IoV) in twenty-first century: a comprehensive study [132]	A complete book for Het-IoV for comprehensive study of IoV different aspects. Enlisted and described the comprehensive analysis of using big data analysis and processing
Joint 3D reconstruction and object tracking for traffic video analysis under IoV environment [133]	Presented the deep analysis for object detection and 3D reconstruction Prepared the simulated trajectory for augmented reality based analsis Video analysis for driver navigation behavior
A collaborative scheduling strategy for IoV computing resources considering location privacy protection in mobile edge computing environment [134]	Presented a complete solution for resource management in IoV Used Deep learning to identify communication delay and privacy of vehicles Also have presented the energy consumption management in IoV
Investigation on deep learning methods for privacy and security challenges of cognitive IoV [135]	Presented a framework for privacy and security management in CIoV Prepared a comparison between several different deep learning methods for process management in IoV Described several key challenges of using deep learning in IoV
Accessibility analysis and modeling for IoV in an urban scene [136]	Presented different phases of accessibility of data among interconnected nodes in IoV Simulated one experiment for testing the navigational behavior of several types of data for optimization of resources such as data packet delivery rate
Evolutionary V2X technologies toward the Internet of Vehicles: challenges and opportunities [137]	Prepared a comprehensive analysis of using V2X technologies for optimal solution Also have presented the challenges and opportunities of evolutionary V2X in IoV
Cyber physical and social networks in IoV (CPSN-IoV): a multimodal architecture in edge-based networks for optimal route selection using 5G technologies [138]	Prepared a comprehensive analysis of using CPSN in IoV Simulated one complete example for navigational pattern in IoV Also have presented the challenges and opportunities and, future directions of IoV
AI, blockchain, and vehicular edge computing for smart and secure IoV: challenges and directions [139]	Presented a comprehensive summary of using smart and secure IoV operations Described the key challenges and future direction of using IoV by using edge computation
Performance comparison of messaging protocols and serialization formats for digital twins in IoV [140]	Prepared a comprehensive summary of several messaging protocols Constrained application protocol (CoAP), advanced message queuing protocol (AMQP), and message queuing telemetry transport (MQTT) are being compared using several analytical schemes

Table 12 (continued)

Study title	Key research contributions
Software-defined networking (SDN) and edge-computing-aided IoV (EC-SDIoV) [141]	Presented the simple summary of using software defined networks by using edge based computing methods Also have described the complete solutions for using SDN for IoV
Smart route: internet-of-vehicles (IoV)-based congestion detection and avoidance (IoV-based CDA) using rerouting planning [142]	Simulated an experiment for using route management and optimal resource management Used different IoV based congestion detection, prevention and avoidance methods
Internet of vehicles: key technologies, network model, solutions and challenges with future aspects [143]	Provided the comprehensive study of all the key areas of IoV including architecture, frameworks, solutions and challenges. Presented the future directions and solutions of using IoV with big data solutions
Reliable computation offloading for edge-computing-enabled software-defined IoV [144]	Integrated the mobile-edge based computing solution with IoV using extensive processing solutions Prepared the comparison summary for mobile and edge based computing nodes and how these nodes can work optimum by using big data
Connected vehicles in the IoV: concepts, technologies and architectures [145]	Presented the comprehensive study of all the key areas of IoV and connected vehicles Summarized the future directions and solutions of using IoV with big data solutions Prepared a consolidated summary for architecture, frameworks, solutions and challenges in IoV and connected vehicles
Internet of vehicles [scanning the issue] [146]	Prepared a comprehensive study for vehicle as intelligent object and equipped with sensing Also have presented the key issues of IoV and big data
Internet of vehicles, vehicular social networks, and cybersecurity [147]	Discussed and presented layered processing architecture of IoV Presented the solution that how big data acquisition and processing can be done in heterogeneous environment
Road quality analysis based on cognitive Internet of Vehicles (CIoV) [148]	Presented the big data acquisition methods and how these solutions can be used for identifying the road quality Provided the solutions for QoS based services for optimal resource management

Table 13 Sensory observation

Name	Category	Range	Type	Application
Vehicular sensors				
Radar sensors	Distance sensors	30–200 m	Real-time	Cruise control (Adaptive), blind spot identification, alert messages
Laser scanners (LIDAR)	Distance sensors	20–200 m	Real-time	Laser, photo-detector, scanner, optics, navigation and positioning system
Ultrasonic sensors	Distance sensors	3 cm–5 m	Real-time	Transducer, reflector, amplifier, detector-compactor
Capacities proximity sensors	Distance sensors	2 m	Real-time	E-field propagation, firmware filtering
Camera vision	Distance sensors	145 dB (dynamic)	Real-time	CMOS image sensors
Night vision sensors	Night vision sensors	Wavelength: up to 14 µm	Real-time	Instrument cluster-by using LCD, navigation system, windshield via a head-up display
Speed sensors	Speed sensors	Frequency: to above 100 kHz	Real-time	Transmission speed sensors, ABS speed sensors
Temperature sensor	Temperature sensor	− 40 to + 125 °C	Real-time	Antilock braking, capacities pressure-sensor signal conditioning, humidity sensing, signal conditioning, fuel pressure
Sensors-barometric pressure	Temperature sensor	600–1100 hPa	Real-time	Weather monitoring & forecasting, evapotranspiration, wind resource assessment, environmental monitoring
Accelerometers	Acceleration inertial sensors	± 1–± 250 g	Real-time	1D, 2D, and 3D based technology, used in the military, aerospace market and consumer electronics and automotive market
Gyroscopes	Acceleration inertial sensors	± 250°/s, ± 500°/s or ± 2000°/s	Real-time	used in space navigation, missile control, under-water guidance, and flight guidance, modern applications are motion-capture and vehicle navigation
People on board				
GPS	Positioning sensor	4.9 m (16 ft.) radius	Real-time	Latitude and longitude of cellular phone and other GPS enabled devices
Accelerometer	Acceleration inertial sensors	± 1 g up to ± 250 g	Real-time	Detect acceleration, vibration, and tilt of the cellular phone
Gyroscope	Acceleration inertial sensors	360 degrees per second	Real-Time	measure rotation, orientation and direction
Magnetometer	Acceleration inertial sensors	294 mAh	Real-time	detect magnetic fields
Proximity sensor	Distance sensors	2.5 m	Real-time	Infrared LED and IR light detector to find out how close the phone is to an outside object
Social media feeds	Event detector and situation awareness	Based on geolocation and event	Real-time, archive	Crowdsourcing, event detection, information disseminator
Fixed sensory devices				
Smart surveillance camera	Imagery surveillance	Based on geolocation and event	Real-time, archive	Event detection, crowdsourcing, real-time

Table 13 (continued)

Name	Category	Range	Type	Application
Smart meters sensors	Sensory observation	Based on geolocation and event	Real-time, archive	Smart metering, automatic meter reading
Air pollution sensors	Sensory observation	Based on geolocation and event	Real-time, archive	Detection of air pollution by different means in smart cities
Road side units (RSU)	Sensory observation and imagery surveillance	Vehicle coverage = 200–10,000 Velocity = 30–100 km/h	Real-time, archive	Use to acquire information regarding vehicle and traffic and generate alert messages
Infrared detectors	Sensory observation	4–15 feet	Real-time	For gathering pedestrians and cyclist navigational pattern
Ultrasonic sensors	Sensory observation	2–400 cm or 1–13 feet	Real-time	Use to acquire information about car parking management

Acknowledgements We are thankful to the anonymous reviewers whose valuable and constructive review comments helped us improving this manuscript.

Author Contributions Initial identification, preparation of research statement and research questions: AAN, AAK, MSF. Evaluate and selection of literature: AAN, MSF, RI, AS. Organization of literature: AAN, AAK, RI, AS. Wrote the paper: AAN, AAK. Revised the manuscript: AAN, AAK, MSF,

Funding The authors have no support or funding to report.

Declarations

Conflict of interest On the behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Kemp G, Vargas-Solar G, Da Silva CF, Ghodous P (2015) Aggregating and managing big realtime data in the cloud-application to intelligent transport for smart cities. In: Proceedings of the 1st international conference on vehicle technology and intelligent transport systems, pp 107–112
2. Lv Y, Duan Y, Kang W, Li Z, Wang F (2015) Traffic flow prediction with big data: a deep learning approach. IEEE Trans Intell Transp Syst 16(2):865–873. https://doi.org/10.1109/TITS.2014.2345663
3. Loewenherz F, Bahl V, Wang Y (2017) Video analytics towards vision zero. Inst Transp Eng ITE J 87(3):25
4. Sherazi HHR, Khan ZA, Iqbal R, Rizwan S, Imran MA, Awan K (2019) A heterogeneous IoV architecture for data forwarding in vehicle to infrastructure communication. Mob Inf Syst 2019:3101276. https://doi.org/10.1155/2019/3101276
5. Sun Y, Song H, Jara AJ, Bie R (2016) Internet of things and big data analytics for smart and connected communities. IEEE Access 4:766–773. https://doi.org/10.1109/ACCESS.2016.2529723
6. Kadam VG, Tamane SC, Solanki VK (2019) Smart and connected cities through technologies. In: Dey N, Tamane S (eds) Big data analytics for smart and connected cities. IGI Global, pp 1–24
7. Dobre C, Xhafa F (2014) Intelligent services for big data science. Fut Gener Comput Syst 37:267–281. https://doi.org/10.1016/j.future.2013.07.014
8. Nguyen DB, Dow CR, Hwang SF (2018) An efficient traffic congestion monitoring system on Internet of Vehicles. Wirel Commun Mob Comput 2018:9136813. https://doi.org/10.1155/2018/9136813
9. Al Nuaimi E, Al Neyadi H, Mohamed N, Al-Jaroodi J (2015) Applications of big data to smart cities. J Internet Serv Appl 6(1):25. https://doi.org/10.1186/s13174-015-0041-5
10. Chen M, Mao S, Liu Y (2014) Big data: a survey. Mob Netw Appl 19(2):171–209. https://doi.org/10.1007/s11036-013-0489-0
11. Yuan W, Deng P, Taleb T, Wan J, Bi C (2013) An unlicensed taxi identification model based on big data analysis. IEEE Trans Intell Transp Syst 17(6):1703–1713. https://doi.org/10.1109/TITS.2015.2498180
12. Artikis A, Weidlich M, Gal A, Kalogeraki V, Gunopulos D (2013) Self-adaptive event recognition for intelligent transport management. IEEE international conference on big data 2013:319–325
13. Karkouch A, Mousannif H, Al Moatassime H, Noel T (2016) Data quality in internet of things: a state-of-the-art survey. J Netw Comput Appl 73:57–81. https://doi.org/10.1016/j.jnca.2016.08.002
14. Warren SD, Brandeis LD (1890) The right to privacy. Harv Law Rev 4:193–220
15. Shi W, Dustdar S (2016) The promise of edge computing. Computer 49(5):78–81
16. Edge computing: vision and challenges—IEEE J Mag. https://ieeexplore.ieee.org/document/7488250. Accessed 18 Sept 2019
17. Ghofrani F, He Q, Goverde RM, Liu X (2018) Recent applications of big data analytics in railway transportation systems: a survey. Transp Res C Emerg Technol 90:226–246
18. Zhu L, Yu FR, Wang Y, Ning B, Tang T (2018) Big data analytics in intelligent transportation systems: a survey. IEEE Trans Intell Transp Syst 20(1):383–398
19. Neilson A, Daniel B, Tjandra S et al (2019) Systematic review of the literature on big data in the transportation domain: concepts and applications. Big Data Res 17:35–44
20. Amy-Mae T (2011) The future of the connected car https://mashable.com/2011/02/26/connected-car/. Accessed 18 Sept 2019

21. Definition of Connected Car – What is the connected car? Defined. auto connected car news, Apr. 22, 2014. https://www.autoconnectedcar.com/definition-of-connected-car-what-is-the-connected-car-defined/. Accessed 18 Sept 2019
22. Stratio Automotive. https://stratioautomotive.com/. Accessed 19 Sept 2019
23. Milakis D, Snelder M, Arem B, Wee B, de Almeida Homem CG (2017) Development and transport implications of automated vehicles in the Netherlands: scenarios for 2030 and 2050. Eur J Transp Infrastruct Res 17:63–85
24. Cochrane methods—twenty years experience in developing systematic review methods | Systematic Reviews | Full Text'. https://systematicreviewsjournal.biomedcentral.com/articles/10.1186/2046-4053-2-76. Accessed 19 Sept 2019
25. Hype cycle for emerging technologies, 2012, Gartner. https://www.gartner.com/en/documents/2100915/hype-cycle-for-emerging-technologies-2012. Accessed 23 Sept 2019
26. Volk M, Bosse S, Turowski K (2017) Providing clarity on big data technologies: a structured literature review. In: 2017 IEEE 19th conference on business informatics (CBI), vol 1, pp 388–397
27. Pautasso M (2013) Ten simple rules for writing a literature review. PLoS Comput Biol 9(7):e1003149. https://doi.org/10.1371/journal.pcbi.1003149
28. Fernandez A, Insfran E, Abrahão S (2011) Usability evaluation methods for the web: a systematic mapping study. Inf Softw Technol 53(8):789–817
29. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33:159–174
30. CORE Rankings Portal—computing research & education. http://core.edu.au/conference-portal. Accessed 30 Sept 2019
31. Dey N, Tamane S (2018) Big data analytics for smart and connected cities. IGI Global, Hershey
32. McQueen B (2017) Big data analytics for connected vehicles and smart cities. Artech House, London
33. Kwon D, Park S, Ryu J-T (2017) A study on big data thinking of the internet of things-based smart-connected car in conjunction with controller area network bus and 4g-long term evolution. Symmetry 9(8):152
34. Costa C, Chatzimilioudis G, Zeinalipour-Yazti D, Mokbel MF (2017) Towards real-time road traffic analytics using telco big data. In: Proceedings of the international workshop on real-time business intelligence and analytics, p 5
35. Ta-Shma P, Akbar A, Gerson-Golan G, Hadash G, Carrez F, Moessner K (2018) An ingestion and analytics architecture for iot applied to smart city use cases. IEEE Internet Things J 5(2):765–774. https://doi.org/10.1109/JIOT.2017.2722378
36. Yuan W, Deng P, Taleb T, Wan J, Bi C (2015) An unlicensed taxi identification model based on big data analysis. IEEE Trans Intell Transp Syst 17(6):1703–1713
37. Zhang M, Wo T, Xie T, Lin X, Liu Y (2017) CarStream: an industrial system of big data processing for internet-of-vehicles. Proc VLDB Endow 10(12):1766–1777. https://doi.org/10.14778/3137765.3137781
38. Zhao K, Tarkoma S, Liu S, Vo H (2016) Urban human mobility data mining: an overview. IEEE international conference on big data (big data) 2016:1911–1920
39. Guo B et al (2015) Mobile crowd sensing and computing: the review of an emerging human-powered sensing paradigm. ACM Comput Surv. CSUR 48(1):7. https://doi.org/10.1145/2794400
40. Itoh M, Yokoyama D, Toyoda M, Tomita Y, Kawamura S, Kitsuregawa M (2014) Visual fusion of mega-city big data: an application to traffic and tweets data analysis of metro passengers. IEEE international conference on big data (big data) 2014:431–440
41. Attanasi A, Silvestri E, Meschini P, Gentile G (2015) Real world applications using parallel computing techniques in dynamic traffic assignment and shortest path search. In: 2015 IEEE 18th international conference on intelligent transportation systems, pp 316–321
42. Baron B, Spathis P, Rivano H, de Amorim MD (2014) Vehicles as big data carriers: road map space reduction and efficient data assignment. In: 2014 IEEE 80th vehicular technology conference (VTC2014-Fall), pp 1–5
43. Costanzo A, Faro A, Giordano D, Spampinato C (2016) An ontological ubiquitous city information platform provided with cyber-physical-social-systems, pp 137–144, https://doi.org/10.1109/CCNC.2016.7444746
44. Big data analytics architecture for real-time traffic control. In: 2017 5th IEEE international conference on models and technologies for intelligent transportation systems (MT-ITS), pp 710–715 (2017)
45. Schatzinger S, Lim CYR (2015) Taxi of the future: big data analysis as a framework for future urban fleets in smart cities. In: International conference on smart and sustainable planning for cities and regions, pp 83–98
46. Chen M, Tian Y, Fortino G, Zhang J, Humar I (2018) Cognitive Internet of Vehicles. Comput Commun 120:58–70. https://doi.org/10.1016/j.comcom.2018.02.006
47. Intel to invest $250 million in autonomous driving—DriverlessGuru.com—your guide to the world of self-driving cars. https://www.driverlessguru.com/blog/intel-to-invest-250-million-into-autonomous-driving. Accessed 24 Oct 2019
48. Daniel A, Subburathinam K, Paul A, Rajkumar N, Rho S (2017) Big autonomous vehicular data classifications: towards procuring intelligence in ITS. Veh Commun 9:306–312
49. St-Aubin P, Saunier N, Miranda-Moreno L (2015) Large-scale automated proactive road safety analysis using video data. Transp Res C Emerg Technol 58:363–379
50. Aleyadeh S, Oteafy S, Hassanein HS (2015) Scalable transportation monitoring using the smartphone road monitoring (SRoM) system. In: Proceedings of the 5th ACM symposium on development and analysis of intelligent vehicular networks and applications, pp 43–50
51. Treboux J, Jara AJ, Dufour L, Genoud D (2015) A predictive data-driven model for traffic-jams forecasting in smart santader city-scale testbed. IEEE wireless communications and networking conference workshops (WCNCW) 2015:64–68
52. Serrano D, Baldassarre T, Stroulia E (2016) Real-time traffic-based routing, based on open data and open-source software. In: 2016 IEEE 3rd world forum on internet of things (WF-IoT), pp 661–665
53. Fusco G, Colombaroni C, Comelli L, Isaenko N (2015). Short-term traffic predictions on large urban traffic networks: applications of network-based machine learning models and dynamic traffic assignment models. https://doi.org/10.1109/MTITS.2015.7223242
54. Golubev A, Chechetkin I, Solnushkin KS, Sadovnikova N, Parygin D, Shcherbakov M (2015) Strategway: web solutions for building public transportation routes using big geodata analysis. In: Proceedings of the 17th international conference on information integration and web-based applications & services, p 91
55. Abdullah E, Emam A (2015) Traffic accidents analyzer using big data. International conference on computational science and computational intelligence (CSCI) 2015:392–397
56. Rizwan P, Suresh K, Babu MR (2016) Real-time smart traffic management system for smart cities by using Internet of Things and big data. In: 2016 International conference on emerging technological trends (ICETT), pp 1–7. https://doi.org/10.1109/ICETT.2016.7873660

57. Liu Y, Chen X, Chen C, Guan X (2016) Traffic big data analysis supporting vehicular network access recommendation. In: 2016 IEEE international conference on communications (ICC), pp 1–6. https://doi.org/10.1109/ICC.2016.7510775
58. Xu W et al (2018) Internet of vehicles in big data era. IEEE-CAA J Autom Sin 5(1):19–35. https://doi.org/10.1109/JAS.2017.7510736
59. The most popular database for modern apps. MongoDB. https://www.mongodb.com. Accessed 10 July 2019
60. Apache Cassandra. http://cassandra.apache.org/. Accessed 10 July 2019
61. Apache HBase—Apache HBaseTM Home. https://hbase.apache.org/. Accessed 10 July 2019
62. Amadeo M, Campolo C, Molinaro A (2016) Information-centric networking for connected vehicles: a survey and future perspectives. IEEE Commun Mag 54(2):98–104. https://doi.org/10.1109/MCOM.2016.7402268
63. (2018) Networks. IEEE Netw 32(6), 160–167. https://doi.org/10.1109/MNET.2018.1700460
64. Welcome to Apache Flume—Apache Flume. https://flume.apache.org/. Accessed 11 July 2019
65. Apache Flink: stateful computations over data streams. https://flink.apache.org/. Accessed 11 July 2019
66. Samza. http://samza.apache.org/. Accessed 11 July 2019
67. Apache Storm. https://storm.apache.org/. Accessed 11 July 2019
68. Liu D (2018) Big data analytics architecture for internet-of-vehicles based on the spark. In: 2018 International conference on intelligent transportation, big data & smart city (ICITBS), pp 13–16
69. Apache Hadoop. https://hadoop.apache.org/. Accessed 11 July 2019
70. Welcome to Apache Pig! https://pig.apache.org/. Accessed 11 July 2019
71. Apache Hive TM. https://hive.apache.org/. Accessed 11 July 2019
72. Xu C, Zhou Z (2018) Vehicular content delivery: a big data perspective. IEEE Wirel Commun 25(1):90–97
73. Chilipirea C, Petre A-C, Groza L-M, Dobre C, Pop F (2017) An integrated architecture for future studies in data processing for smart cities. Microprocess Microsyst 52:335–342
74. Apache SparkTM - unified analytics engine for big data. https://spark.apache.org/. Accessed 11 July 2019
75. Gebremeskel GB, Chai Y, Yang Z (2014) The paradigm of big data for augmenting internet of vehicle into the intelligent cloud computing systems. In: Internet of vehicles—technologies and services. Cham 247–261
76. Ahmed M, Kostakos V (2016) Cyclist-aware traffic lights through distributed smartphone sensing. Pervasive Mob Comput 31:22–36. https://doi.org/10.1016/j.pmcj.2016.01.012
77. Cottrill CD, Derrible S (2015) Leveraging big data for the development of transport sustainability indicators. J Urban Technol 22(1):45–64
78. Darwish TSJ, Bakar KA (2018) Fog based intelligent transportation big data analytics in the Internet of Vehicles environment: motivations, architecture, challenges, and critical issues. IEEE Access 6:15679–15701. https://doi.org/10.1109/ACCESS.2018.2815989
79. Arooj A, Farooq MS, Umer T, Shan RU (2019) Cognitive Internet of Vehicles and disaster management: a proposed architecture and future direction. Trans Emerg Telecommun Technol. https://doi.org/10.1002/ett.3625
80. Zhu M, Liu X-Y, Qiu M, Shen R, Shu W, Wu M-Y (2016) Traffic big data based path planning strategy in public vehicle systems. In: 2016 IEEE/ACM 24th international symposium on quality of service (IWQoS), pp 1–2
81. Bell MG, Trozzi V, Hosseinloo SH, Gentile G, Fonzone A (2012) Time-dependent Hyperstar algorithm for robust vehicle navigation. Transp Res Policy Pract 46(5):790–800
82. Al Najada H, Mahgoub I (2016) Autonomous vehicles safe-optimal trajectory selection based on big data analysis and predefined user preferences. In: 2016 IEEE 7th annual ubiquitous computing. Electronics & mobile communication conference (uEMCON), pp 1–6
83. Aris IB, Sahbusdin RKZ, Amin AFM (2015) Impacts of IoT and big data to automotive industry. In: 2015 10th Asian control conference (ASCC), pp 1–5
84. Ahn S, Choi J (2019) Internet of vehicles and cost-effective traffic signal control. Sensors 19(6):1275
85. Zeng G (2015) Application of big data in intelligent traffic system. IOSR J Comput Eng 17(1):01–04
86. Shi Q, Abdel-Aty M (2015) Big data applications in real-time traffic operation and safety monitoring and improvement on urban expressways. Transp Res C Emerg Technol 58:380–394
87. Ali et al RY (2015) Future connected vehicles: challenges and opportunities for spatio-temporal computing. In: Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems, p 14
88. Yang J, Han Y, Wang Y, Jiang B, Lv Z, Song H (2017) Optimization of real-time traffic network assignment based on IoT data using DBN and clustering model in smart city. Fut Gener Comput Syst. https://doi.org/10.1016/j.future.2017.12.012
89. Contreras-Castillo J, Zeadally S, Ibañez JAG (2016) Solving vehicular ad hoc network challenges with big data solutions. IET Netw 5(4):81–84
90. Zhou Z, Gao C, Xu C, Zhang Y, Mumtaz S, Rodriguez J (2018) Social big-data-based content dissemination in Internet of Vehicles. IEEE Trans Ind Inform 14(2):768–777. https://doi.org/10.1109/TII.2017.2733001
91. Pandey MK, Subbiah K (2016) Social networking and big data analytics assisted reliable recommendation system model for Internet of Vehicles. In: International conference on Internet of Vehicles, pp 149–163
92. Torre-Bastida AI, Del Ser J, Laña I, Ilardia M, Bilbao MN, Campos-Cordobés S (2018) Big data for transportation and mobility: recent advances, trends and challenges. IET Intell Transp Syst 12(8):742–755
93. Li Y, Cao Y, Qiu H, Gao L, Du Z, Chen S (2016) Big wave of the intelligent connected vehicles. China Commun 13(2):27–41
94. Ayed AB, Halima MB, Alimi AM (2015) Big data analytics for logistics and transportation. In: 2015 4th international conference on advanced logistics and transport (ICALT), pp 311–316
95. Banikhalaf M, Manasrah AM, AlEroud AF, Hamadneh N, Qawasmeh A, Al-Dubai AY (2019) A reliable route repairing scheme for Internet of Vehicles. Int J Comput Appl Technol 61(3):229–238
96. Huang C, Lu R, Choo K-KR (2017) Vehicular fog computing: architecture, use case, and security and forensic challenges. IEEE Commun Mag 55(11):105–111
97. Nkenyereye L, Jang J-W (2015) A study of big data solution using hadoop to process connected vehicle's diagnostics data. In: Information science and applications. Springer, pp 697–704
98. Datta SK, Costa RPFD, Härri J, Bonnet C (Jun. 2016) Integrating connected vehicles in Internet of Things ecosystems: challenges and solutions. In: 2016 IEEE 17th international symposium on a world of wireless, mobile and multimedia networks (WoWMoM), pp 1–6. https://doi.org/10.1109/WoWMoM.2016.7523574
99. Tbatou S, Ramrami A, Tabii Y (2017) Security of communications in connected cars modeling and safety assessment. In: Proceedings of the 2nd international conference on big data, cloud and applications, p 56

100. Xu W et al (2016) Supporting large scale connected vehicle data analysis using HIVE. IEEE international conference on big data (big data) 2016:2296–2304
101. Shukla S, Balachandran K, Sumitha V (2016) A framework for smart transportation using big data. International conference on ICT in business industry & government (ICTBIG) 2016:1–3
102. Amato G, Carrara F, Falchi F, Gennaro C, Meghini C, Vairo C (2017) Deep learning for decentralized parking lot occupancy detection. Expert Syst Appl 72:327–334. https://doi.org/10.1016/j.eswa.2016.10.055
103. Kumar N, Misra S, Rodrigues JJPC, Obaidat MS (2015) Coalition games for spatio-temporal big data in Internet of Vehicles environment: a comparative analysis. IEEE Internet Things J 2(4):310–320. https://doi.org/10.1109/JIOT.2015.2388588
104. Zhang W, Zhang Z, Chao H (2017) Cooperative fog computing for dealing with big data in the Internet of Vehicles: architecture and hierarchical resource management. IEEE Commun Mag 55(12):60–67. https://doi.org/10.1109/MCOM.2017.1700208
105. Cárdenas-Benńtez N, Aquino-Santos R, Magaña-Espinoza P, Aguilar-Velazco J, Edwards-Block A, Medina Cass A (2016) Traffic congestion detection system through connected vehicles and big data. Sensors 16(5):599
106. Wang F, Hu L, Zhou D, Sun R, Hu J, Zhao K (2015) Estimating online vacancies in real-time road traffic monitoring with traffic sensor data stream. Ad Hoc Netw 35:3–13
107. Ding Z, Yang B, Chi Y, Guo L (2015) Enabling smart transportation systems: A parallel spatio-temporal database approach. IEEE Trans Comput 65(5):1377–1391
108. Zhang C, Ota K, Jia J, Dong M (2018) Breaking the blockage for big data transmission: gigabit road communication in autonomous vehicles. IEEE Commun Mag 56(6):152–157
109. Lu N, Cheng N, Zhang N, Shen X, Mark JW (2014) Connected vehicles: solutions and challenges. IEEE Internet Things J 1(4):289–299
110. Yin L, Cheng Q, Wang Z, Shao Z (2015) Big data'for pedestrian volume: exploring the use of google street view images for pedestrian counts. Appl Geogr 63:337–345
111. Wang C, Li X, Zhou X, Wang A, Nedjah N (2016) Soft computing in big data intelligent transportation systems. Appl Soft Comput 38:1099–1108
112. Guo L et al (2017) A secure mechanism for big data collection in large scale internet of vehicle. IEEE Internet Things J 4(2):601–610. https://doi.org/10.1109/JIOT.2017.2686451
113. Lin Y, Wang P, Ma M (2017) Intelligent transportation system (ITS): concept, challenge and opportunity. In: 2017 IEEE 3rd international conference on big data security on cloud (bigdatasecurity), IEEE international conference on high performance and smart computing (HPSC), and IEEE international conference on intelligent data and security (IDS), pp 167–172
114. Sun et al Y (2015) Security and privacy in the Internet of Vehicles, pp 116–121. https://doi.org/10.1109/IIKI.2015.33
115. McGuinness DL, Van Harmelen F et al (2004) OWL web ontology language overview. W3C Recomm 10(10):2004
116. Darji P (2015) Why connected cars are poised to have a fundamental impact. Diginomica. https://diginomica.com/why-connected-cars-are-poised-to-have-a-fundamental-impact. Accessed 20 Nov 2019
117. Chen W, Guha RK, Kwon TJ, Lee J, Hsu Y-Y (2011) A survey and challenges in routing and data dissemination in vehicular ad hoc networks. Wirel Commun Mob Comput 11(7):787–795
118. Mavromoustakis CX, Mastorakis G, Dobre C (2017) Advances in mobile cloud computing and big data in the 5G era. Springer, Berlin
119. Symantec—Global leader in next-generation cyber security | Symantec. Accessed 25 Nov 2019. https://www.symantec.com/
120. Butt TA, Iqbal R, Salah K, Aloqaily M, Jararweh Y (2019) Privacy management in social Internet of Vehicles: review, challenges and blockchain based solutions. IEEE Access 7:79694–79713
121. Iqbal R, Butt TA, Afzaal M, Salah K (2019) Trust management in social Internet of vehicles: factors, challenges, blockchain, and fog solutions. Int J Distrib Sens Netw 15(1):1550147719825820
122. Intelligent Transportation Systems—research—ITS strategic plan 2015–2019'. https://www.its.dot.gov/research_areas/strategicplan2015.htm. Accessed 10 May 2020
123. Sae S (2016) j3061, cybersecurity guidebook for cyber-physical vehicle systems. Nr 1:52
124. lynn.greenbauer.ctr@dot.gov, enforcement & justice services. NHTSA (2018). https://www.nhtsa.gov/enforcement-justice-services. Accessed 10 May 2020
125. Committee and others ITS (2013) IEEE standard for wireless access in vehicular environments–security services for applications and management messages. IEEE Std
126. Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (3GPP TS 36.101 version 14.3.0 Release 14)
127. Kitchenham B, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering. Citeseer
128. Cheng N et al (2018) Big data driven vehicular networks. IEEE Netw 32(6):160–167. https://doi.org/10.1109/MNET.2018.1700460
129. Kitchin R (2014) The real-time city? Big data and smart urbanism. GeoJ 79(1):1–14
130. Lv Z, Qiao L, Cai K, Wang Q (2020) Big data analysis technology for electric vehicle networks in smart cities. IEEE Trans Intell Transp Syst 22:1807–1816
131. Wang W (2020) Vehicle refueling behavior model based on spatio-temporal big data monitoring platform. In: International conference on application of intelligent systems in multi-modal information analytics, pp 708–713
132. Richa TP, Sharma Sharma AK (2020) Heterogeneous-Internet of Vehicles (Het-IoV) in twenty-first century: a comprehensive study. In: Gupta BB, Perez GM, Agrawal DP, Gupta D (eds) Handbook of computer networks and cyber security: principles and paradigms. Springer, Cham, pp 555–584
133. Cao M, Zheng L, Jia W, Liu X (2020) Joint 3D reconstruction and object tracking for traffic video analysis under IoV environment. IEEE Trans Intell Transp Syst. https://doi.org/10.1109/TITS.2020.2995768
134. Pang M, Wang L, Fang N (2020) A collaborative scheduling strategy for IoV computing resources considering location privacy protection in mobile edge computing environment. J Cloud Comput 9(1):52. https://doi.org/10.1186/s13677-020-00201-x
135. Sassi MSH, Fourati LC (Jun. 2020) Investigation on deep learning methods for privacy and security challenges of cognitive IoV. In: 2020 International wireless communications and mobile computing (IWCMC), pp 714–720. https://doi.org/10.1109/IWCMC48107.2020.9148417
136. Cheng J et al (2020) Accessibility analysis and modeling for IoV in an urban scene. IEEE Trans Veh Technol 69(4):4246–4256. https://doi.org/10.1109/TVT.2020.2970553
137. Zhou H, Xu W, Chen J, Wang W (2020) Evolutionary V2X technologies toward the Internet of Vehicles: challenges and opportunities. Proc IEEE 108(2):308–323. https://doi.org/10.1109/JPROC.2019.2961937
138. Arooj A, Farooq MS, Umer T, Rasool G, Wang B (2020) Cyber physical and social networks in IoV (CPSN-IoV): a multimodal architecture in edge-based networks for optimal route selection using 5G technologies. IEEE Access 8:33609–33630
139. Hammoud A, Sami H, Mourad A, Otrok H, Mizouni R, Bentahar J (2020) AI, blockchain, and vehicular edge computing for smart

and secure IoV: challenges and directions. IEEE Internet Things Mag 3(2):68–73. https://doi.org/10.1109/IOTM.0001.1900109
140. Proos DP, Carlsson N (2020) Performance comparison of messaging protocols and serialization formats for digital twins in IoV. In: 2020 IFIP networking conference (networking), pp 10–18
141. Hou X et al (2020) Reliable computation offloading for edge-computing-enabled software-defined IoV. IEEE Internet Things 7(8):7097–7111. https://doi.org/10.1109/JIOT.2020.2982292
142. Khan Z, Koubaa A, Farman H (2020) Smart route: internet-of-vehicles (IoV)-based congestion detection and avoidance (IoV-based CDA) using rerouting planning. Appl Sci. https://doi.org/10.3390/app10134541
143. Qureshi KN, Din S, Jeon G, Piccialli F (2020) Internet of vehicles: key technologies, network model, solutions and challenges with future aspects. IEEE Trans Intell Transp Syst. https://doi.org/10.1109/TITS.2020.2994972
144. Dutta AK, Elhoseny M, Dahiya V, Shankar K (2020) An efficient hierarchical clustering protocol for multihop Internet of Vehicles communication. Trans Emerg Telecommun Technol 31(5):e3690. https://doi.org/10.1002/ett.3690
145. Mahmood Z (2020) Connected vehicles in the IoV: concepts, technologies and architectures. In: Mahmood Z (ed) Connected vehicles in the internet of things: concepts, technologies and frameworks for the IoV. Springer, Cham, pp 3–18
146. Shen X, Fantacci R, Chen S (2020) Internet of vehicles [scanning the issue]. Proc IEEE 108(2):242–245. https://doi.org/10.1109/JPROC.2020.2964107
147. Kim S, Shrestha R (2020) Internet of vehicles, vehicular social networks, and cybersecurity. In: Kim S, Shrestha R (eds) Automotive cyber security: introduction, challenges, and standardization. Springer, Singapore, pp 149–181
148. Rahimi H, Dharmalingam D (2020) Road quality analysis based on cognitive Internet of Vehicles (CIoV). ArXiv200409287 Cs. Accessed 13 Oct 2020. arXiv:2004.09287