# Localization System Enhanced with CDLPE: A Low-Cost, Resilient Map-Matching Algorithm

Author Names Omitted for Anonymous Review. Paper-ID [12]

*Abstract*—Accurate and robust localization is a critical component in intelligent vehicles, playing a significant role in route planning and efficient navigation. There's a rising trend towards affordable positioning solutions. These use common vehicular sensors like GPS, IMU, and cameras to improve navigation accuracy. This paper presents a comprehensive, low-cost localization framework with a lightweight map. The framework introduces two key novelties. Firstly, we propose a method known as the Cross-Dimensional Lane and Pose Estimator (CDLPE), designed to effectively resist scenarios with poor satellite signals. In addition, our system delivers a reliable localization service by effectively integrating matching results and capitalizing on the benefits of the sensors used, coupled with an understanding of the environment. We have verified the robustness of our method under different driving scenarios. Compared to the classical Iterative Closest Point (ICP) algorithm, the lane identification accuracy has improved by 4.42% and 9.23% during normal and weak satellite signal conditions, respectively. Videos in: https://youtu.be/DsYXSeWQhWc
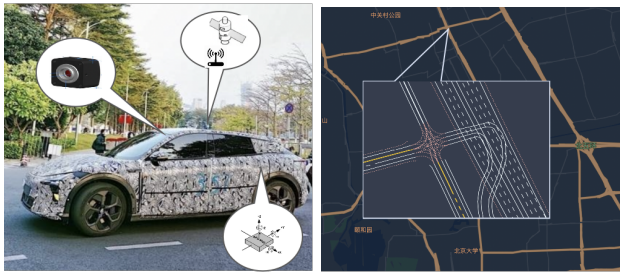
Fig. 1: Test Vehicle Configuration and Lightweight Map

## I. INTRODUCTION

Localization is a critical technology that helps intelligent vehicles determine their precise position and orientation within the environment. This process is essential for path planning, obstacle avoidance, and ensuring safe and efficient navigation. Vehicle localization typically involves the integration of various sensors such as Global Navigation Satellite System (GNSS), LiDAR, radar, and cameras, along with sophisticated algorithms to process sensor data, and may also utilize map data and other information sources to accurately determine the vehicle's position and orientation. In recent years, the approach of using high-precision maps and Lidar for auxiliary positioning has been quite common due to its high accuracy, with Nayak et al. [14] serving as a prime illustration. However, due to the high maintenance cost of large-scale 3D high-precision maps, and with the rapid improvement of image-based algorithm accuracy due to the development of deep learning, people are increasingly inclined to use a low-cost

positioning solution that combines lightweight maps such as Poggenhans et al. [15] and images with other commonly used vehicular sensors such as GPS, IMU for precise positioning. Each sensor possesses distinct attributes: GNSS is effective in open spaces, IMU consistently computes position, orientation, and speed but is prone to integration drift, and cameras necessitate road features and a good field of view. Maps not only assist in positioning but also provide prior driving information, significantly aiding in the integration of data from all sensors.

This paper introduces a robust and precise localization system for intelligent vehicle navigation in both urban and highway settings. Our work adroitly amalgamates data from an array of sensors with environmental knowledge and employs an innovative map matching algorithm on a lane-level light map. The system is designed to consistently provide accurate location services even in complex scenarios such as tunnels, intersections, and multi-layered roads. In summary, our main contributions are as follows:

1. We present an accurate and low-cost complete localization framework for intelligent vehicle, which effectively integrates map-matching results and leverages the advantages of multiple sensors with the understanding of environment.

2. We have developed a map matching algorithm that effectively resists poor GNSS signal scenarios and significantly enhances the accuracy of lane association. This method, called the Cross-Dimensional Lane and Pose Estimator (CDLPE), represents a substantial improvement in the field.

3. Our localization system has been rigorously tested on vehicles navigating crowded urban streets on a daily basis. The results demonstrate that our system consistently provides promising localization outcomes across a variety of driving scenarios.

## II. RELATED WORKS

In this section, we contextualize our contributions within relevant sub-fields.

### A. Low-cost vehicular localization systems

With the advancement of computer vision, utilizing low-cost cameras [5] is the direction of development for intelligent vehicles. In earlier works, it's common to utilize 3d point cloud based high-definition map [33, 13, 26] in localization. These works [32, 16, 34] usually transfer the image data to a same format by specific algorithm firstly and than exploit rigid registration algorithms such as Iterative Closest Point(ICP) [2]

to estimate pose. Due to the high maintenance costs of high-definition maps, people [19, 27, 12] are now more inclined to use semantics representation which are more robust against illumination variation and seasonal changes [15] from data in vector-format maps for positioning, such as lane marking, curb, and so on. To leverage additional sensor inputs, some learning-based method [21, 31] try to utilize a coarse GPS location and gravity direction. However, multi-sensor fusion by filter methods has been long leveraged to build accurate and reliable localization systems [9, 7, 8]. Our work falls into the category of vision-based localization, integrating vision systems in a loosely coupled manner with the inertial navigation system (INS).

### B. Map matching algorithms

An critical issue in map-matching methods is that the detected feature can falsely correspond with a one stored on a offline map in the lateral direction. To tackle this problem, RANSAC algorithm [28, 20], probabilistic description [10, 23] and tracking methods [24] are applied, and several researches believe ego-lane identification is important, [25, 30] explored lane-change detection system. [3, 9] divide the location task into several parts, including road level, lane level and ego-lane level localization. [1] proposed two distinct map matching algorithms fused by EKF. [4, 6] present approaches for lane-level localization in a coarse-to-fine fashion. Some works [18, 11, 29] explored deep learning algorithms. Nonetheless, the effectiveness of most skills is highly dependent on the quality of perception and GNSS inputs in contributing to the pose estimation process. Thus, this work introduces the Cross-Dimensional Lane and Pose Estimator (CDLPE), a method that optimally utilizes previous maps and visual data by cross-dimensionally estimating the vehicle's pose with different algorithms and fusing the result adaptively, thereby bolstering the system's resilience in scenarios of poor visibility or weak GNSS signals.

### III. APPROACH

### A. System Overview

Figure 2 shows the block diagram of the proposed vehicle localization system. It utilizes a GPS, IMU, built-in wheel speed sensor, single front camera, a digital map and the planned route. The system combines GPS, dead reckoning using INS feedback and wheel odometry, and multi-layer map matching algorithms through an error state Kalman filter (ESKF). The sensor fusion module uses a unicycle kinematic model, and ESKF helps avoid parameter singularity and ensures their linearization. Like many public researches, the system also uses a neural network for lane lines detection and tracking, fitting the detected lines to a cubic curve, then calibrated using multi-frame detection results. These curves are cross-validated with the map to filter out unstable results, playing a crucial role in map matching.

### B. Cross-Dimensional Lane and Pose Estimator (CDLPE)

As shown in the Figure 2, CDLPE cross-dimensionally estimates its ego-lane and pose in multiple modules. The algorithm starts from matching at road level, which employs a standard forward algorithm in Hidden Markov Modeling (HMM). Then, in order to make better use of features from different dimensions, the lane-level matching of features and geometry are run independently with different algorithms. Finally, the environment understanding module determine the output pose as well as estimation confidence.

*1) Lane-level Feature Matching:* This module aims to rely on rough initial positioning information, located road segment, lane-level map data as a priori, and combined with perception lane lines features to determine thes approximate location of the vehicle. This module applies a histogram filter, which approximates the posteriors by decomposing the state space into finitely many spaces, and representing the cumulative posterior for each region by a single probability value [22]. Its formula is as follows

$$\bar{p}_{i,t} = \sum_j p(X_t = x_i | u_t, X_{t-1} = x_j) p_{j,t-1} \quad (1)$$

$$p_{i,t} = \eta p(z_t | X_t = x_i) \bar{p}_{i,t} \quad (2)$$

where $p_{i,t}$ represents the belief of each state $x_i$ at time $t$, $u_t$ is the control input, and $z_t$ is the measurement vector. Equations (1) and (2) correspond to the prediction and update step. As the Figure 3(a) shows, the $p_{i,t}$ means the probability of that vehicle is located in the lateral grid $i$.

In our work, the prediction step is performed based on the knowledge of odometry velocity, meanwhile according to Bayes theorem, the update step can be rewritten as

$$p_{i,t} = \eta p(z_t | X_t = x_i) \bar{p}_{i,t} \ = \eta \frac{p(X_t = x_i | z_t) p(z_t)}{p(X_t = x_i)} \bar{p}_{i,t} \quad (3)$$

in which $p(X_t = x_i) = \frac{d_i^{right} - d_i^{left}}{width_{road}}$, and $p(X_t = x_i | z_t)$ is calculated by fusing all the measurement probabilities. Meanwhile, $p(z_t)$ can be merged into the normalization factor $\eta$, thus we can derive

$$p_{i,t} = \bar{\eta} \frac{p(X_t = x_i | z_t)}{p(X_t = x_i)} \bar{p}_{i,t} \quad (4)$$

The problem is abstracted as a one-dimensional positioning problem with discrete and limited status, where state $x_i$ means the vehicle is located in the lateral grid $i$. The lane line position data, given as points (latitude and longitude) on the map, is fitted with a cubic polynomial to match the perception. As Figure 3(a) shows, for ease of subsequent computations, the leftmost lane line is used as the origin, adjusting the intercepts accordingly, i.e $d_i = d_i - d_0$.

Next, we introduce the process of prediction and update corresponding to Equation (1) and (4).
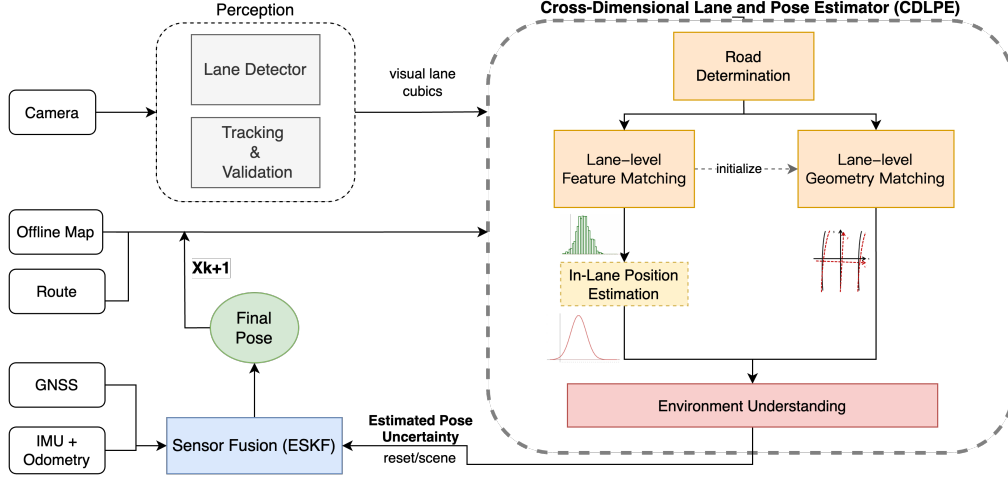
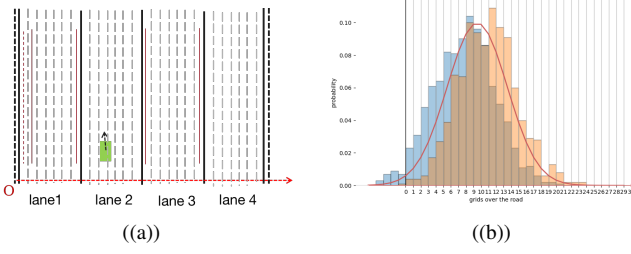Fig. 2: Overview of the Proposed Localization System Enhanced with CDLPE



Fig. 3: (a) The grey dashed lines represent the lateral grids $D = \{d_1, d_2, ..., d_k\}$ in the road, with each grid measuring 20 cm in width, and the red line shows the one-dimensional coordinate system within the lane; (b) The blue and orange bars represent probability distribution of prediction $\bar{p}_{i,t}$ and measurement $p(x_i|z_{1:m})$ respectively. The red curve denotes final fitted Gaussian function.

*a) The Prediction Step:* The prediction step is to predict on the new belief of the state with the historical distribution of the filter and the motion model. Since the state only consider horizontal localization, we calculate the new lateral position of the car on the road at the current time $t$ based on the posterior distribution fitted at the time $t-1$ and the odometry information.

$$\mu_i = \mu_{i-1} + velocity_x * \delta t \qquad (5)$$

$$\bar{p}_{i,t} \sim N(\mu_i, \sigma_{i-1}^2) \qquad (6)$$

which can be interpreted as a slight movement based on the probability distribution within the road of the previous moment.

*b) The Update Step:* There are only two states of a vehicle in a certain grid, i.e., it is within the grid or not. Therefore, in this solution, static binary Bayes filtering is used to calculate the probability $p(x_i'|z_{1:m}')$, which indicates the confidence of vehicle in the $i$th grid under the observation

of $z_1 : z_m$ at time $t$. For simplicity, we ignore the symbol $t$ representing the time:

$$l_k^i = l_{k-1}^i + \log \frac{p(x_i|z_k)}{1 - p(x_i|z_k)} - \log \frac{p(x_i)}{1 - p(x_i)} \qquad (7)$$

in which $l$ is the logistic probability, i.e. $l_k^i = \log \frac{p(x_i|z_{1:k})}{1-p(x_i|z_{1:k})}$. After obtain $l_m^i$ by the above Equation (7), it can be derived that the probability of vehicle is on the grid $i$ under the measurements $z_1 : z_m$ at time $t$, as shown in the following equation:

$$p(x_i|z_{1:m}) = 1 - \frac{1}{1 + e^{l_m^i}} \qquad (8)$$

Taking the number of lanes detected on the left and right side of the vehicle as an example of observation, we illustrate the calculation of observation probability $p(X_t = x_i|z_t)$ in Equation (7). When the vehicle observes $n_{left}$ lane lines on the left and $n_{right}$ lane lines on the right through the camera, the probability can be expressed by Equation (9):

$$p(x_i|z_{NoL}) = \frac{1}{4width} \int_{d_i^l}^{d_i^r} \left[ 1 + erf(\frac{y - d_{max(n_l-1,0)}}{\sqrt{2}\delta}) \right]$$
$$* \left[ 1 - erf(\frac{y - d_{min(n-n_r,n-1)}}{\sqrt{2}\delta}) \right] dy \qquad (9)$$

This method utilizes the concept of filtering and offers impressive scalability. It allows for seamless integration of new observational data into the existing model.

*c) In-lane Pose Estimation:* After deriving the grid probability distribution $p(x_i)$, the vehicle's position must be determined. In ambiguous situations, this is achieved using an optimization solution. We employ the Ceres Solver, a nonlinear optimization library, to fit a standard Gaussian function to our probability distribution, as Figure 3(b) shows. The cost function can be articulated through the subsequent equations.
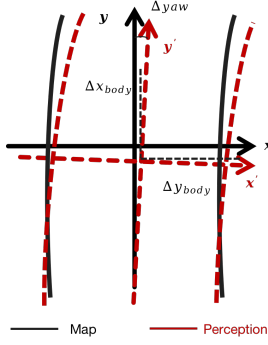
$$J = \sum^i (p(x_i) - f(x)) \qquad (10)$$

Fig. 4: The map and perceived red lines under the current vehicle pose assumption.

$$\text{s.t.} \quad f(x) = \exp(-\frac{(x-\mu)^2}{2var^2}) \qquad (11)$$

This estimated pose serves as both the output for in-lane positioning and the prior position for the next moment's prediction process.

*2) Lane-level Geometry Matching:* This section focuses on achieving centimeter-level positioning and maintaining stability based on geometric relationships, unaffected by accidental incorrect ego-lane determination in feature matching. The plan is divided into two parts: the first part addresses the data association problem between perceived and prior map lane lines data. The second part obtains accurate pose optimization based on the established association relationship.

*a) Data Association:* This part abstracts the data association problem between perception and map lane lines into a classic maximum matching set solving problem of a weighted bipartite graph. It calculates weights from several lane lines data features that affect positioning accuracy and assigns solutions as a whole to ensure the optimal matching set is directly obtained during the data association stage. The key point is to identify and eliminate erroneous data effectively. The weight of the $i$-th group of matching pairs can be calculated by Equation (12)

$$weight_i = c_i * \sum_j (1 - \frac{error_{i,j}}{\sum_i error_{i,j}}) \qquad (12)$$

where $j$ represents the type of error, and $c_i$ is the correction coefficient, which is related to the properties of the perception-acquired lane line itself in the $i$-th group of matching pairs. After obtaining the association weight matrix, the Hungarian matching algorithm is directly used to solve it to obtain the maximum matching set. This solution can maintain good correlation accuracy when there are partial misdetections in the perception data such as ground cracks or potential problems in the map such as temporary road repairs.

*b) Pose Optimization:* For vehicle pose correction using perception and map data matching results, it's necessary to convert the map lane lines data to the body system, which involves projection transformation. Equidistant and equiangular projections can both introduce errors. As the vehicle moves further from the initial reference point, these projection errors

increase. To mitigate this, our approach uses the vehicle's current position as the reference point for converting map lane lines data to the body system, as shown in Figure 4. Then, we solve the pose error based on the data association results of the map data and the perceived lane lines equation in the body system. The state to optimize is the error of the vehicle pose, defined as

$$X = [\Delta x_{body}, \Delta y_{body}, \Delta yaw]^T \qquad (13)$$

, in which $\Delta x_{body}$, $\Delta y_{body}$ and $\Delta yaw$ represents lateral, longitude and orientation error respectively. The residual term comes from the fitness of the curve equation of the lane lines perceived from the points from the prior map as well as the Mahalanobis distance between the map and the perceived object. This non-linear problem can be formulated as

$$\min_X \frac{1}{2} \sum_j^m \sum_i^m p_i^j (fitness\_error_i^j(X)^2) + \frac{1}{2} p(X^2)$$
$$+ \frac{1}{2} \sum_k^l p_k (obj\_err_k(X)^2) \qquad (14)$$

$$\text{s.t.} \quad fitness\_error_i^j(X) = f(x_{i,map}^j) - y_{i,map}^j \qquad (15)$$

$$X_{lower} < X < X_{upper} \qquad (16)$$

where $p$ is the robust kernel function in optimization, $f$ is the detected cubic curve function, and $obj\_error$ is used to evaluate the matching degree of other detected static features in the map such as stop lines, etc.

*3) Environment Understanding:* This module consists of uncertainty estimation and scene recognition. Uncertainty estimation calculates the value $UncV$ for the fusion algorithm to determine the variance while fusing matching result. This uncertainty stemming from matching uncetainty and map inaccuracy, as (17) shows.

$$UncV = \lambda * (UncV_{geometry} + UncV_{feature}) + UncV_{map} \qquad (17)$$

$UncV_{feature} = \frac{1}{\sum_i p(i)} \sum_i p(i)(d_{opt} - d_i)^2$ indicates the uncertainty in the estimated probability distribution across road grids in feature matching. On the other hand, $UncV_{geometry}$ represents the variance value of the optimized $[longitude, latitude, yaw]^T$ in geometry matching module, given by the Ceres solver in the tangent space. Note that the position with lower uncertainty will be output as matching result.

Currency issues like changes in lanes, width, and heading are common in maps. To detect potential anomalies, we measure this uncertainty value by the difference between confident failed visual detection results and the map data.

$$UncV_{map} = \frac{1}{N_{failed}} \sum_i^{N_{failed}} (\frac{abs(yaw_i - yaw_{map})}{scale\_factor * velocity} + \frac{abs(offset_i - offset_{map})}{width}) \qquad (18)$$
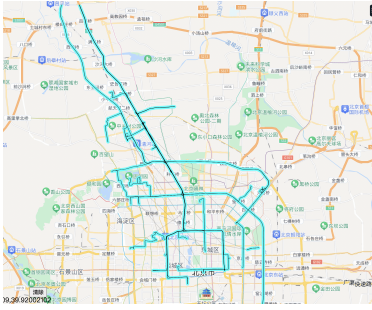
Fig. 5: The Coverage of Beijing Test Dataset

As a supplement to the use of $UncV$, the scene recognition identifying whether it is a scene that have a significant impact on map matching or need a special strategy in fusion.

In essence, the road determination module outputs the vehicle's current located road and recommendation position, useful for navigation planning and deviation assessment in extreme scenarios. The lane-level feature matching calculates the vehicle's probability distribution within the road grids, using Bayesian methods based on determined road and perception data. It estimates the vehicle's approximate location, mitigating lane matching errors during significant GPS deviation by lane lines features. The lane-level geometry matching focuses on producing precise and stable within-lane pose, which will not be influenced by errors in feature-based matching after itself being initialized successfully. The environment understanding module not only guarantees the algorithm's applicability in complex road networks, but also provides uncertainty information for fusion algorithm to choose estimated pose from matching.

## IV. EVALUATION RESULTS

### A. Experimental Setup and Preliminary Performance

Our algorithm is road-tested daily within the city by our intelligent vehicles. Each test vehicle is equipped with a front camera, GNSS receiver, IMU and vehicle's wheel odometry. To construct the ground truth value, we have additionally assembled NovAtel devices.The test set consists of approximately 300km of data in Beijing, shown in Figure 5. Due to the lack of ground truth in some scenarios without GNSS signals, we have provided a preliminary performance estimate based on manual statistics, as shown in Table I below. Qualitative and quantitative experiment will further evaluate the accuracy and robustness of the method.

| Scenario | Distance | Lateral RMS < 1.0m |
|---|---|---|
| Highway/Expressway | 145km | 0.999 |
| Urban with servere shelter | 42km | 0.988 |
| Tree-lined road | 23km | 0.992 |
| Tunnel | 53km | 1.0 |
| Under stacked road | 25km | 0.975 |
| Ordinary or open urban road | 98km | 0.994 |

TABLE I: The performance of various scenarios in Beijing testset
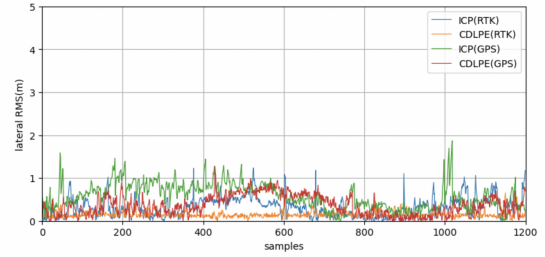


((a)) Lateral RMS Errors
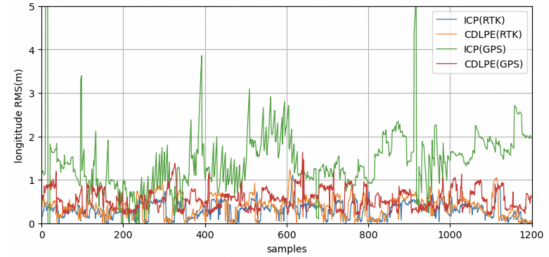


((b)) Longititude RMS Errors

Fig. 6: Comparison of errors set by different experimental methods

### B. Quantitative Analysis

We benchmark the performance of our localization approach against the classical ICP method [17] on our Beijing test set. Based on the perception detection of lane lines points and map lane lines sample points, we used ICP for matching, replacing the CDLPE module. For fairness, the type of lane line is used as an expanded dimension in the ICP method. To clearly illustrate the contributions of various sources, we present the test results in four distinct modes. 1) Our system with RTK and ICP; 2) Our system with RTK and CDLPE; 3) Our system with GPS and ICP; 4) Our system with GPS only and CDLPE. In Table II ,we show the quantitative results in both regular and weak GNSS roads. Meanwhile, in order to display the evaluation accuracy more intuitively, we have randomly selected a continuous 1200-second data segment from the regular scenes. The Figure 6 illustrates the respective results of the lateral and longitudinal errors of the above four methods when the sampling rate is reduced from 20Hz to 1Hz.

### C. Ablation Study

The experimental data in the table shows that our algorithm is very robust under both ordinary GPS and weak GNSS scenarios, which is the benefit of our multi-dimensional estimation. In order to further illustrate the role of environment understanding, we have separately designed the following experiments based on ordinary GPS: 1) Our system with CDLPE excluding geometry matching. 2) Our system with CDLPE excluding feature matching. 3) Our system with CDLPE. We tested on the Beijing dataset, uniformly sampled 1000 points from the error distribution of the results of each method, and plotted them in two dimensions. As Figure 7 shows, it is clear that our method effectively integrates information and models from different dimensions.

| Scenes | Method | Lateral RMS(m) | | Longitudinal RMS(m) | | Yaw(deg) | Ego Lane |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | max | mean | max | mean | mean | Accuracy |
| Regular | RTK + ICP | 4.942 | 0.544 | 4.863 | 0.871 | 1.740 | 0.9530 |
| | RTK + CDLPE | 0.746 | 0.175 | 1.794 | 0.414 | 0.692 | 0.9965 |
| | GPS + ICP | 7.540 | 0.746 | 5.008 | 1.424 | 2.658 | 0.9312 |
| | GPS + CDLPE | 1.598 | 0.213 | 2.379 | 0.631 | 0.978 | 0.9877 |
| Weak GNSS | RTK + ICP | 11.214 | 1.565 | 13.621 | 1.424 | 3.138 | 0.8426 |
| | RTK + CDLPE | 1.174 | 0.289 | 4.658 | 1.296 | 1.025 | 0.9641 |
| | GPS + ICP | 12.057 | 1.947 | 15.362 | 1.876 | 4.357 | 0.8206 |
| | GPS + CDLPE | 1.975 | 0.485 | 5.680 | 1.684 | 1.687 | 0.9394 |

TABLE II: The comparison of different localization methods based on pose and lane accuracy.
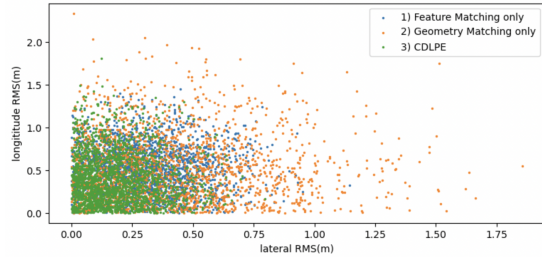


((a)) Distribution of 2-dimensional RMS Error
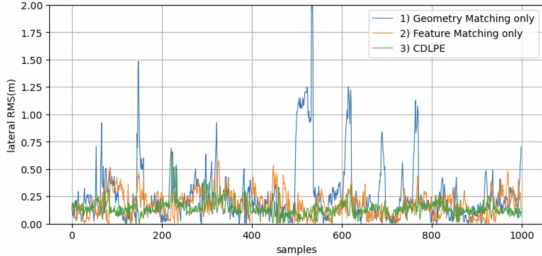


((b)) Lateral RMS Error

Fig. 7: Errors of different modules in the ablation experiment.

### D. Qualitative Results

An accurate localization ensures that the projection of map elements on image is completely consistent with the semantic perception. Projection results in different scenes in test set are shown in Figure 8. Sub-figures d, e, f depict a typical turning process, where longitudinal errors become lateral ones and GNSS has a margin of error. Our algorithm ensures a smooth and accurate correction at intersections, preventing wrong lane placements, by utilizing matching result in multiple dimensions. Sub-figures g and h represent weak GNSS signal scenarios like tunnels and underpasses. Despite GNSS drift (light blue block), our stable fused result (green block) is achieved through scene understanding and map-matching verification.

### V. CONCLUSION AND FUTURE WORKS

In this paper, we present a comprehensive, end-to-end, vision-based vehicle localization framework. Our framework is further bolstered by a novel map-matching algorithm we've developed, named Cross-Dimensional Lane and Pose Estimator (CDLPE). We have showcased the reliability of our algorithm in a variety of challenging driving conditions. Furthermore, our high level of accuracy and resilience to weak signal interference has been confirmed through experimental comparisons of pose and lane accuracy under varying GNSS signal strengths. Lightweight maps often have limitations in terms of absolute accuracy. In future work, it is necessary for the algorithm to estimate these mapping errors in real-time for precise positioning.

### REFERENCES

[1] Rabbia Asghar, Mario Garzón, Jérôme Lussereau, and Christian Laugier. Vehicle localization based on visual lane marking and topological map matching. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 258–264, 2020. doi: 10.1109/ICRA40945.2020.9197543.

[2] P.J. Besl and Neil D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992. doi: 10.1109/34.121791.

[3] Kyoungtaek Choi, Jae Suhr, and Ho Gi Jung. In-lane localization and ego-lane identification method based on highway lane endpoints. *Journal of Advanced Transportation*, 2020:1–16, 02 2020. doi: 10.1155/2020/8684912.

[4] Liuyuan Deng, Ming Yang, Bing Hu, Tianyi Li, Hao Li, and Chunxiang Wang. Semantic segmentation-based lane-level localization using around view monitoring system. *IEEE Sensors Journal*, 19(21):10077–10086, 2019. doi: 10.1109/JSEN.2019.2929135.

[5] Xingshuai Dong and Massimiliano L. Cappuccio. Applications of computer vision in autonomous vehicles: Methods, challenges and future directions, 2023.

[6] Chengcheng Guo, Minjie Lin, Heyang Guo, Pengpeng Liang, and Erkang Cheng. Coarse-to-fine semantic localization with hd map for autonomous driving in structural scenes, 2021.

((c)) normal urban road     ((d)) express way     ((e)) traffic jam

((f)) turning     ((g)) turning     ((h)) turning

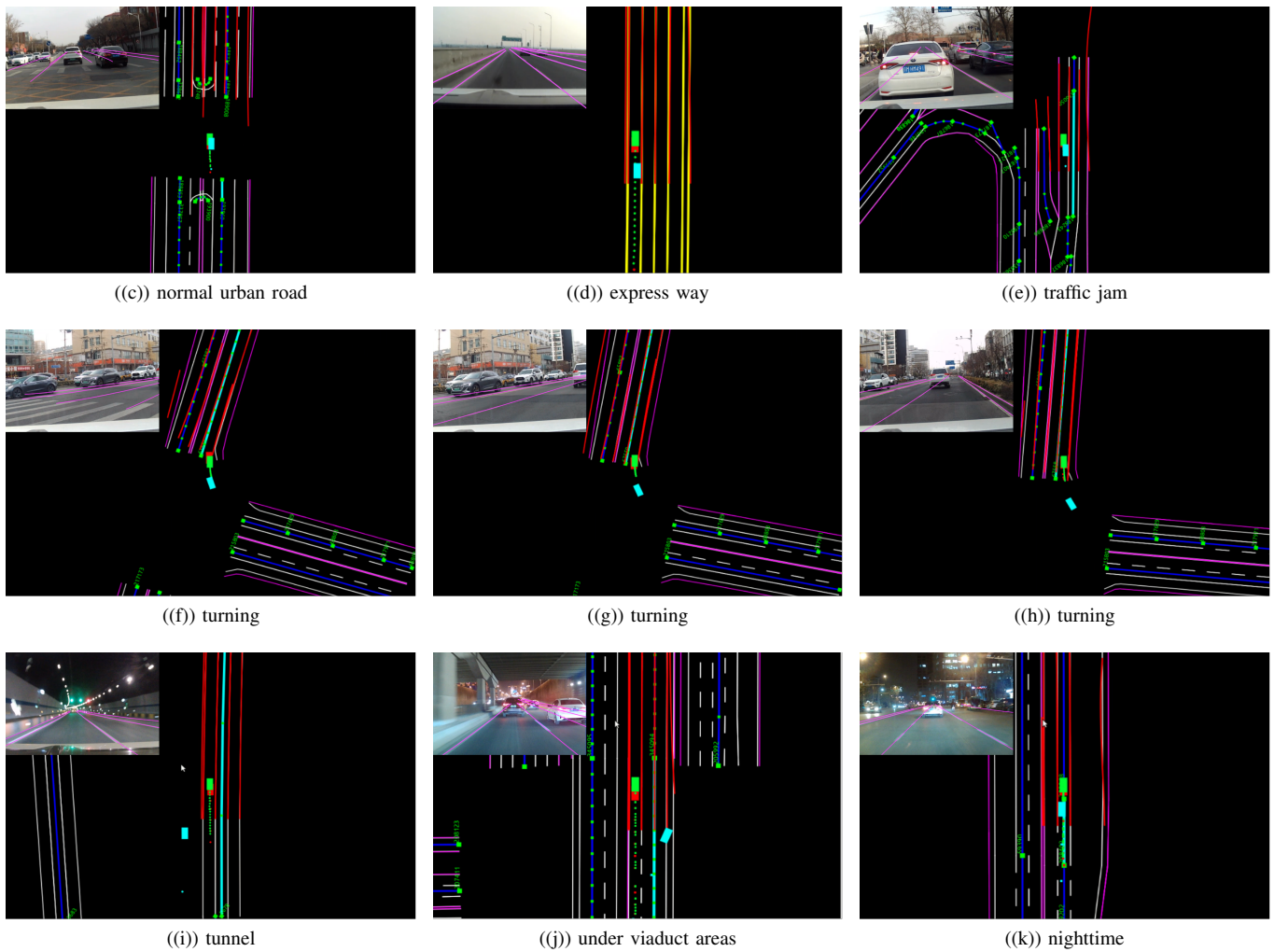((i)) tunnel     ((j)) under viaduct areas     ((k)) nighttime

Fig. 8: The top left corner depicts the projected lane lines on the image. The figures display the simulation interface of the algorithm within the global coordinate system. In this system, the lane lines are denoted in red, the final positioning result is indicated by the green block, the RTK result is represented by the sky-blue block, and the map-matching result is shown by the red block.

[7] Toni Heidenreich, Jens Spehr, and Christoph Stiller. Laneslam – simultaneous pose and lane estimation using maps with lane-level accuracy. pages 2512–2517, 09 2015. doi: 10.1109/ITSC.2015.404.

[8] Chih-Yuan Hsu and Nong-Hong Lin. A visual slam based-method for vehicle localization. Technical report, SAE Technical Paper, 2024.

[9] Abderrahim Kasmi, Johann Laconte, Romuald Aufrere, Dieumet Denis, and Roland Chapuis. End-to-end probabilistic ego-vehicle localization framework. *IEEE Transactions on Intelligent Vehicles*, 6(1):146–158, 2021. doi: 10.1109/TIV.2020.3017256.

[10] Julius Kümmerle, Marc Sons, Fabian Poggenhans, Tilman Kühner, Martin Lauer, and Christoph Stiller. Accurate and efficient self-localization on roads using basic geometric primitives. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 5965–5971,

2019. doi: 10.1109/ICRA.2019.8793497.

[11] Dongfang Liu, Yiming Cui, Xiaolei Guo, Wei Ding, Baijian Yang, and Yingjie Chen. Visual localization for autonomous driving: Mapping the accurate location in the city maze. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 3170–3177, 2021. doi: 10.1109/ICPR48806.2021.9411961.

[12] Yan Lu, Jiawei Huang, Yi-Ting Chen, and Bernd Heisele. Monocular localization in urban environments using road markings. In *2017 IEEE Intelligent Vehicles Symposium (IV)*, pages 468–474, 2017. doi: 10.1109/IVS.2017. 7995762.

[13] Raul Mur-Artal and Juan D. Tardos. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE Transactions on Robotics*, 33(5): 1255–1262, October 2017. ISSN 1941-0468. doi: 10. 1109/tro.2017.2705103. URL http://dx.doi.org/10.1109/

TRO.2017.2705103.

[14] Abhijeet Nayak, Daniele Cattaneo, and Abhinav Valada. Ralf: Flow-based global and metric radar localization in lidar maps. *ArXiv*, abs/2309.09875, 2023. URL https://api.semanticscholar.org/CorpusID:262045190.

[15] Fabian Poggenhans, Jan-Hendrik Pauls, Johannes Janosovits, Stefan Orf, Maximilian Naumann, Florian Kuhnt, and Matthias Mayr. Lanelet2: A high-definition map framework for the future of automated driving. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 1672–1679, 2018. doi: 10.1109/ITSC.2018.8569929.

[16] Tong Qin, Yuxin Zheng, Tongqing Chen, Yilun Chen, and Qing Su. A light-weight semantic map for visual localization towards autonomous driving. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 11248–11254, 2021. doi: 10.1109/ICRA48506.2021.9561663.

[17] S. Rusinkiewicz and M. Levoy. Efficient variants of the icp algorithm. In *Proceedings Third International Conference on 3-D Digital Imaging and Modeling*, pages 145–152, 2001. doi: 10.1109/IM.2001.924423.

[18] Paul-Edouard Sarlin, Daniel DeTone, Tsun-Yi Yang, Armen Avetisyan, Julian Straub, Tomasz Malisiewicz, Samuel Rota Bulo, Richard Newcombe, Peter Kontschieder, and Vasileios Balntas. Orienternet: Visual localization in 2d public maps with neural matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21632–21642, 2023.

[19] Markus Schreiber, Carsten Knöppel, and Uwe Franke. Laneloc: Lane marking based localization using highly accurate maps. In *2013 IEEE Intelligent Vehicles Symposium (IV)*, pages 449–454, 2013. doi: 10.1109/IVS.2013.6629509.

[20] Jae Suhr, Jeungin Jang, Daehong Min, and Ho Gi Jung. Sensor fusion-based low-cost vehicle localization system for complex urban environments. *IEEE Transactions on Intelligent Transportation Systems*, 18:1–9, 08 2016. doi: 10.1109/TITS.2016.2595618.

[21] Linus Svärm, Olof Enqvist, Fredrik Kahl, and Magnus Oskarsson. City-scale localization for cameras with known vertical direction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(7):1455–1461, 2017. doi: 10.1109/TPAMI.2016.2598331.

[22] Sebastian Thrun, Wolfram Burgard, and Dieter Fox. *Probabilistic robotics.* Intelligent robotics and autonomous agents. MIT Press, 2005. ISBN 978-0-262-20162-9.

[23] Huayou Wang, Changliang Xue, Yanxing Zhou, Feng Wen, and Hongbo Zhang. Visual semantic localization based on hd map for autonomous vehicles in urban scenarios. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 11255–11261, 2021. doi: 10.1109/ICRA48506.2021.9561459.

[24] Daniel Wilbers, Christian Merfels, and Cyril Stachniss. Localization with sliding window factor graphs on third-party maps for automated driving. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 5951–5957, 2019. doi: 10.1109/ICRA.2019.8793971.

[25] Zhichen Wu, Jianda Li, Jiadi Yu, Yanmin Zhu, Guangtao Xue, and Minglu Li. L3: Sensing driving conditions for vehicle lane-level localization on highways. In *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, pages 1–9, 2016. doi: 10.1109/INFOCOM.2016.7524436.

[26] Xin Xia, Neel P. Bhatt, Amir Khajepour, and Ehsan Hashemi. Integrated inertial-lidar-based map matching localization for varying environments. *IEEE Transactions on Intelligent Vehicles*, 8(10):4307–4318, 2023. doi: 10.1109/TIV.2023.3298892.

[27] Zhongyang Xiao, Kun Jiang, Shichao Xie, Tuopu Wen, Chunlei Yu, and Diange Yang. Monocular vehicle self-localization method based on compact semantic map. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 3083–3090, 2018. doi: 10.1109/ITSC.2018.8569274.

[28] Zhongyang Xiao, Diange Yang, Tuopu Wen, Kun Jiang, and Ruidong Yan. Monocular localization with vector hd map (mlvhm): A low-cost method for commercial ivs. *Sensors*, 20(7), 2020. ISSN 1424-8220. doi: 10.3390/s20071870. URL https://www.mdpi.com/1424-8220/20/7/1870.

[29] Liqi Yan, Yiming Cui, Yingjie Chen, and Dongfang Liu. Hierarchical attention fusion for geo-localization, 2021.

[30] Wei Yan, Ning Xiao, Pan Jiang, Hongkai Wang, Yilong Yuan, Liang Lin, and Chang Liu. Ego lane estimation using visual information and high definition map. In *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 603–608, 2023. doi: 10.1109/PLANS53410.2023.10140029.

[31] Bernhard Zeisl, Torsten Sattler, and Marc Pollefeys. Camera pose voting for large-scale image-based localization. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 2704–2712, 2015. doi: 10.1109/ICCV.2015.310.

[32] Haihan Zhang, Chun Xie, Hisatoshi Toriya, Hidehiko Shishido, and Itaru Kitahara. Vehicle localization in a completed city-scale 3d scene using aerial images and an on-board stereo camera. *Remote Sensing*, 15(15), 2023. ISSN 2072-4292. doi: 10.3390/rs15153871. URL https://www.mdpi.com/2072-4292/15/15/3871.

[33] Ji Zhang and Sanjiv Singh. Loam : Lidar odometry and mapping in real-time. *Robotics: Science and Systems Conference (RSS)*, pages 109–111, 01 2014.

[34] Xingxing Zuo, Patrick Geneva, Yulin Yang, Wenlong Ye, Yong Liu, and Guoquan Huang. Visual-inertial localization with prior lidar map constraints. *IEEE Robotics and Automation Letters*, 4(4):3394–3401, 2019. doi: 10.1109/LRA.2019.2927123.