Perception in Snow-covered Environments

Leah Chowenhill
lchowenh@andrew.cmu.edu

Swathi Jadav
sjadav@andrew.cmu.edu

Lakshay Virmani
lvirmani@andrew.cmu.edu

1 Summary

This project shall seek to address the current shortcomings of perception in snow-covered environments. While autonomous vehicles can easily detect road boundaries in clear and unobstructed conditions, this functionality becomes difficult in degraded environments. Such environments are those with low visual contrast or airborne obscurities like falling snow. We propose a multi-head cross attention network utilizing fused RGB images and LiDAR point clouds to detect drivable paths on snow-covered roads. We hypothesize that this model shall perform better than a convolutional neural network (CNN) utilizing fused RGB images, LiDAR, and radar data. The multi-head cross attention model shall be trained using the DENSE dataset. If time permits, we shall also seek to implement an ensemble method which detects both drivable paths and objects. Finally, we shall apply this model to the Canadian Adverse Driving Conditions Dataset (CADC) [3] to assess its robustness.

2 Motivation

Snow and other Winter weather conditions make it especially challenging for autonomous vehicles to accurately "see" the roadway and surrounding objects. In these conditions, traditional visual cues for perception - such as lane markers or road reflectors - are not visible and cannot be reliably utilized. Moreover, sensor accuracy is decreased by snow. For example, monocular cameras experience decreased visibility and contrast, as well as increased glare. Additionally, signal scattering, attenuation, and absorption inhibit the function of LiDAR sensors. While radar sensing is unaffected by snow, this data alone cannot be used to correctly classify objects. Therefore, it is necessary to determine methods through which the usefulness of camera images and LiDAR point clouds can be increased when occlusions like snow are present.

3 Goals

The key goals of this project are as follows:

1. Drivable path detection [4]: Provide the ego vehicle with drivable roadway boundaries
2. Object detection and tracking (extra credit): Identify and track objects in the vicinity of the ego vehicle
3. Annotate Canadian Adverse Driving Conditions Dataset (CADC) with drivable path labels.
4. Ensure the system is robust to:
    - Snow conditions - Actively snowing vs previously snowed, small vs large amounts of accumulated snow on the road, etc.
    - Lighting conditions - Day vs night, clear vs fog, etc.
5. Ensure the system is efficient and meets the ego vehicle demands of near real-time decision making.

4 Use Cases

The use cases of perception in snow-covered environments can be broken down into three categories, as follows:

1. **Safety**
   In the United States (U.S.), inclement weather is the fifth most common cause of accidents. In fact, the Federal Highway Administration reports that approximately 900 people are killed and 76,000 people are injured annually by automotive accidents in snow or sleet [3]. Autonomous vehicles which can navigate winter weather conditions could decrease the number of accidents each year due to snow, thereby preventing death and bodily injury of vehicle passengers or vulnerable road users (VRUs).
2. **Feasibility**
   While autonomous vehicles continue to fail in snowy conditions, their deployment will be geographically limited. Approximately 70% of the lower 48 states in the U.S. receive snow each year, making these areas inaccessible by current autonomous vehicle technology [3]. This hurdle is further emphasized by the current availability of autonomous vehicle services such as Cruise and Waymo, which offer services only in warm-weather climates like San Francisco, CA and Phoenix, AZ. Enabling reliable and accurate perception in snow can increase the geographical range in which autonomous driving is feasible.
3. **Industry**
   Snow removal is a difficult and costly undertaking for most government bodies. The Federal Highway Administration estimates that the nation spends upwards of $4 billion dollars annually on snow and ice removal [3]. Additionally, in recent years, states have struggled to find enough people who are able and willing to become snowplow operators. Operating plow trucks and salt spreaders is not only dangerous work, but it often requires a Class C driver's license [4]. Therefore, introducing autonomous driving to the snow removal industry could be more cost-effective and efficient than utilizing human drivers.

5 Data Sets

5.1 DENSE

The DENSE dataset [1] was created by Ulm University in Baden-Württemberg, Germany. This dataset contains LiDAR, radar, and RGB monocular camera data which captures the driving environment in all weather conditions, as well as in both daytime and nighttime. Moreover, the DENSE dataset uses the HDL64-S3 LiDAR, which has 64 channels, and contains a total of 13,000 frames [2]. The DENSE dataset is labeled with drivable path boundaries, to be used in determining roadway edges.

5.2 Canadian Adverse Driving Conditions Dataset (CADC)

The CADC [3] was created by the University of Waterloo in partnership with the University of Toronto, Canada. This dataset contains RGB monocular camera and LiDAR data that captures snowy driving environments in both daytime and nighttime. This dataset contains 7,000 frames and is labeled to detect ten classes of objects such as cars, pedestrians, busses, bicycles, traffic guidance objects, and animals [1]. As the CADC uses a VLP-32C LiDAR, which has only 32 channels, the DENSE dataset will need to be downsampled before being used for model training so that the same model can be applied to the CADC dataset without error.

6 Methodology

6.1 Inputs

6.1.1 RGB Image

We'll be using images from front-facing camera of the ego vehicle to detect drivable regions as in [5]. We believe that a network should be able to learn the relationship between road and the visual cues offered by objects on its sides like poles, curbs, vegetation etc.

6.1.2 LIDAR

We believe that in instances where the visual cues aren't enough to identify the drivable path, LIDAR data will be of help. It can help boost the confidence about objects on the road and in the vicinity. Also, we think that in many cases, since there is a difference in the depth of the road and the surrounding areas, a model trained on LIDAR data of an ego vehicle should be able to infer this and predict the regions in which the ego vehicle can drive.

6.2 Experiments

6.2.1 RGB

We'll start with building a model which only used RGB data to predict the drivable regions and the objects in the vicinity of the ego vehicle.

6.2.2 LIDAR

We'll experiment with another model which only uses the LIDAR data. The results from these two experiments should help us understand the pros and cons of using each.

6.2.3 RGB + LIDAR : CNN based fusion

We'll then proceed with fusing both the RGB and LIDAR data in order to build upon the strenghts of both and have them encounter each others weaknesses. We'll start our fusion experiment with a simple CNN based fusion network. The network comprises of a Encoder network to extract features from both the sensor modalities. The extracted features are then fused and passed to another decoder network to predict image segmentation of the drivable path.

6.3 RGB + LIDAR : Multi-headed cross-attention based fusion

In this method, we propose to use a Multi-head cross-attention network to fuse the sensor modalities to enhance the performance and robustness of the module. The CNN-based approach does not mix visual and LiDar information. Theoretically, we believe that the network will infer a better mapping between the 3D LiDar point cloud and the 2D image segmentation if the features are coupled through the use of a Muti-head cross-attention network. Figure 1 explains how the two modalities interact with each other through the network. As a proof of concept, a similar idea has been used in [2] for the prediction of bounding boxes of objects. We believe that the same idea should be able to improve the performance and robustness of our sensor fusion model.

6.4 Training phases

6.4.1 Real world non autonomous driving data

We'll start with models which are trained on real world data that are not related to autonomous driving. This will help us use transfer learning in order to build a model which generalizes better.

6.4.2 Autonomous driving data with clear weather conditions

We then plan to fine-tune these models with autonomous driving data comprising of clear weather conditions. This should help the model easily pick up on cues which might not be easy to detect in adverse weather.

6.4.3 Autonomous driving data with snow and adverse weather conditions

Finally, we'll conduct another phase of finetuning which only includes the autonomous driving data with snow and adverse weather conditions.

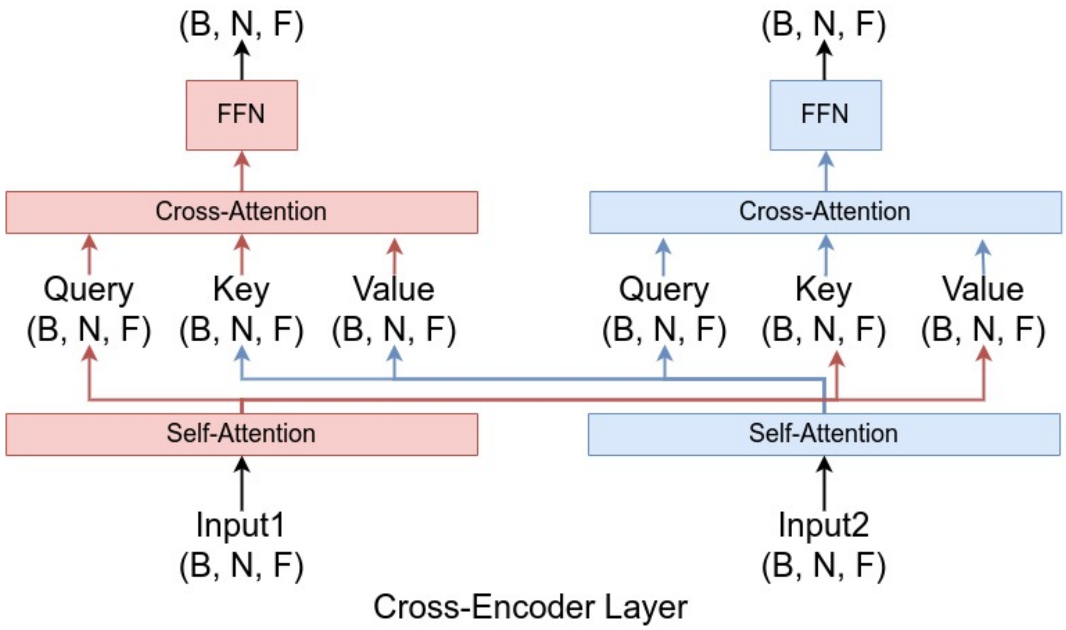

Figure 1: Multi-Head Cross Attention Mechanism

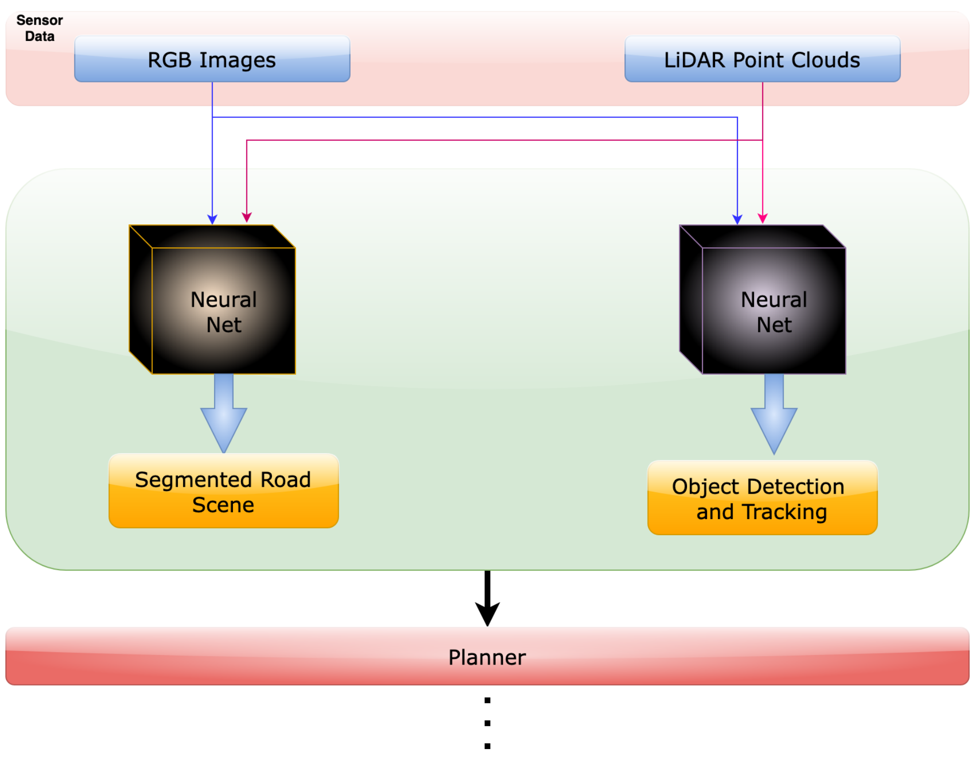

Figure 2: Perception in snow covered environments

7 System Design

8 Development Milestones

The development schedule for this project is outlined in 3 below.

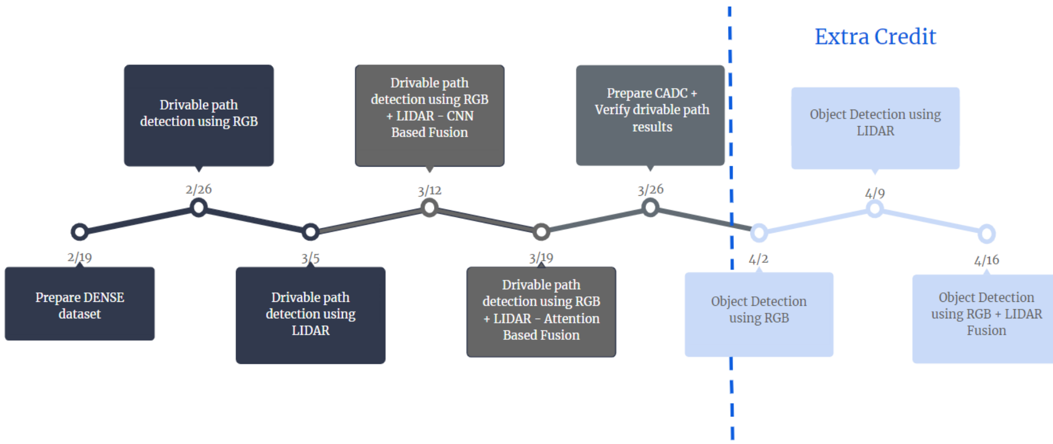

Figure 3: Development Milestones

9 Demonstration Sequences

The demonstration sequences for this project are outlined in 4 below.

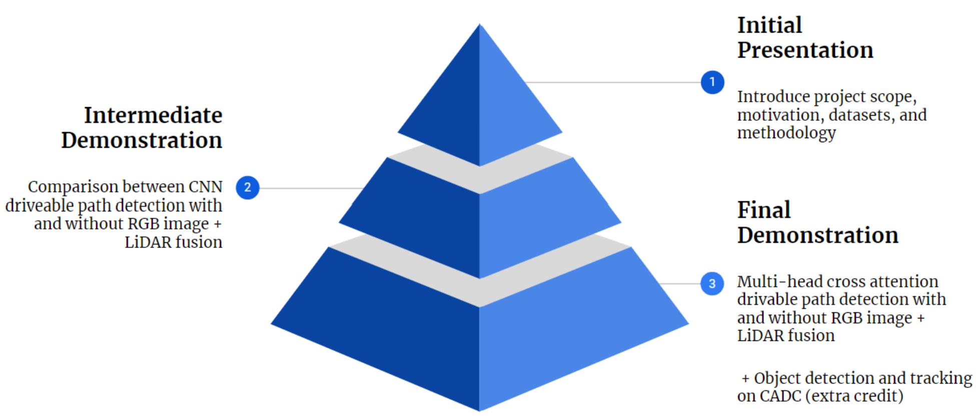

Figure 4: Demonstration Sequences

10 Work Partitioning

1. **Leah Chowenhill**
   (a) Implement CNN model using RGB images + LiDAR
   (b) Add multi-headed cross attention to model
   (c) Apply CNN model to CADC dataset
   (d) Fuse semantic segmentation + object detection on CADC dataset

2. **Swathi Jadav**
   (a) Downsample DENSE dataset
   (b) Implement CNN model using DENSE RGB images
   (c) Implement CNN model using DENSE RGB images
   (d) Add multi-headed cross attention to model
   (e) Implement object detection model using CADC RGB images
3. **Lakshay Virmani**
   (a) Implement CNN model using LiDAR
   (b) Add multi-headed cross attention to model
   (c) Implement object detection model using CADC LiDAR
   (d) Fuse semantic segmentation + object detection on CADC dataset

11 Conclusion

In conclusion, through this Autonomous Driving project, we would like to explore the power and potential of deep learning and artificial intelligence in improving sensor fusion techniques. We would like to showcase the ability of deep learning-based sensor fusion techniques to accurately detect drivable areas under adverse weather conditions and also detect objects and track objects to enhance the driver's safety in bad weather environments.

Through the development of this project, we would like to gain a deeper understanding of the various sensor fusion techniques and their applications in real-world scenarios. We would also like to discover the importance of choosing the right algorithms, and data sets to achieve optimal performance and accuracy.

Overall, this Autonomous Driving project has been an enriching experience that has allowed us to push the boundaries of what is possible with modern technology. With further research and development, we look forward to unlocking the full potential of learning-based sensor fusion techniques and creating a better autonomous driving experience in adverse weather conditions.

References

[1] M. Bijelic, T. Gruber, F. Mannan, F. Kraus, W. Ritter, K. Dietmayer, and F. Heide. Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather. In *The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.

[2] G. K. Erabati and H. Araujo. Msf3ddetr: Multi-sensor fusion 3d detection transformer for autonomous driving, 2022.

[3] M. Pitropov, D. E. Garcia, J. Rebello, M. Smart, C. Wang, K. Czarnecki, and S. Waslander. Canadian adverse driving conditions dataset. *The International Journal of Robotics Research*, 40(4-5):681–690, dec 2020.

[4] N. Rawashdeh, J. Bos, and N. Abu-Alrub. Drivable path detection using cnn sensor fusion for autonomous driving in the snow. page 5, 04 2021.

[5] T. Yasuno, J. Fujii, H. Sugawara, and M. Amakata. Road surface translation under snow-covered and semantic segmentation for snow hazard index, 2021.

Appendix