International Co-operation on Theories and
Concepts in Traffic Safety

Towards Safer Roads: A Data-Driven Ground-Aerial Fusion Approach for Enhanced Ego-Localization

[1]*Ben Zekri, A.E., [2]Bahmanyar, R. & [3]Chaabouni-Chouayakh, H.
*lead presenter
[1] alaa.benzekri@dlr.de, German Aerospace Center (DLR), Remote Sensing Technologies Institute, Germany
[2] German Aerospace Center (DLR), Remote Sensing Technologies Institute, Germany
[3] Digital Research Center of Sfax, Sm@rts Laboratroy, Tunisia

Accurate localization is crucial for autonomous systems to ensure safe and effective operation on modern roadways. While consumer-grade GPS sensors are widely used, they often fall short due to limitations imposed by satellite geometry, atmospheric disturbances, and multipath interference. These factors can result in errors of several meters, a discrepancy that is unacceptable for the centimeter-level accuracy required for autonomous driving. Inaccurate positioning not only affects navigation, but also increases the risk of collisions, thus undermining road safety.

To overcome these challenges, autonomous vehicles are now integrating a variety of sensors such as inertial measurement units, LiDAR, radar, and computer vision to provide a more reliable picture of the environment. These systems combine data streams to improve overall accuracy and resilience to individual sensor failures. Filtering algorithms, including Kalman and particle filters, have been instrumental in refining localization estimates. Computer vision techniques such as visual odometry further mitigate scenarios. More recently, deep learning approaches have emerged as powerful tools capable of learning complex mappings between sensor inputs and precise location outputs. These advances not only enhance autonomous navigation, but also contribute to road safety by reducing the risk of accidents and promoting smoother interactions among all road users.

As research continues to evolve, the development of comprehensive, high-accuracy datasets remains critical to training adaptive models that can thrive in diverse driving conditions. Existing simulators and data collection platforms often lack the centimeter-level accuracy, environmental diversity, and temporal dynamics required to train robust models. Improving these datasets is also essential to fully exploit emerging methods and ultimately pave the way for safer, more reliable autonomous navigation systems.

In this work, we develop a robust solution for ego-localization. We propose to create a carefully curated dataset that overcomes current shortcomings by providing centimeter-level GPS annotations, diverse scenarios, and rich temporal dynamics. In addition, we design an innovative architecture that fuses ground-view images with rough GPS localization data to accurately predict true positions.

To generate the ego-localization dataset, we use the CARLA simulator. We equip each simulated vehicle with six RGB cameras, a Bird's Eye View (BEV) camera for semantic segmentation, and a GPS sensor. Vehicle data is captured from multiple perspectives and locations. In addition, we collect aerial imagery by moving a camera with a GPS sensor through different environments while obtaining fine-grained semantic segmentation annotations.

Our localization correction approach consists of three modules: BEV mapping, aerial image segmentation, and localization error estimation.

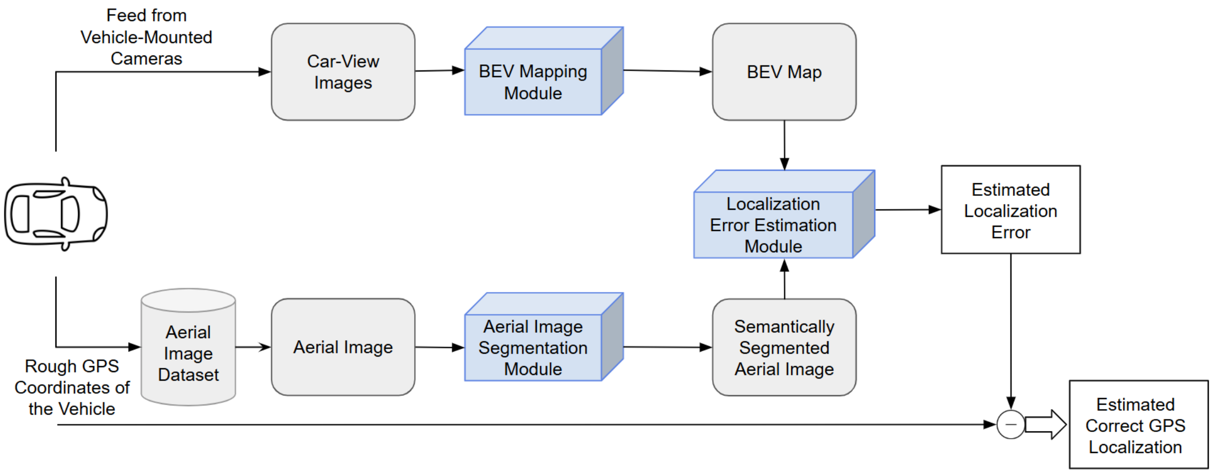

Figure 1 Overview of our localization correction approach

A segmentation BEV map of the vehicle's environment is generated using six images from the vehicle's cameras. An aerial image centered on the vehicle's GPS position is processed to create an aerial segmentation map. The two segmentation maps are then matched to estimate the true location of the vehicle by determining the center of the BEV map on the aerial segmentation map. Assuming accurate geo-registration of the aerial images, the geo-coordinates of the estimated point represent the true vehicle position. Finally, the distance between this estimated point and the GPS-derived center of the aerial image quantifies the localization error.

We expect the proposed approach to achieve high sub-meter localization accuracy under varied simulated urban conditions, providing a robust benchmark for evaluating safety-critical localization tasks. Initial experiments will be conducted on the synthetic dataset to validate performance across diverse environments. Subsequently, real-world samples will assess the model's ability to generalize beyond simulation. This sim-to-real evaluation will determine how effectively the system supports accurate and reliable localization under real driving conditions—a prerequisite for safer navigation and reduced collision risks.

This work presents a novel localization correction approach that fuses ground-view and aerial data to significantly improve positioning accuracy. By reducing GPS-related errors, especially in urban environments, the system enhances the reliability of autonomous navigation—a key enabler of traffic safety.

Keywords: Ego-localization, Sensor fusion, Aerial imagery, BEV mapping, Traffic safety