negative obstacle (points lower than the road plane). Moreover, since water bodies cannot be detected by LiDAR due to refraction and reflection, the authors propose a technique to detect potholes filled with water by scanning the image for large areas of missing data.

Shuo *et al.* [29] improves the aforementioned method by projecting the points on the camera plane and interpolating the depth values of the projected points to receive a depth image. They use both horizontal and vertical histograms to coarsely detect the road area and refine it respectively. Although they state their method as sensor fusion between the monocular camera and LiDAR, they do not utilize the color values of the camera images. Both works [15] and [29] use the KITTI dataset as a benchmark and achieve great results, comparable to machine learning methods.

Other techniques for pothole detection may include laser scanning, ground penetrating radar, ultrasonic sensor, as well as multi-sensor fusion, especially concerning fusion with imaging information. An extensive review of such techniques falls beyond the scope of this article. However, an interested reader may be referred to the survey in [31].

2) Point cloud saliency: One of main challenges in techniques utilizing point clouds is the inherent noise and the increased computational cost due to the unordered data structure of point clouds. To address such challenges, saliency map extraction has been proposed as a powerful step in point cloud processing to reduce noise and data dimensionality, leading to more robust solutions and computational efficiency [32] [33]. Yet, the use of local saliency in pothole detection has not been sufficiently examined. Saliency maps were constructed from point clouds obtained from Mobile Laser Scanning (MLS) in [34] for road crack detection. MLS point clouds contain spatial information (i.e., Euclidean coordinates) and intensity information, and thus the extracted features could leverage both height and intensity information. Feature saliency was estimated by calculating the distances from the normal of each point to the principal normal of the input point clouds. In a similar setting, Wang *et al.* [35] extracted saliency maps in MLS point clouds by projecting the distance of each point's normal vector to the point cloud's dominant normal vector into a hyperbolic tangent function space.

3) Cooperative driving: While significant advances have been made for single-agent perception, many applications require multiple sensing agents and cross-agent communication for more accurate results. Objects, captured by the single-agent's sensor devices, may be heavily occluded or far away from the sensors' view, resulting in sparse observations. Nevertheless, failing to detect and predict the accurate position or moving intention of these occluded or "hard-to-see" objects might have harmful consequences in safety-critical situations, and especially if the reaction time is very narrow [36]. The development of multi-agent solutions can lead to collaborative perception and, through information sharing, may improve the driving performance and experiences, providing endless possibilities for safe driving.

Recently, cooperative autonomous driving has been considered as a possible solution to improve the performance and safety of autonomous vehicles [37]. Cooperative perception for 3D object detection can be performed via early or late fusion of information, i.e., combination of multiple sensing points of view or fusion of object detection results, respectively.Both fusion approaches can extend the perception of the sensing system, however, only the early fusion approach can actually exploit complementary information. A major challenge that arises regarding cooperative perception is how to effectively merge sensors' data received from different vehicles to obtain a precise and comprehensive perception outcome. Additionally, despite the attention that cooperative driving has attracted recently, the absence of a suitable open dataset for benchmarking algorithms has made it difficult to develop and assess cooperative perception technologies.

Xu *et al.* [38] presented the first open dataset and used it to benchmark fusion strategies for V2V (vehicle-to-vehicle) perception. They also plan to extend the dataset with more tasks as well as sensor suites and investigate more multimodal sensor fusion methods in the V2V and V2I (vehicle-to-infrastructure) settings. Arnold *et al.* [37] proposed a system that produces a perception of complex road segments (e.g., complex T-junctions and roundabouts) using a network of roadside infrastructure sensors with fixed positions. Chen *et al.* [39] studied the raw-data level cooperative perception for enhancing the detection ability of self-driving systems. They fuse the sensor data collected from different positions and angles of connected vehicles, relying on LiDAR 3D point clouds. Liu *et al.* [40] addressed the collaborative perception problem, where one agent is required to perform a perception task and can communicate and share information with other agents on the same task.

Chen *et al.* [41] proposed a point cloud feature-based cooperative perception framework for connected autonomous vehicles to increase object detection precision. The features are selected to be rich enough for the training process, and at the same time have an intrinsically small size to achieve real-time edge computing. Guo *et al.* [42] proposed a cooperative fusion method to combine spatial feature maps for achieving a higher 3D object detection performance. Yuan *et al.* [43] proposed a 3D keypoints feature fusion scheme for cooperative driving detection to remedy the problem of low bounding box localization accuracy. Fang *et al.* [44] presented an iterated split covariance intersection filter-based cooperative localization strategy with a decentralized framework. In addition, they adopted a point cloud registration method to obtain the relative pose estimation using mutually shared information from neighbour vehicles. Kim and Liu [45] presented the concept of cooperative autonomous driving using mirror neuron-inspired intention awareness and cooperative perception, providing information on the upcoming traffic situations ahead, even beyond line-of-sight and field-of-view.

4) Situational awareness and AR infotainment: In the case of semi-autonomous vehicles, where the operator/driver may be asked to take manual control of the car at any moment, it is of great importance [46] to implement notification paradigms that direct the operator's, possibly reduced, attention to the event that triggered the take-over request [47], [48]. Recently, the automotive industry started to invest funds and efforts into AR technology and its integration with In-Vehicle Information