



A Review of Crowdsourcing Update Methods for High-Definition Maps

Yuan Guo ¹, Jian Zhou ²,*⁰, Xicheng Li ³, Youchen Tang ² and Zhicheng Lv ⁴

- School of Resource and Environmental Sciences, Wuhan University, Wuhan 430205, China; guoyuan@whu.edu.cn
- ² State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University, Wuhan 430205, China; youchentang@whu.edu.cn
- ³ School of Geodesy and Geomatics, Wuhan University, Wuhan 430070, China
- ⁴ School of Information and Communication Engineering, Hubei University of Economics, Wuhan 430000, China; lion112@139.com
- * Correspondence: jianzhou@whu.edu.cn

Abstract: High-definition (HD) maps serve as crucial infrastructure for autonomous driving technology, facilitating vehicles in positioning, environmental perception, and motion planning without being affected by weather changes or sensor-visibility limitations. Maintaining precision and freshness in HD maps is paramount, as delayed or inaccurate information can significantly impact the safety of autonomous vehicles. Utilizing crowdsourced data for HD map updating is widely recognized as a superior method for preserving map accuracy and freshness. Although it has garnered considerable attention from researchers, there remains a lack of comprehensive exploration into the entire process of updating HD maps through crowdsourcing. For this reason, it is imperative to review and discuss crowdsourcing techniques. This paper aims to provide an overview of the overall process of crowdsourced updates, followed by a detailed examination and comparison of existing methodologies concerning the key techniques of data collection, information extraction, and change detection. Finally, this paper addresses the challenges encountered in crowdsourced updates for HD maps.

Keywords: autonomous driving; high-definition map; crowdsourcing update; change detection

1. Introduction

In the late 20th century, manual mapping methods were gradually replaced by digitalized cartography [1–4]. With the continuous increase in urbanization and the widespread use of automobiles, there has been a substantial rise in transportation demands, leading to the emergence of navigable digital road maps [5]. At the same time, with the development of information and communication technology, sensor technology, and artificial intelligence, the field of autonomous driving technology has rapidly progressed. According to the Society of Automotive Engineers (SAE) J3016 standard, there are six levels in vehicle automation from L0 to L5 [6]. As autonomous driving technology advances from L2 to beyond L3, it imposes higher demands on navigable digital maps. High-definition (HD) maps have emerged and have gradually become a fundamental component of autonomous driving technology [7,8].

HD maps are specialized electronic maps primarily used for advanced driver-assistance systems (ADAS) and autonomous driving. They contain rich, high-precision geographical road information and are stored and managed in standardized formats [9]. As shown in Figure 1, HD maps generally consist of road networks, lane information, localization features, and traffic infrastructure details. The road network primarily describes the geometric and attribute information related to roads, such as road types, grades, and widths. Lane information focuses on specific lane markings, detailing lane direction, numbers, and speed limits within each lane. Localization features encompass points used by autonomous



Citation: Guo, Y.; Zhou, J.; Li, X.; Tang, Y.; Lv, Z. A Review of Crowdsourcing Update Methods for High-Definition Maps. *ISPRS Int. J. Geo-Inf.* 2024, *13*, 104. https:// doi.org/10.3390/ijgi13030104

Academic Editors: Wolfgang Kainz and Wei Huang

Received: 21 December 2023 Revised: 15 February 2024 Accepted: 14 March 2024 Published: 20 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). vehicles for positioning, including their location, type, texture, and shape. The signal layer contains geometric and semantic information about traffic signs, traffic lights, and specific road markings, detailing their types, heights, and other characteristics. The road network in HD maps aids autonomous vehicles in global navigation and path planning. Lane information is utilized for fine-grained, lane-level path planning in autonomous driving [10,11]. Localization features and signals assist autonomous vehicles in environmental perception and achieving high-precision positioning [12–14].



Figure 1. The main contents of HD maps.

As an indispensable component of autonomous driving, the production and updates of HD maps significantly influence the development of autonomous driving technology. Research institutions, autonomous driving companies, and map producers are actively involved in the research and development of HD maps. Companies such as HERE, TomTom, and Waymo offer HD maps covering major highways globally, provide detailed lane information in key cities worldwide, and supply Advanced Driver-Assistance System (ADAS) maps with global coverage [15,16].

There are two commonly used methods for producing and updating HD maps. The first method involves using mobile mapping systems (MMS) equipped with high-precision light detection and ranging (LiDAR), high-definition cameras, and accurate positioning devices to collect data in specific areas [17]. Subsequently, the collected data undergo processing and feature extraction to generate HD maps. This approach is known as centralized production and the updating of HD maps. While maps produced using this method exhibit high accuracy, they require expensive sensors and skilled professionals to conduct map acquisition and production. With advancements in sensor technology, the cost of cameras and global navigation satellite system (GNSS) positioning devices has decreased. Consequently, more vehicles are equipped with these sensors, enabling partial implementation of advanced driver-assistance features. Simultaneously, these sensors can be utilized for data collection and perception. The collected data are transmitted to the mapping cloud platform for data cleaning, fusion, and information extraction, facilitating updates to HD maps. This approach, known as the crowdsourced updating of highdefinition maps, eliminates the need for expensive sensors and allows for large-scale map updates. However, compared to centralized methods, this approach involves a more complex technical path.

Some map providers adopt a combination of centralized and crowdsourced methods for real-time updates of HD maps. For instance, HERE MMS for collecting foundational HD maps employs crowdsourced vehicles equipped with visual sensors to update road information in real time. Other map providers opt for a single crowdsourced updating approach to build HD maps. For example, Mobileye has introduced a method called Road Experience Management (REM), wherein road information is crowdsourced through vehicle-mounted cameras and deep-learning technology is utilized to recognize and collect various types of road information. Compared to centralized updating methods, crowdsourced updating methods have advantages such as a lower cost and higher update frequency, as shown in Table 1. Therefore, the crowdsourced approach has gradually become the primary method for HD map updates.

Table 1. Comparison of HD map-update methods.

	Centralized Update	Crowdsourced Update
Sensors	High-precision sensors	Low-precision sensors
	(LiDAR, stereo camera, GNSS, IMU)	(GNSS, camera)
Accuracy	Centimeter level	Decimeter level
Update Frequency	Low-update frequency	High-update frequency
	(Quarterly/monthly)	(Monthly/weekly/daily)
Cost	High cost	Low cost

With the advancement of autonomous driving technology, HD maps have gained significant attention. Liu et al. presented a literature review of HD maps, focusing on HD map structure, functionalities, accuracy requirements, and standardization aspects [5]. Bao et al. introduced the concept of HD maps and their usefulness in autonomous vehicles and provided an overview of HD map-creation methods [18]. Elghazaly et al. provided a review of the applications of HD maps in autonomous driving and reviewed the different approaches and algorithms to build HD maps [4]. Existing review studies often focus broadly on the overall construction and application of HD maps; they lack detailed discussions on the crowdsourced updating methods. Therefore, this paper delves into the critical technologies involved in the crowdsourced updating of HD maps. It reviews the latest advancements in these key technologies and concludes by discussing the challenges faced in the crowdsourced updating methods.

The structure of this paper is organized as follows: Section 2 provides a comprehensive overview of the overall framework for crowdsourced updating of HD maps, exploring the key technologies involved in the updating methods. Section 3 reviews the data-collection methods for crowdsourced updates. Section 4 summarizes and reviews the methods for extracting information in HD maps based on crowdsourced data. Section 5 introduces the methods for change detection and updates in HD maps based on crowdsourced data. Section 6 discusses the challenges existing in current crowdsourced updating methods for HD maps. Finally, Section 7 concludes the paper.

2. The Framework of Crowdsourced Updating

With the decreasing cost of sensors, an increasing number of vehicles are equipped with various sensors to record the vehicle-driving status to ensure driving safety. For instance, some taxis are equipped with GNSS receivers, while other vehicles may have dashcams. Moreover, the rise of intelligent driving has led to a growing number of vehicles that integrate multiple sensors, including cameras, LiDAR, and radar, enabling autonomous driving capabilities. The crowdsourced updating method leverages the sensors mounted on vehicles to collect real-time data on vehicle-positioning trajectories and road-sequence imagery, enabling the real-time perception of lane markings, traffic signs, and other elements. The gathered diverse and dispersed data are centrally processed through a map cloud-service platform. A series of algorithms involving data cleaning, data fusion, information extraction, change detection, and map updating are applied to analyze and process the crowdsourced data. This process enables rapid updates of HD maps, which are then transmitted to autonomous vehicles to assist in lane-level navigation, high-precision positioning, environmental perception, and path planning, as shown in Figure 2. Therefore, the key technologies involved in crowdsourced updating methods encompass data collection, real-time environmental perception, data processing, and change detection. These technologies synergistically contribute to constructing the crowdsourced updating system for HD maps. This paper summarizes and analyzes the current state of crowdsourced updating research, starting from these pivotal technologies.



Figure 2. The overall framework of crowdsourced updates for HD maps.

3. Crowdsourced Data Collection for HD Maps Updating

Data collection is the primary step in crowdsourced updating, with an emphasis on cost-effectiveness. While, theoretically, any vehicle equipped with sensors could contribute to the crowd-sourced updating of HD maps, the current trend leans toward utilizing public buses or taxis, which are equipped with standard sensors such as IMU, GNSS, and cameras, to address computational resource constraints and privacy considerations. Some HD map providers, like HERE, employ advanced sensors such as LiDAR or radar for crowd-sourced data collection. As depicted in Figure 3, IMUs and GNSS receivers are capable of collecting vehicle-trajectory data, cameras gather road-image sequences, and LiDAR and radar sensors capture point-cloud data of the road.



Figure 3. Crowdsourced updating sensors and data.

3.1. GNSS-Based Data-Collection Methods

As the cost of positioning devices decreases, most vehicles such as buses and taxis are now equipped with GNSS devices capable of recording the time, speed, and position information of vehicles. These data are commonly referred to as floating car data (FCD), offering advantages of broad coverage and real-time capabilities [19]. Leveraging abundant FCD not only facilitates the acquisition of lane information but also allows for inferring changes in road usage, such as traffic congestion, road closures, and temporary restrictions. These data serve as foundational support for the dynamic updating of HD maps.

According to the sampling frequency, FCD can be categorized into high-frequency and low-frequency types. High-frequency FCD usually have a sampling rate higher than 1 Hz, while low-frequency FCD are typically collected every 10–60 s. In general, high-frequency FCD are often sourced from high-precision GNSS recorders or smartphones. High-precision GNSS recorders can achieve sampling frequencies of 10–100 Hz and positional accuracy up to 3 m. Smartphones, on the other hand, typically have sampling frequencies ranging from 1 Hz to 5 Hz, with positional accuracy slightly lower than high-precision GNSS recorders but still reaching up to 5 m [20,21]. Dolancic et al. utilized devices with diverse spatial and temporal resolutions to collect GNSS trajectories, enabling the derivation of the topology and geometric information required for high-accuracy lane maps [22]. Arman et al. proposed a method that is capable of identifying lanes in highway segments based on GPS trajectories. The GPS data were collected through the Touring Mobilis smartphone app Be-Mobile. It includes time, coordinates, speed, and headings, and it is stored at 1 Hz frequency [23]. High-frequency FCD typically possesses superior spatiotemporal resolution and provides more accurate speed and acceleration information, enabling the extraction of more precise road details. However, this data-collection method increases energy consumption and demands larger storage space for trajectory data. Additionally, high-frequency GNSS receivers are relatively expensive.

Low-frequency FCD typically originate from public transportation vehicles such as taxis and buses. The sampling frequency for this type of data generally falls within the range of 10 s to 60 s, with positioning accuracy ranging from 10 m to 30 m [24–26]. Li et al. proposed a method that utilizes FCD data collected from taxis in Wuhan to detect auxiliary lanes at intersections [24]. Similarly, Kan et al. proposed a method to detect traffic congestion based on FCD data collected by taxis, with a sampling frequency of 60 s [25]. The acquisition of low-frequency FCD is characterized by extended time intervals, contributing to reduced energy consumption and minimized data volume. However, this method yields trajectory data with lower spatial and temporal resolution, leading to limited precision in extracting road information.

In summary, high-frequency FCD data provide higher accuracy and detailed road information. In contrast, low-frequency FCD data often contain more noise, requiring data cleaning before acquiring map information. While FCD serve as a crucial data source for multi-source updated HD maps, it is essential to note that both low-frequency and high-frequency GNSS data may exhibit accuracy instability. This instability can be attributed to factors such as weather, obstruction, or signal reflection, limiting the accuracy of the information obtained. Moreover, while FCD data can typically capture road geometry, inferring specific semantics and attribute information can be challenging. To meet the requirements of HD maps, crowdsourced updating necessitates richer and more precise data.

3.2. Camera-Based Data-Collection Methods

For the crowdsourced updating of HD maps, existing research tends to favor the use of low-cost industrial cameras combined with GNSS receivers to gather road data. This approach allows for the capturing of richer details of the roads. The vision-based crowdsourced data-collection method can be classified based on the number of cameras used as either monocular-vision or stereo-vision approaches. Additionally, depending on whether calibration is required between the cameras and GNSS receivers, these methods can be further categorized into those that require calibration and those that do not.

Monocular vision is a prevalent method for crowdsourced data collection due to its light weight, easy installation, and cost-effectiveness. However, this approach cannot directly capture depth information of the scene. To achieve more accurate 3D road reconstruction, single-camera systems typically integrate with Inertial Navigation Systems (INS) or wheel encoders [27,28]. For instance, Guo et al. proposed a low-cost scheme for

extracting lane information for HD maps. The authors utilized vehicles equipped with standard GPS, INS, and a real camera to collect road-image data during daily driving activities while synchronizing the vehicle's trajectory [29]. Jang et al. also proposed a method for automatically constructing HD maps using a monocular camera setup. Their approach involved obtaining odometry data from the wheel encoder of the vehicle, estimating the vehicle's position, and detecting HD maps information from the obtained images [30].

Today, some massive commercial vehicles are equipped with onboard sensors such as IMU, GPS, wheel encoders, and cameras, which can gather substantial road-related data for crowdsourced updates to HD maps [31], as shown in Figure 4a. Bu et al. utilized a monocular camera setup with a GPS system onboard a commuter bus traveling between downtown Pittsburgh and Washington, Pennsylvania, to autonomously detect changes in pedestrian crosswalks at urban intersections [32]. Moreover, Korean IT company SK Telecom installed camera devices to collect crowdsourced data, which are called the Road Observation Data (ROD) on public buses in Seoul. Kim et al. used this data to update lane information for HD maps [33]. Apart from buses, many private vehicles and ride-hailing cars are also equipped with cameras [34], at present as shown in Figure 4b. Yan et al. cooperated with automobile manufacturers to collect image or video data, along with GPS information during driving, with the users' consent to update lane information in HD maps [35].



Figure 4. The monocular vision-based crowdsourced collection devices. (**a**) Crowdsourced collection device installed on buses [32]; (**b**) crowdsourced collection device installed on ride-hailing cars [34].

Utilizing monocular vision for crowdsourced data collection offers a cost-effective solution that can be easily deployed on various types of vehicles. This method is well-suited for large-scale crowdsourced data-collection projects, allowing for the capturing of a substantial amount of road information in a relatively short time. However, it lacks the ability to directly acquire in-depth information about the scene. To address this limitation and enhance the construction of HD maps, the integration with other sensors becomes necessary.

The aforementioned methods all require calibration to obtain the intrinsic and extrinsic parameters of the monocular camera. This step is necessary to transform the extracted local HD map information from the camera coordinate system to the 3D geographic coordinate system through perspective transformation. However, the process of camera calibration typically demands expertise, relying on experimentation and calculations to acquire the camera's intrinsic and extrinsic parameters. Therefore, some researchers have proposed HD map-construction methods that do not rely on the intrinsic and extrinsic parameters of cameras or smartphones [36–38]. Chawla et al. proposed a method for extracting 3D positions of landmarks in HD maps [36]. This approach is used only on a monocular color camera and GPS without assuming the known camera's intrinsic and extrinsic parameters. Methods that do not require calibration simplify the deployment process of monocular camera setup and make it easier to install on different platforms. However, the absence of calibration may result in the inaccurate estimation of the camera intrinsic and extrinsic

parameters, impacting the accuracy and precision of the collected data. To address this issue, more complex post-processing is needed to correct distortions in the images.

The monocular camera-based method faces limitations in practical application due to constraints related to the field of view, weather conditions, and variations in lighting. Consequently, some scholars have opted for stereo-vision approaches for crowdsourced data collection, aiming to enhance the accuracy of the collected data [39]. Jeong et al. utilized a vehicular setup equipped with a forward-looking stereo ZED camera, an IMU, and two wheel encoders to extract lane information for HD maps [40]. Lagahit et al. utilized a custom-made stereo camera comprising two Sony α 6000 cameras to update HD maps of traffic cones [41]. Compared to monocular cameras, the use of stereo cameras allows for more accurate and detailed information in HD maps. Monocular cameras, when using data from a single trip, are susceptible to road obstructions and variations in lighting conditions, making it challenging to meet the accuracy requirements for HD maps. Hence, it is necessary to cluster data from multiple trips to compensate for their accuracy limitations. However, monocular cameras are relatively inexpensive and easier to deploy on various types of vehicles, facilitating a more widespread adoption for updating HD maps through crowd-sourced means.

3.3. LiDAR-Based Data-Collection Methods

LiDAR creates point-cloud data by emitting laser beams and measuring their return time, capturing both static and dynamic information on roads. Compared to GNSS-based and camera-based data-collection methods, LiDAR achieves higher accuracy in information acquisition. However, due to the higher cost, LiDAR-based data-collection methods are less commonly used for crowdsourced updates in HD maps. Instead, they are often integrated into MMS for centralized data collection in HD map creation. However, there are still some researchers using LiDAR for crowdsourced updates of HD maps. Dannheim et al. proposed a cost-effective method for constructing HD maps [42]. They employed postprocessing algorithms to extract HD map information formed from data collected by a LiDAR, three cameras, a GNSS receiver, an IMU to gather data, and the vehicle's CAN BUS. Kim et al. utilized low-cost GNSS and various specifications of LiDAR for crowdsourced data collection in a simulated environment. This enabled the updating of point-cloud maps in HD maps [43,44]. Liu et al. developed an acquisition component that consists of LiDAR and RGB-D cameras to update the dynamic information in HD maps in real time [45].

The high cost of LiDAR has impacted its application in crowdsourced updating for HD maps. However, with advancements in sensor technology, the cost of LiDAR is decreasing, especially with the emergence of solid-state LiDAR. While there is limited existing research on crowdsourced data collection based on LiDAR, it holds potential for application in updating HD maps through crowdsourcing efforts. Furthermore, radar data have received relatively less emphasis due to limited semantic analysis and lower resolution. However, the emergence of novel radar technologies introduces prospects for its application in crowdsourced updates of HD maps.

This section discusses various crowdsourced data-collection methods from a sensor perspective, categorizing them into three types: GNSS-based methods, camera-based methods, and LiDAR-based methods. Each method has its own advantages and disadvantages, as shown in Table 2. The GNSS-based method is cost-effective, easy to deploy in various vehicles, has extensive coverage, and updates rapidly. However, it suffers from lower accuracy and lacks detailed information about the lanes and traffic signs. Solely relying on GNSS data is challenging to meet the requirements for HD map updates. The LiDAR-based approach offers higher accuracy and is less affected by changes in environmental lighting. Nevertheless, it faces the challenge of high costs and difficulty in widespread adoption in crowdsourced vehicular platforms. The combined GNSS data-and-camera-based crowd-sourced data-collection method is currently the mainstream approach. It has relatively lower costs, is easily scalable, and, although it is susceptible to environmental changes, the increase in data volume helps compensate for this issue.

	GNSS-Based	Camera-Based	LiDAR-Based
Advantages	Cost-effective, extensive coverage, updates rapidly	Cost-effective, easily scalable, information- abundant	hHgh accuracy and robustness
Disadvantages	Low accuracy, limited information content	Environmentally sensitive	High cost, difficulty in widespread

Table 2. Comparison of crowdsourced data-collection methods.

The collection of crowdsourced data employs a decentralized approach, gathering road information from various types of sensors. There exist inconsistencies in both the methods and accuracy of data collection among different vehicles and devices. Additionally, different vehicles might have distinct formats and coordinate systems. Hence, research into algorithms for aligning and integrating crowdsourced data is crucial for the crowdsourced updates of HD maps.

4. Information Extraction Methods Form Crowdsourced Data

Extracting the road network, lanes, and traffic sign information based on crowdsourced data is one of the crucial steps in updating HD maps. The process of extracting information from crowdsourced data generally involves two stages, as shown in Figure 5. After collecting crowdsourced data, at the vehicle end, a local road map is extracted through real-time perception algorithms and uploaded to the cloud. At the cloud end, the perception results from vehicles are fused, and through algorithms involving data cleaning, clustering, mining, and more, the final HD map information is extracted.



Figure 5. The process of extraction HD map information from crowdsourced data [34].

4.1. Vehicle-End Information Extraction

The extraction of HD map elements at the vehicle end involves two processes: realtime road-information detection and local map reconstruction. The extraction of road information involves the real-time extraction of elements such as lanes and signs from images using either traditional or deep-learning-based image-segmentation methods. On the other hand, local map reconstruction involves estimating the camera's motion by tracking feature points between adjacent image frames to reconstruct the environment. In this section, the paper reviews the current research status from these two aspects.

4.1.1. Real-Time Detection of Road Information

Real-time road-information detection can be categorized into two methods: traditional and deep-learning-based methods. Traditional methods rely on brightness and geometric features of road markings or traffic signs for detection. Sobel and Canny operators are commonly used for the edge detection of road markings [46–48]. However, the practicality of traditional methods is constrained as they tend to be specific to certain scenarios, lacking robustness in more complex environments.

Due to the heightened robustness and superior capacity for generalization, deeplearning methods have rapidly gained attention. Deep-learning methods utilize semantic segmentation or target-detection techniques to extract HD map features from sensor data. Target-detection methods like R-CNN [49], Faster-RCNN [50], FPN [51], and YOLO [52] primarily focus on location detection and category determination of elements in the scene, such as signs and traffic lights. Semantic-segmentation methods like U-Net [53], DEEPLABv3 [54], Monolayout [55], and Segnet primarily target a pixel-level determination of road areas, lane markings, and other features in the scene, enhancing geometric positioning accuracy. Additionally, for different instances of the same category of objects, neural networks can employ instance-segmentation methods like LaneNet [56] and CLRNET [57] to uniquely label each target.

For the extraction of HD map features, in addition to utilizing deep learning for feature perception, it is necessary to perform a projection transformation to convert map features from 2D coordinate space to 3D coordinate space [58], as shown in Figure 6. For instance, Zhou et al. introduced a Lane Mask-Propagation Network method to detect lane information in images. Subsequently, the authors projected lane markings from a perspective space to a 3D space using GNSS positional data, extracting lane information for HD maps [34]. Qin et al. employed a CNN-based segmentation method to segment the image into multiple categories such as ground, lanes, stop lines, road signs, curbs, vehicles, bikes, and pedestrians, which is used for constructing a semantic map. Similarly, based on intrinsic and extrinsic parameters, the semantic pixels are projected from the perspective image plane to the ground plane [31]. The transformation from 2D space to 3D space involves projecting the extracted features using the sensors' intrinsic and extrinsic parameters. This process significantly influences the accuracy of the extracted map features. In addition, the utilization of deep-learning algorithms, particularly based on segmentation or detection, for real-time perception and the extraction of HD map elements has marked substantial advancement. Nevertheless, this approach often relies on single-frame image recognition, lacking temporal information. Additionally, during the spatial-projection process, significant distortion commonly exists, limiting the accuracy of extracted elements.



Figure 6. A framework diagram for deep-learning-based HD map-element extraction [58].

4.1.2. Local Map Reconstruction

Considering the low accuracy of the image-segmentation results of a single frame, in order to extract higher-precision road information, some researchers have proposed to use Structure from Motion (SfM), Visual Odometry (VO), or Simultaneous Localization and Mapping (SLAM) algorithms to reconstruct local semantic maps. SfM is a traditional method for 3D reconstruction that reconstructs the 3D structure of an environment from a set of images by recovering the camera's motion. In SfM, the initial step involves determining the camera-motion trajectory through the detection and matching of feature points in images. Subsequently, leveraging this camera-motion trajectory alongside the positional information of feature points enables the computation of the 3D structure of the observed scene. Finally, the optimization of these parameters is performed to enhance the precision of 3D reconstruction [59–61]. However, the existing SFM-based local map reconstruction encounters challenges associated with high-computational complexity. To address these challenges, Zhanabatyrova et al. proposed a framework based on SfM technology that supports lightweight change detection. This framework selectively updates the semantic 3D map of the HD map, focusing solely on areas where changes occur [62].

VO is a specific scenario within SfM and also the front-end part of SLAM algorithms [63]. It is the process of gradually estimating vehicle poses from a sequence of continuous images, facilitating the reconstruction of environmental information. VO algorithms can be categorized into model-based approaches, as well as deep-learningbased approaches. Model-based VO algorithms perform bit-position estimation through mathematical modeling and geometric computation, which can be further classified into the feature-based method and the direct method [64]. The feature-based approach uses SIFT [65], SURF [66], BRIEF [67], ORB [68] et al. for feature extraction and matching to achieve data association between images. The direct method completes the data correlation directly from the optical flow and estimates the camera motion by minimizing the photometric error [69,70]. Model-based VO methods necessitate processes such as feature extraction, feature matching, and camera calibration. These algorithms exhibit a relatively high level of complexity and are vulnerable to environmental noise and sensor inaccuracies. Consequently, recent research has witnessed an increasing emphasis on exploring deep-learning-based VO algorithms. Deep-learning-based VO methods can be used to recover the camera pose, directly from the image pairs, using an end-to-end approach. Commonly used deep-learning-based VO methods include DeepVO [71], SuperPoint [72], CNN-SVO [73], LoFTR [74], etc. Deep-learning-based methods are also applied to the construction of local maps in HD maps. For instance, Qin et al. used the VO method after semantic segmentation and perspective transformation of the image to construct a local semantic map by stacking semantic points from multiple frames [31], as shown in Figure 7. This approach enables the construction of more comprehensive and detailed local maps with richer semantic information, contributing to a HD map with enhanced content.



Figure 7. Local semantic-map construction based on visual odometry [31].

Different from SfM and VO, SLAM aims to create maps on a global scale. Its goal is to achieve a comprehensive estimation of the entire vehicle's trajectory. This approach necessitates loop-closure detection and global optimization to attain a more precise environmental map. Traditional slam algorithms usually include several steps: front-end visual odometry, back-end optimization, loopback detection, and map building. In recent years, the combination of deep learning and SLAM algorithms has received widespread attention [75,76]. Deep-learning-based SLAM methods can better understand the semantic information in the environment, improve the accuracy of data association, and enhance the robustness and generalization ability of the algorithms. Research on local semantic-map construction for HD map updating based on SLAM algorithms has also received extensive attention [30,40,77–79]. These methods typically leverage deep-learning techniques to segment data, extract lane features, and perform loop-closure detection to achieve the construction of local maps. However, the process of local map construction based on SFM, VO, or SLAM algorithms usually involves multiple complex steps, resulting in a lengthy algorithmic process with high computational demands. While deep learning has made notable progress in certain steps, the overall workflow still requires ample computational resources and time. The increased computational complexity poses challenges for real-time performance in these methods.

Presently, academia and several autonomous driving companies have introduced various online algorithms for local semantic-map creation. These enable end-to-end HD map construction at the vehicle end. Several classical algorithms include HDMapNet [80], VectorMap [81], MapTR [82,83], and StreamMapNet [84], among others [85–87]. These methods are capable of autonomously generating localized HD map through BEV perception. HDMapNet, in particular, stands out as the pioneering approach to directly apply deep neural networks to semantic-map construction. It introduces a feature projection module from perspective view to BEV, capable of encoding image features or point-cloud data to predict vectorized map elements in the BEV. HDMapNet consists of an image and point-cloud encoder alongside a BEV Decoder [80], as shown in Figure 8. However, it suffers from high-computational requirements and extended processing times. To address this issue, VectorMap deviates from constructing maps via semantic segmentation and instead relies on object-detection methodologies, abstracting map elements into essential key points. It directly detects observations in a vectorized format, thereby minimizing the necessity for subsequent post-processing. This achieves an end-to-end generation of local vector maps [81]. These methods primarily rely on single-frame data input for map construction, raising concerns about the robustness. StreamMapNet addresses the issue by incorporating temporal information between data frames in online map construction. While its main predictor resembles other single-frame network structures, it utilizes a buffer to store propagated memory features [84]. In addition to online local map construction, researchers have increasingly focused on other aspects related to HD maps, such as topological prediction [88,89]. This research involves modeling the topological structure of road networks as a set of values that can be learned through neural networks.



Figure 8. The model overview of HDMapNet [80].

Table 3 provides a comparison between methods for local map construction based on SfM, VO, SLAM, and end-to-end online map construction. Generally, SLAM-based methods exhibit better accuracy, while end-to-end methods prioritize real-time performance in dynamic scenarios. Additionally, SLAM methods perform well in structured environments, whereas end-to-end methods demonstrate stronger adaptability to diverse environments. It is worth noting that both SLAM-based methods and end-to-end methods acquire local maps rather than complete offline HD maps. At the vehicular end, whether based on SLAM or end-to-end methods for local HD map construction, there is a high demand for computational power, making it challenging to widely propagate crowdsourced HD map updates. Additionally, the locally constructed semantic maps at the vehicular end are typically short range and lack in representation for longer distances. Furthermore, the precision of singular local maps generated at the vehicle end often falls short of meeting the requirements for HD maps.

SfM, VO, SLAM-Based Methods			End-to-End Methods	
	Traditional	Deep Learning-Based	- Life to Life Wethous	
Advantages	Mature and power-efficient	Strong adaptability to complex environments, efficient	Simplified workflow, adaptability, real-time capability	
Disadvantages	Environment- sensitive, complex computations	High-computational resource requirements, high-data-volume demands	Challenges in acquiring large-scale training data, high-computational power requirements	

Table 3. Comparison of local map-reconstruction methods.

4.2. Cloud-End Information Extraction

To build a globally comprehensive HD map, it is essential to transmit the multi-vehicle local map data to the cloud end for fusion, cleaning, extraction, and updating. Existing research tends to focus more on constructing local maps at the vehicle end, while there is comparatively less emphasis on research for updating global HD maps in the cloud. Das et al. proposed an HD map-construction framework that combines vehicle- and cloudbased components [39], as illustrated in Figure 9. At the vehicle end, the authors modeled data as a pose graph for fusion. At the cloud end, they used an approach similar to the multi-agent SLAM, combining crowdsourced data from multiple vehicles to generate a unified environmental map. However, the authors did not specifically tackle the noise issue within crowdsourced data. In response, Pannen et al. [90], subsequent to uploading the data to the cloud, implemented a clustering and cleansing algorithm utilizing DBSCAN. Following this, the B-spline curves were employed to fit the data, consequently extracting high-precision lane information. Kim et al. generated a new feature layer for HD maps at the vehicle end and uploaded it to the map cloud through a mobile network system. Within the map cloud, multiple feature layers collected by crowdsourced vehicles were integrated using data-association algorithms. Subsequently, they employed a recursive least-squares algorithm to update the feature layers of the HD maps [91].

Researchers are increasingly focusing on methods for extracting and updating HD map information within map clouds. However, current research primarily concentrates on aligning and fusing multisource data, with limited exploration into methods for crowd-sourced data clean. However, compared to data collected using MMS, crowd-sourced data tend to have lower accuracy. Data cleaning in the cloud can help enhance the accuracy of crowdsourced maps.



Figure 9. The HD map-construction framework integrates both vehicle and cloud components [39].

5. Crowdsourcing Update for HD Maps

After extracting the crowdsourced HD maps in the cloud, conducting change detection on the existing map is a crucial step to ensure the freshness of the HD map. The specific process is illustrated in Figure 10. Initially, the crowdsourced map is associated with existing HD maps. Following this, change detection is carried out using probabilitybased or learning-based algorithms. Then, within the update cycle, there is a continuous estimation of the credibility of map-feature updates. Then, the map is updated when the credibility reaches a certain threshold. One of the challenges in updating HD maps lies in the uncertainty of the change locations. Therefore, it is necessary to associate and match each element between the crowdsourced maps and the existing HD maps to determine the locations where changes have occurred. Change-detection methods can be categorized into probabilistic-based approaches and learning-based approaches.



Figure 10. The process of map-change detection and updates.

5.1. Probabilistic-Based Approaches

The probability-based change-detection methods are the current mainstream approaches for detecting changes in HD maps [92]. Commonly used methods include particle filtering [93,94], cost functions [33,95], and random variable approaches [90]. To detect and manage the physical changes in the HD maps, probability-based methods typically categorize map features into three classes: normal features, deleted features, and newly added features [93,96], as shown in Figure 11. Subsequently, by assessing the degree of match between feature information and the existing HD map, the probability of feature existence is estimated. Data with higher existence probabilities are retained in the HD maps, Welte et al. used location information to assess map quality and proposed a Kalman filter-smoothing method with residuals. They employed the covariance intersection of

residuals for error detection [96]. Probability-based methods are an effective means for HD map updating. This approach aligns features from crowdsourced map with those in the HD map, calculating the change probability based on the alignments. While this method entails longer computational cycles and involves complex algorithms, it plays a crucial role in maintaining the precision of HD maps. By continuously monitoring the HD map with multiple crowdsourced maps over a period, it provides real-time updates to change probabilities, thereby enhancing the reliability of map updates.



Figure 11. Classification of HD map features [93].

5.2. Learning-Based Approaches

Learning-based methods, by adopting an end-to-end approach, use raw data and HD maps as inputs and directly output the change information. However, there is currently a limited number of learning-based change-detection methods available. Heo et al. designed an HD map change-detection algorithm based on deep-metric learning, offering a unified framework that directly maps input images to estimated probabilities of HD changes, as illustrated in the Figure 12. The authors employed adversarial learning to implement the difference between RGB images and HD maps into a common feature space. Subsequently, they utilized a local change detector to estimate pixel-level change probabilities [97]. Similarly, Lambert and Hays proposed a learning-based HD map change-detection algorithm. They formulated the problem as predicting whether the map needs an update by combining HD maps and crowdsourced data. Additionally, by mining thousands of hours of data from over 9 months of autonomous vehicle-fleet operations, they assembled the first dataset for HD map-change detection based on real observed map alterations [98].



Figure 12. A learning-based HD map change-detection system [97].

In general, updating HD maps is challenging due to the uncertainty regarding the timing and location of changes. Moreover, different elements within HD maps have varied update frequencies; for instance, the update cycle for roadside infrastructure might be longer than that for lane markings. Additionally, the precision of crowdsourced data from a single source can be somewhat unreliable, necessitating the calculation of the credibility of crowdsourced data to ensure the accuracy of HD map updates.

6. Challenges

Updating HD maps by crowdsourced data stands as a pivotal factor in ensuring map currency and precision, which is crucial for the advancement of autonomous driving technologies. However, this process encompasses multiple stages, from data collection and processing to the extraction of information elements and, finally, updating the HD map, forming a relatively complex, multistage workflow. Presently, there are ongoing studies targeting crowdsourced updates for HD maps. However, achieving a comprehensive crowdsourced-based map-updated process remains challenging. Most research focuses on specific aspects such as the real-time perception of information elements, local map construction, or map-change detection. There remain numerous unresolved issues within the entire crowdsourced update procedure.

The accuracy of HD maps stands as one of their most critical features. However, the precision and reliability of crowdsourced data carry considerable uncertainty, stemming from the diversity in data-collection methods and variations in sensor accuracy. Effectively cleansing and enhancing the quality of crowdsourced data poses a significant challenge. It necessitates the development of robust algorithms to identify low-quality data within crowdsourced information and establish corresponding data-cleansing procedures to enhance data quality. This improvement in data quality serves as a reliable foundation for the updating and maintenance of HD maps.

Apart from the precision issues associated with data sources, the real-time construction of local maps on the vehicle also faces several challenges, including insufficient accuracy and computational limitations. Despite the various online map-construction methods, these methods serve as supplements to existing maps and have not made breakthroughs in significantly improving the accuracy of existing HD maps. Firstly, in terms of accuracy, methods for real-time local map construction on the vehicle side are often constrained by the limitations of perception devices. Additionally, these methods require extensive annotated data for training and may struggle to adapt to complex and dynamic real-world road environments in practical applications. Secondly, insufficient computational power is also a limiting factor in the real-time construction of local maps on the vehicle end. Source devices in vehicles, such as smartphones or regular cameras, are typically constrained by power consumption, making it challenging to meet the high computational requirements of current algorithms.

The third point concerns open datasets for crowdsourced updates of HD maps. While there are some open datasets related to autonomous driving, they are not specifically designed for crowdsourced updates of HD maps. An ideal dataset for crowdsourced updates should include a large amount of road data collected by different types of sensors, and the dataset should encompass both the pre-update and post-update HD maps for quantitative analysis of algorithmic results. The lack of such datasets presents a significant challenge for researchers in acquiring HD maps and crowdsourced data. The introduction of datasets for crowdsourced updates of HD maps will significantly advance research and development in this field.

Finally, the collective update of HD maps also involves concerns regarding personal privacy. For instance, crowdsourced data might contain sensitive personal information like license plates or facial details. These pieces of information, when correlated with trajectory data, could potentially infer individual behavioral patterns or routines. Therefore, effective measures for privacy protection are necessary to ensure that the collection and utilization

of crowdsourced data do not compromise user privacy rights. Only by safeguarding user privacy can the broader adoption of crowdsourced updates for HD maps be facilitated.

7. Conclusions

HD maps play a vital role in advancing autonomous driving technology and ensuring their accuracy and freshness hinges on crowdsourced updates. While the methodologies for crowdsourced updates in HD maps have garnered attention from numerous researchers, there has been a limited summarization of these endeavors in the existing literature. Hence, this paper conducts a review of the technologies employed in recent years for crowdsourced updates in HD maps. It begins by contrasting two methods: crowdsourced updates versus centralized updates, analyzing the pivotal technologies involved in crowdsourced updates. Furthermore, it compares and synthesizes the methods related to crowdsourced data collection, information extraction, and map updates. Additionally, this paper delves into the challenges facing crowdsourced updates in HD maps.

Some challenges that require further research include: 1. Streamlining the entire process of crowdsourced updates; 2. enhancing techniques for crowdsourced data cleaning; 3. implementing privacy measures for crowdsourced data; 4. advancing the development of datasets that combine crowdsourced data with HD maps.

Author Contributions: Conceptualization, Yuan Guo; Methodology, Jian Zhou; Formal analysis, Yuan Guo, Jian Zhou, and Zhicheng Lv; Investigation, Xicheng Li, Youcheng Tang, and Jian Zhou; Writingoriginal draft preparation, Yuan Guo and Xicheng Li; Supervision, Youchen Tang and Zhicheng Lv.; Funding acquisition, Yuan Guo and Jian Zhou. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant No. 42201480, Grant No. 42101448), China Postdoctoral Science Foundation (Grant No. 2023M732683).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

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

References

- Jones, A.; Blake, C.; Davies, C.; Scanlon, E. Digital maps for learning: A review and prospects. *Comput. Educ.* 2004, 43, 91–107. [CrossRef]
- 2. Jiang, L.; Liang, Q.; Qi, Q.; Ye, Y.; Liang, X. The heritage and cultural values of ancient Chinese maps. J. Geogr. Sci. 2017, 27, 1521–1540. [CrossRef]
- Black, J. Maps and History: Constructing Images of the Past; Yale University Press: New Haven, CT, USA, 2000; ISBN 978-0-300-08693-5.
- 4. Elghazaly, G.; Frank, R.; Harvey, S.; Safko, S. High-Definition Maps: Comprehensive Survey, Challenges, and Future Perspectives. *IEEE Open J. Intell. Transp. Syst.* **2023**, *4*, 527–550. [CrossRef]
- Liu, R.; Wang, J.; Zhang, B. High Definition Map for Automated Driving: Overview and Analysis. J. Navig. 2020, 73, 324–341. [CrossRef]
- 6. SAE J3016 Automated-Driving Graphic. Available online: https://www.sae.org/site/news/2019/01/sae-updates-j3016-automated-driving-graphic (accessed on 22 November 2023).
- Ye, S.; Fu, Y.; Wang, W.; Pan, Z. Creation of high definition map for autonomous driving within specific scene. In Proceedings of the International Conference on Smart Transportation and City Engineering 2021, Chongqing, China, 6–8 August 2021; SPIE: Bellingham, WA, USA, 2021; Volume 12050, pp. 1365–1373.
- 8. Seif, H.G.; Hu, X. Autonomous Driving in the iCity—HD Maps as a Key Challenge of the Automotive Industry. *Engineering* **2016**, 2, 159–162. [CrossRef]
- Poggenhans, F.; Pauls, J.-H.; Janosovits, J.; Orf, S.; Naumann, M.; Kuhnt, F.; Mayr, M. Lanelet2: A high-definition map framework for the future of automated driving. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 1672–1679.
- 10. Guo, X.; Cao, Y.; Zhou, J.; Huang, Y.; Li, B. HDM-RRT: A Fast HD-Map-Guided Motion Planning Algorithm for Autonomous Driving in the Campus Environment. *Remote Sens.* **2023**, *15*, 487. [CrossRef]
- Jian, Z.; Zhang, S.; Chen, S.; Lv, X.; Zheng, N. High-Definition Map Combined Local Motion Planning and Obstacle Avoidance for Autonomous Driving. In Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 9–12 June 2019; 2019; pp. 2180–2186.

- 12. Huang, Y.; Zhou, J.; Li, X.; Dong, Z.; Xiao, J.; Wang, S.; Zhang, H. MENet: Map-enhanced 3D object detection in bird's-eye view for LiDAR point clouds. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *120*, 103337. [CrossRef]
- Fang, J.; Zhou, D.; Song, X.; Zhang, L. MapFusion: A General Framework for 3D Object Detection with HDMaps. In Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Prague, Czech Republic, 27 September–1 October 2021; pp. 3406–3413.
- 14. Shin, D.; Park, K.; Park, M. High Definition Map-Based Localization Using ADAS Environment Sensors for Application to Automated Driving Vehicles. *Appl. Sci.* 2020, *10*, 4924. [CrossRef]
- HERE HD Live Map | Autonomous Driving System | Platform | HERE. Available online: https://www.here.com/platform/HDlive-map (accessed on 2 February 2024).
- 16. HD Map. Available online: https://www.tomtom.com/products/hd-map/ (accessed on 2 February 2024).
- Ye, C.; Zhao, H.; Ma, L.; Jiang, H.; Li, H.; Wang, R.; Chapman, M.A.; Junior, J.M.; Li, J. Robust Lane Extraction From MLS Point Clouds Towards HD Maps Especially in Curve Road. *IEEE Trans. Intell. Transport. Syst.* 2022, 23, 1505–1518. [CrossRef]
- 18. Bao, Z.; Hossain, S.; Lang, H.; Lin, X. A review of high-definition map creation methods for autonomous driving. *Eng. Appl. Artif. Intell.* **2023**, 122, 106125. [CrossRef]
- 19. Li, X.; Zhang, Y.; Xiang, L.; Wu, T. Urban Road Lane Number Mining from Low-Frequency Floating Car Data Based on Deep Learning. *ISPRS Int. J. Geo-Inf.* 2023, 12, 467. [CrossRef]
- Neuhold, R.; Haberl, M.; Fellendorf, M.; Pucher, G.; Dolancic, M.; Rudigier, M.; Pfister, J. Generating a Lane-Specific Transportation Network Based on Floating-Car Data. In *Advances in Human Aspects of Transportation*; Stanton, N.A., Landry, S., Di Bucchianico, G., Vallicelli, A., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 1025–1037.
- 21. Yang, X.; Tang, L.; Stewart, K.; Dong, Z.; Zhang, X.; Li, Q. Automatic change detection in lane-level road networks using GPS trajectories. *Int. J. Geogr. Inf. Sci.* 2018, 32, 601–621. [CrossRef]
- Dolancic, M. Automatic lane-level road network graph-generation from Floating Car Data Page. GI Forum 2016, 4, 231–242. [CrossRef]
- Arman, M.A.; Tampère, C.M.J. Road centreline and lane reconstruction from pervasive GPS tracking on motorways. *Procedia* Comput. Sci. 2020, 170, 434–441. [CrossRef]
- 24. Li, X.; Wu, Y.; Tan, Y.; Cheng, P.; Wu, J.; Wang, Y. Method Based on Floating Car Data and Gradient-Boosted Decision Tree Classification for the Detection of Auxiliary Through Lanes at Intersections. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 317. [CrossRef]
- 25. Kan, Z.; Tang, L.; Kwan, M.-P.; Ren, C.; Liu, D.; Li, Q. Traffic congestion analysis at the turn level using Taxis' GPS trajectory data. *Comput. Environ. Urban Syst.* 2019, 74, 229–243. [CrossRef]
- Fan, L.; Zhang, J.; Wan, C.; Fu, Z.; Shao, S. Lane-Level Road Map Construction considering Vehicle Lane-Changing Behavior. J. Adv. Transp. 2022, 2022, e6040122. [CrossRef]
- Dabeer, O.; Ding, W.; Gowaiker, R.; Grzechnik, S.K.; Lakshman, M.J.; Lee, S.; Reitmayr, G.; Sharma, A.; Somasundaram, K.; Sukhavasi, R.T.; et al. An end-to-end system for crowdsourced 3D maps for autonomous vehicles: The mapping component. In Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Vancouver, BC, Canada, 24–28 September 2017; pp. 634–641.
- Herb, M.; Weiherer, T.; Navab, N.; Tombari, F. Crowd-sourced Semantic Edge Mapping for Autonomous Vehicles. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, 4–8 November 2019; pp. 7047–7053.
- 29. Guo, C.; Kidono, K.; Meguro, J.; Kojima, Y.; Ogawa, M.; Naito, T. A Low-Cost Solution for Automatic Lane-Level Map Generation Using Conventional In-Car Sensors. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 2355–2366. [CrossRef]
- Jang, W.; An, J.; Lee, S.; Cho, M.; Sun, M.; Kim, E. Road Lane Semantic Segmentation for High Definition Map. In Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 2018; pp. 1001–1006.
- 31. Qin, T.; Huang, H.; Wang, Z.; Chen, T.; Ding, W. Traffic Flow-Based Crowdsourced Mapping in Complex Urban Scenario. *IEEE Robot. Autom. Lett.* **2023**, *8*, 5077–5083. [CrossRef]
- 32. Bu, T.; Mertz, C.; Dolan, J. Toward Map Updates with Crosswalk Change Detection Using a Monocular Bus Camera. In Proceedings of the 2023 IEEE Intelligent Vehicles Symposium (IV), Anchorage, AK, USA, 4–7 June 2023; pp. 1–8.
- Kim, K.; Cho, S.; Chung, W. HD Map Update for Autonomous Driving with Crowdsourced Data. *IEEE Robot. Autom. Lett.* 2021, 6, 1895–1901. [CrossRef]
- 34. Zhou, J.; Guo, Y.; Bian, Y.; Huang, Y.; Li, B. Lane Information Extraction for High Definition Maps Using Crowdsourced Data. *IEEE Trans. Intell. Transp. Syst.* 2023, 24, 7780–7790. [CrossRef]
- Yan, C.; Zheng, C.; Gao, C.; Yu, W.; Cai, Y.; Ma, C. Lane Information Perception Network for HD Maps. In Proceedings of the 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece, 20–23 September 2020; pp. 1–6.
- Chawla, H.; Jukola, M.; Brouns, T.; Arani, E.; Zonooz, B. Crowdsourced 3D Mapping: A Combined Multi-View Geometry and Self-Supervised Learning Approach. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October 2020–24 January 2021; pp. 4750–4757.
- Szabó, L.; Lindenmaier, L.; Tihanyi, V. Smartphone Based HD Map Building for Autonomous Vehicles. In Proceedings of the 2019 IEEE 17th World Symposium on Applied Machine Intelligence and Informatics (SAMI), Herlany, Slovakia, 24–26 January 2019; pp. 365–370.

- Golovnin, O.K.; Rybnikov, D.V. Video Processing Method for High-Definition Maps Generation. In Proceedings of the 2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), Vladivostok, Russia, 6–9 October 2020; pp. 1–5.
- Das, A.; IJsselmuiden, J.; Dubbelman, G. Pose-graph based Crowdsourced Mapping Framework. In Proceedings of the 2020 IEEE 3rd Connected and Automated Vehicles Symposium (CAVS), Victoria, BC, Canada, 18 November–16 December 2020; pp. 1–7.
- Jeong, J.; Cho, Y.; Kim, A. Road-SLAM: Road marking based SLAM with lane-level accuracy. In Proceedings of the 2017 IEEE Intelligent Vehicles Symposium (IV), Los Angeles, CA, USA, 11–14 June 2017; pp. 1736–1743.
- 41. Lagahit, M.L.R.; Tseng, Y.H. A preliminary study on updating high definition maps: Detecting and positioning a traffic cone by using a stereo camera. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, XLII-4-W19, 271–274. [CrossRef]
- 42. Dannheim, C.; Maeder, M.; Dratva, A.; Raffero, M.; Neumeier, S.; Icking, C. A cost-effective solution for HD Maps creation. In Proceedings of the AmE 2017—Automotive meets Electronics; 8th GMM-Symposium, Dortmund, Germany, 7–8 March 2017; pp. 1–5.
- 43. Kim, C.; Cho, S.; Sunwoo, M.; Resende, P.; Bradai, B.; Jo, K. Updating Point Cloud Layer of High Definition (HD) Map Based on Crowd-Sourcing of Multiple Vehicles Installed LiDAR. *IEEE Access* **2021**, *9*, 8028–8046. [CrossRef]
- 44. Kim, C.; Cho, S.; Sunwoo, M.; Resende, P.; Bradaï, B.; Jo, K. Cloud Update of Geodetic Normal Distribution Map Based on Crowd-Sourcing Detection against Road Environment Changes. J. Adv. Transp. 2022, 2022, e4486177. [CrossRef]
- Liu, Q.; Han, T.; Xie, J.; Kim, B. Real-Time Dynamic Map With Crowdsourcing Vehicles in Edge Computing. *IEEE Trans. Intell.* Veh. 2023, 8, 2810–2820. [CrossRef]
- Chen, G.H.; Zhou, W.; Wang, F.J.; Xiao, B.J.; Dai, S.F. Lane Detection Based on Improved Canny Detector and Least Square Fitting. Adv. Mater. Res. 2013, 765–767, 2383–2387. [CrossRef]
- Talib, M.L.; Ramli, S. A Review of Multiple Edge Detection in Road Lane Detection Using Improved Hough Transform. *Adv. Mater. Res.* 2015, 1125, 541–545. [CrossRef]
- 48. Wijaya, B.; Jiang, K.; Yang, M.; Wen, T.; Tang, X.; Yang, D. Crowdsourced Road Semantics Mapping Based on Pixel-Wise Confidence Level. *Automot. Innov.* 2022, *5*, 43–56. [CrossRef]
- Liang, T.; Bao, H.; Pan, W.; Pan, F. Traffic Sign Detection via Improved Sparse R-CNN for Autonomous Vehicles. J. Adv. Transp. 2022, 2022, e3825532. [CrossRef]
- Gavrilescu, R.; Zet, C.; Foşalău, C.; Skoczylas, M.; Cotovanu, D. Faster R-CNN:an Approach to Real-Time Object Detection. In Proceedings of the 2018 International Conference and Exposition on Electrical And Power Engineering (EPE), Iasi, Romania, 18–19 October 2018; pp. 0165–0168.
- 51. Ren, K.; Huang, L.; Fan, C.; Han, H.; Deng, H. Real-time traffic sign detection network using DS-DetNet and lite fusion FPN. *J. Real-Time Image Proc.* 2021, *18*, 2181–2191. [CrossRef]
- Wang, J.; Chen, Y.; Dong, Z.; Gao, M. Improved YOLOv5 network for real-time multi-scale traffic sign detection. *Neural Comput. Appl.* 2023, 35, 7853–7865. [CrossRef]
- 53. Yang, X.; Li, X.; Ye, Y.; Lau, R.Y.K.; Zhang, X.; Huang, X. Road Detection and Centerline Extraction Via Deep Recurrent Convolutional Neural Network U-Net. *IEEE Trans. Geosci. Remote Sens.* 2019, 57, 7209–7220. [CrossRef]
- Mahmud, M.N.; Osman, M.K.; Ismail, A.P.; Ahmad, F.; Ahmad, K.A.; Ibrahim, A. Road Image Segmentation using Unmanned Aerial Vehicle Images and DeepLab V3+ Semantic Segmentation Model. In Proceedings of the 2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 27–28 August 2021; pp. 176–181.
- Mani, K.; Daga, S.; Garg, S.; Shankar, N.S.; Krishna Murthy, J.; Krishna, K.M. Mono Lay out: Amodal scene layout from a single image. In Proceedings of the 2020 IEEE Winter Conference on Applications of Computer Vision (WACV), Snowmass Village, CO, USA, 1–5 March 2020; pp. 1678–1686.
- Wang, Z.; Ren, W.; Qiu, Q. LaneNet: Real-Time Lane Detection Networks for Autonomous Driving. arXiv 2018, arXiv:1807.01726. [CrossRef]
- 57. Zheng, T.; Huang, Y.; Liu, Y.; Tang, W.; Yang, Z.; Cai, D.; He, X. CLRNet: Cross Layer Refinement Network for Lane Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 898–907. Available online: https://openaccess.thecvf.com/content/CVPR2022/html/Zheng_CLRNet_Cross_Layer_ Refinement_Network_for_Lane_Detection_CVPR_2022_paper.html (accessed on 4 February 2024).
- Tian, W.; Ren, X.; Yu, X.; Wu, M.; Zhao, W.; Li, Q. Vision-based mapping of lane semantics and topology for intelligent vehicles. *Int. J. Appl. Earth Obs. Geoinf.* 2022, 111, 102851. [CrossRef]
- 59. Liang, D.; Guo, Y.; Zhang, S.; Zhang, S.-H.; Hall, P.; Zhang, M.; Hu, S. LineNet: A Zoomable CNN for Crowdsourced High Definition Maps Modeling in Urban Environments. *arXiv* **2018**, arXiv:1807.05696. [CrossRef]
- Szántó, M.; Kobál, S.; Vajta, L.; Horváth, V.G.; Lógó, J.M.; Barsi, Á. Building Maps Using Monocular Image-feeds from Windshieldmounted Cameras in a Simulator Environment. *Period. Polytech. Civ. Eng.* 2023, 67, 457–472. [CrossRef]
- 61. Fisher, A.; Cannizzaro, R.; Cochrane, M.; Nagahawatte, C.; Palmer, J.L. ColMap: A memory-efficient occupancy grid mapping framework. *Robot. Auton. Syst.* 2021, 142, 103755. [CrossRef]
- 62. Zhanabatyrova, A.; Souza Leite, C.F.; Xiao, Y. Automatic Map Update Using Dashcam Videos. *IEEE Internet Things J.* 2023, 10, 11825–11843. [CrossRef]
- 63. Eyvazpour, R.; Shoaran, M.; Karimian, G. Hardware implementation of SLAM algorithms: A survey on implementation approaches and platforms. *Artif. Intell. Rev.* 2023, *56*, 6187–6239. [CrossRef]

- 64. Lim, K.L.; Bräunl, T. A Review of Visual Odometry Methods and Its Applications for Autonomous Driving. Available online: https://arxiv.org/abs/2009.09193v1 (accessed on 13 December 2023).
- 65. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. Int. J. Comput. Vis. 2004, 60, 91–110. [CrossRef]
- Bay, H.; Tuytelaars, T.; Van Gool, L. SURF: Speeded Up Robust Features. In Proceedings of the Computer Vision—ECCV 2006, Graz, Austria, 7–13 May 2006; Leonardis, A., Bischof, H., Pinz, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 404–417.
- 67. Calonder, M.; Lepetit, V.; Strecha, C.; Fua, P. BRIEF: Binary Robust Independent Elementary Features. In Proceedings of the Computer Vision—ECCV 2010, Heraklion, Greece, 5–11 September 2010; Daniilidis, K., Maragos, P., Paragios, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 778–792.
- 68. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2564–2571.
- 69. Newcombe, R.A.; Lovegrove, S.J.; Davison, A.J. DTAM: Dense tracking and mapping in real-time. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 2320–2327.
- Engel, J.; Schöps, T.; Cremers, D. LSD-SLAM: Large-Scale Direct Monocular SLAM. In Proceedings of the Computer Vision— ECCV 2014, Zurich, Switzerland, 6–12 September 2014; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 834–849.
- Wang, S.; Clark, R.; Wen, H.; Trigoni, N. DeepVO: Towards end-to-end visual odometry with deep Recurrent Convolutional Neural Networks. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; pp. 2043–2050.
- 72. DeTone, D.; Malisiewicz, T.; Rabinovich, A. SuperPoint: Self-Supervised Interest Point Detection and Description. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–23 June 2018; pp. 224–236. Available online: https://openaccess.thecvf.com/content_cvpr_2018_workshops/w9/html/DeTone_SuperPoint_Self-Supervised_Interest_CVPR_2018_paper.html (accessed on 4 February 2024).
- Loo, S.Y.; Amiri, A.J.; Mashohor, S.; Tang, S.H.; Zhang, H. CNN-SVO: Improving the Mapping in Semi-Direct Visual Odometry Using Single-Image Depth Prediction. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 5218–5223.
- Sun, J.; Shen, Z.; Wang, Y.; Bao, H.; Zhou, X. LoFTR: Detector-Free Local Feature Matching With Transformers. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2021; pp. 8922– 8931. Available online: https://openaccess.thecvf.com/content/CVPR2021/html/Sun_LoFTR_Detector-Free_Local_Feature_ Matching_With_Transformers_CVPR_2021_paper.html (accessed on 4 February 2024).
- 75. Chen, C.; Wang, B.; Lu, C.X.; Trigoni, N.; Markham, A. A Survey on Deep Learning for Localization and Mapping: Towards the Age of Spatial Machine Intelligence. *arXiv* 2020, arXiv:2006.12567. [CrossRef]
- 76. Duan, C.; Junginger, S.; Huang, J.; Jin, K.; Thurow, K. Deep Learning for Visual SLAM in Transportation Robotics: A review. *Transp. Saf. Environ.* **2019**, *1*, 177–184. [CrossRef]
- Doer, C.; Henzler, M.; Messner, H.; Trommer, G.F. HD Map Generation from Vehicle Fleet Data for Highly Automated Driving on Highways. In Proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA, 19 October–13 November 2020; pp. 2014–2020.
- Liebner, M.; Jain, D.; Schauseil, J.; Pannen, D.; Hackeloer, A. Crowdsourced HD Map Patches Based on Road Model Inference and Graph-Based SLAM. In Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 9–12 June 2019; pp. 1211–1218.
- 79. Zhang, P.; Zhang, M.; Liu, J. Real-Time HD Map Change Detection for Crowdsourcing Update Based on Mid-to-High-End Sensors. *Sensors* **2021**, *21*, 2477. [CrossRef]
- Li, Q.; Wang, Y.; Wang, Y.; Zhao, H. HDMapNet: An Online HD Map Construction and Evaluation Framework. In Proceedings of the 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–27 May 2022; pp. 4628–4634.
- Liu, Y.; Yuan, T.; Wang, Y.; Wang, Y.; Zhao, H. VectorMapNet: End-to-end Vectorized HD Map Learning. In Proceedings of the 40th International Conference on Machine Learning, Honolulu, HI, USA, 23–29 July 2023; pp. 22352–22369.
- 82. Liao, B.; Chen, S.; Wang, X.; Cheng, T.; Zhang, Q.; Liu, W.; Huang, C. MapTR: Structured Modeling and Learning for Online Vectorized HD Map Construction. *arXiv* 2023, arXiv:2208.14437. [CrossRef]
- Liao, B.; Chen, S.; Zhang, Y.; Jiang, B.; Zhang, Q.; Liu, W.; Huang, C.; Wang, X. MapTRv2: An End-to-End Framework for Online Vectorized HD Map Construction. *arXiv* 2023, arXiv:2308.05736. Available online: http://arxiv.org/abs/2308.05736 (accessed on 30 November 2023).
- Yuan, T.; Liu, Y.; Wang, Y.; Wang, Y.; Zhao, H. StreamMapNet: Streaming Mapping Network for Vectorized Online HD Map Construction. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 1–10 January 2024; pp. 7356–7365. Available online: https://openaccess.thecvf.com/content/WACV2024/html/Yuan_ StreamMapNet_Streaming_Mapping_Network_for_Vectorized_Online_HD_Map_Construction_WACV_2024_paper.html (accessed on 5 February 2024).
- 85. Qiao, L.; Ding, W.; Qiu, X.; Zhang, C. End-to-End Vectorized HD-Map Construction with Piecewise Bezier Curve. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 13218–13228. Available online: https://openaccess.thecvf.com/content/CVPR2023/html/Qiao_End-to-End_Vectorized_ HD-Map_Construction_With_Piecewise_Bezier_Curve_CVPR_2023_paper.html (accessed on 30 November 2023).

- 86. Yu, J.; Zhang, Z.; Xia, S.; Sang, J. ScalableMap: Scalable Map Learning for Online Long-Range Vectorized HD Map Construction. *arXiv* 2023, arXiv:2310.13378. [CrossRef]
- Ding, W.; Qiao, L.; Qiu, X.; Zhang, C. PivotNet: Vectorized Pivot Learning for End-to-end HD Map Construction. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 2–3 October 2023.
- 88. Liao, B.; Chen, S.; Jiang, B.; Cheng, T.; Zhang, Q.; Liu, W.; Huang, C.; Wang, X. Lane Graph as Path: Continuity-preserving Path-wise Modeling for Online Lane Graph Construction. *arXiv* 2023, arXiv:2303.08815. [CrossRef]
- Ben Charrada, T.; Tabia, H.; Chetouani, A.; Laga, H. TopoNet: Topology Learning for 3D Reconstruction of Objects of Arbitrary Genus. Comput. Graph. Forum 2022, 41, 336–347. [CrossRef]
- Pannen, D.; Liebner, M.; Hempel, W.; Burgard, W. How to Keep HD Maps for Automated Driving Up To Date. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; pp. 2288–2294.
- 91. Kim, C.; Cho, S.; Sunwoo, M.; Jo, K. Crowd-Sourced Mapping of New Feature Layer for High-Definition Map. *Sensors* 2018, 18, 4172. [CrossRef]
- Ding, W.; Hou, S.; Gao, H.; Wan, G.; Song, S. LiDAR Inertial Odometry Aided Robust LiDAR Localization System in Changing City Scenes. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; pp. 4322–4328.
- 93. Jo, K.; Kim, C.; Sunwoo, M. Simultaneous Localization and Map Change Update for the High Definition Map-Based Autonomous Driving Car. *Sensors* 2018, *18*, 3145. [CrossRef]
- 94. Pannen, D.; Liebner, M.; Burgard, W. HD Map Change Detection with a Boosted Particle Filter. In Proceedings of the 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 20–24 May 2019; pp. 2561–2567.
- Li, B.; Song, D.; Kingery, A.; Zheng, D.; Xu, Y.; Guo, H. Lane Marking Verification for High Definition Map Maintenance Using Crowdsourced Images. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October 2020–24 January 2021; pp. 2324–2329.
- Welte, A.; Xu, P.; Bonnifait, P.; Zinoune, C. HD Map Errors Detection using Smoothing and Multiple Drives. In Proceedings of the 2021 IEEE Intelligent Vehicles Symposium (IV Workshops), Nagoya, Japan, 11–17 July 2021; p. 37. Available online: https://hal.science/hal-03521309 (accessed on 29 November 2023).
- Heo, M.; Kim, J.; Kim, S. HD Map Change Detection with Cross-Domain Deep Metric Learning. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October 2020–24 January 2021; pp. 10218–10224.
- 98. Lambert, J.; Hays, J. Trust, but Verify: Cross-Modality Fusion for HD Map Change Detection. *arXiv* 2022, arXiv:2212.07312. [CrossRef]

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