Chapter 6

Conclusion and Future Work

This chapter concludes the proposed works in the previous thesis chapter with the significant contributions and findings. This chapter also gives future research direction for researchers by explaining open research issues in the field of different modules of AVs.

6.1 Conclusions

Enabling computers to understand images and videos in real scenarios is a crucial step toward developing intelligent systems. Intelligent systems like robots, Autonomous Vehicles etc., have various applications in industries (Farmhouse, construction, military), Transportation (shuttle, mass transits), and consumers (Commute, task, shared and retail).

Focus of this thesis is to develop some effective algorithms for different modules of autonomous vehicles. These developed methods should be as robust as possible with the appearance of different objects in crowd scenarios. This thesis proposes novel methods using deep-learning approaches for different modules of autonomous vehicles, followed by their testing and evaluation compared to other state-of-the-art methods. The contributions in the thesis tried to overcome the challenges faced by previously reported object detection, object tracking, trajectory prediction and motion planning module of AVs.

Chapter 1 of the thesis introduces autonomous vehicles, their modules and sub-modules. Further, this chapter also defines the motivation behind the research work. This chapter also lists the objectives of the thesis and its contributions of the thesis.

Chapter 2 discusses the theoretical background for various sub-modules of AVs. This chapter has covered an overview of deep learning and machine learning approaches, and further, in this chapter, a literature survey of prominent approaches for object detection, object tracking, trajectory prediction and motion planning. Issues and challenges of the fields were also mentioned. The benchmark databases for submodules of AVs and the performance metrics were also described in detail.

Chapter 3 proposed a model for the object detection. The model is a channel-spatial attention-based method for object detection. The attention is used to learn the semantic features. The model is focused on giving faster and more accurate object detection. The impact of channel-spatial attention has been evaluated on two object detection datasets, namely KITTI and BDD. From the experimental results, it can be observed that the result obtained is comparable to the state-of-the-art models.

Chapter 4 proposed a deep-learning-based multi-object tracking model. The first one is used an end-to-end model by calculating the relative scale of objects for object re-identification. The proposed model tries to overcome the challenges of state-of-the-art techniques. Experimental analysis has been tested on the most promising datasets, such as MOT and WAYMO. The performance of the proposed model is compared with state-of-the-art methods, which shows that the presented approach is better.

Chapter 5 proposed two methods, one for trajectory Prediction and one for motion planning of Autonomous vehicles. This chapter also proposes a backbone network with a spatial-temporal graph Neural network for the TP. Multi-scale spatial features with an LSTM network achieve significant prediction results on three publicly available datasets Argoverse, Lyft and Apolloscape. The second has presented a multi-view image-based approach that considers three images from different angles to understand the ego vehicle's surroundings and PID controller to operate the motion using steering, throttle and speed. The proposed model has achieved more state-of-the-art results than the other existing methods.

6.2 Suggestions for Future Research

The research work presented in this thesis can be taken further in different directions; the scope for future work is as follows:

- Further, new deep-learning based models can be proposed to benchmark the datasets.
- Some other modalities may also be explored for accuracy and efficiency for autonomous vehicles to avoid collisions and abide by the traffic rules.
- Sensor fusion can encompass more meaningful information for autonomous vehicles.
- Explore the models for dense crowds and streets (Narrow paths).
- Can perform hardware implementation of the models and methods.

6.3 Future Work

Despite the above-proposed models, the researchers can explore the methods for the other submodules of the AVs such as Lane detection, traffic light and sigh detection, behavior prediction, trajectory planning etc. Adding these submodules will make them more efficient and secure navigation for the AVs. Some existing databases have a very sparse crowd, like Lyft; some have only day conditions. Generating a database that captures nearly all the scenarios can help to create better models.

Heterogeneity in objects is still not entirely resolved—a study of mechanisms for better separation of background and foreground in cases where they are similar. The existing methods have been categorized into various individual tasks for the AVs; A combined approach must be developed for a complete task of the AVs that can simulate with hardware efficiently for the navigation.

The proposed model can be extended and improved using feature extraction from better deep learning models. The deep learning models can be developed using prior information from traditional models. Many improvements in the proposed image annotation model are possible, in terms of better features to be extracted and superior models.