
Self-Supervised Steering Angle Prediction for Vehicle Control Using Visual Odometry

Qadeer Khan^{1,2}

Patrick Wenzel^{1,2}

Daniel Cremers^{1,2}

¹ Technical University of Munich

² Artisense

{qadeer.khan, patrick.wenzel, cremers}@tum.de

Abstract

Vision-based learning methods for self-driving cars have primarily used supervised approaches that require a large number of labels for training. However, those labels are usually difficult and expensive to obtain. In this paper, we demonstrate how a model can be trained to control a vehicle’s trajectory using camera poses estimated through visual odometry methods in an entirely self-supervised fashion. We propose a scalable framework that leverages trajectory information from several different runs using a camera setup placed at the front of a car. Experimental results on the CARLA simulator demonstrate that our proposed approach performs at par with the model trained with supervision.

1 Introduction

The success of current learning-based approaches on various tasks can be attributed to the availability of a large amount of labeled data. In most circumstances, data is being generated at a rate that is far superior than it is humanly possible to annotate it. Even if possible, it would be expensive, time-consuming, and prone to errors. Vision-based sensorimotor control tackles the following problem: Given an image captured by a camera, what should be the corresponding steering command? Controlling a car using supervised steering labels has shown fairly good performance [Bojarski et al., 2016, Chen et al., 2015, Codevilla et al., 2018b, Muller et al., 2006]. However, this performance is mostly limited to the domain in which the data was collected [Wenzel et al., 2018]. In order to generalize across a variety of conditions, a car fitted with drive-

by-wire needs to be driven by an *expert* driver for all such possible conditions. This is an in-feasible, inefficient, and expensive solution for the following reasons:

- Dynamic conditions such as lighting and weather are beyond human control. Hence, data collection for a specific scenario cannot be pre-planned.
- The trained model may overfit to the specific position at which the camera is placed. Any deviation, in the position of the camera at test time, may lead to deteriorated performance.
- Collection of supervised data requires drive by wire cars. Since many vehicles driven on the street by drivers do not have this option, we lose out on bootstrapping this huge treasure of valuable data.

It may be argued that modern cars have the CAN bus which may be used to extract out important information including steering commands which can then be used for supervision. However, common people are not expected to modify their vehicle, if even allowed by the law of the country they reside in. Moreover, self-modifications also have implications on the warranty/insurance premiums. Therefore, we have referred to an *expert* as a dedicated driver who has the necessary license and permits from the relevant traffic and regulatory authorities to drive a modified vehicle on public roads. Traditional people are not expected to hold these permits.

In contrast to supervised approaches, an alternative would be to deploy reinforcement learning (RL) based techniques, wherein instead of collecting expert supervised labels, a reward function is defined [Abbeel et al., 2007, Abbeel and Ng, 2004, Lillicrap et al., 2015]. The model learns a policy by a hit and trial method using the reward function as an indicator for updating the model’s weights. However, this is an extremely data-inefficient solution as it requires far more sessions to train a model when compared with supervised techniques. Moreover, car manufacturers would be unwilling to use an unstable learning-based

method for their autonomous functions. Recently, the authors of [Kendall et al., 2019] demonstrated an example of using deep RL for training a policy for maintaining the car on the road in a controlled environment. However, this approach requires an expert driver to intervene for any correction whenever the car attempts to deviate away from the road.

In this paper, we propose a self-supervised approach which that does not suffer from the limitations of the supervised and RL-based approaches as described earlier. We use a camera setup, that can be placed at the front of any car model. Ideally, we would like to retrieve the precise ground truth poses of the cameras as the car moves along its trajectory. This could be done using precise GNSS positions or through localization against landmarks. However, acquiring accurate camera poses using these methods may not always be possible such as when a car enters a tunnel or drives through an urban canyon due to the unavailability of GNSS signals or landmarks [Weiss et al., 2011]. Therefore, as an alternative, we demonstrate that using visual odometry (VO) methods [Engel et al., 2017, Mur-Artal et al., 2015, Wang et al., 2017] to estimate the camera poses in a self-supervised manner can yield similar performance as with supervised labels. Note that, the VO methods can themselves be integrated with GNSS [Schleicher et al., 2009, Carlson, 2010], or inertial measurement units (IMUs) [Qin et al., 2017, von Stumberg et al., 2018] to improve the precision. However, to disambiguate the contribution of our method from additional sensor information, we only consider visual information, i.e. camera images.

The overall framework works as follows. Given an image, we train a model which predicts the lateral movement of the relative translation vector between 2 camera frames. The relative translation vector predicted by our model at inference time can be considered as a *virtual* target indicating to the vehicle where to traverse next. Therefore, our framework can be seen as a waypoint-based control approach. However, the main difference to other approaches [Kaufmann et al., 2018, Agarwal et al., 2019] is that we do not require physical waypoints nor accurate GNSS reception at any given time. The authors of [Bansal et al., 2019] also use *virtual* waypoints to navigate a robot indoors, however, they assume that the odometry is perfectly known. Figure 1 shows the high-level framework of the proposed self-supervised vehicle control method. Our emphasis is on retrieving the relative translation vector which would be the same irrespective of which car the module is placed on. Hence, our approach of placing the camera module at the front of any vehicle can be scaled up to as many numbers of running cars on the road as possible without drive-by-wire limitations or regulatory restrictions. To this end, we make the following contributions:

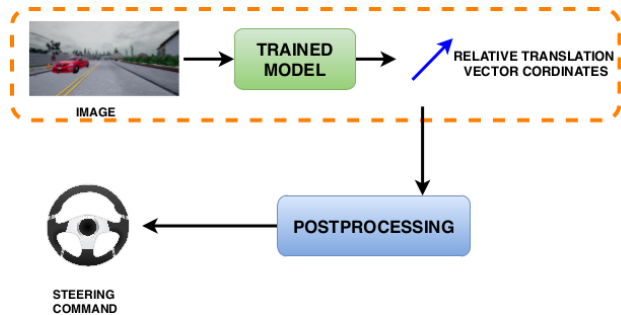


Figure 1: This figure shows the high-level overview for vehicle control. An image is fed to a deep neural network which is trained in a self-supervised manner using camera poses obtained by visual odometry. The model predicts the relative lateral translation vector for the next camera position, hence indicating the direction of where the car ought to move next. This vector value is then passed on to the post-processing module to produce an appropriate steering command.

1. Ability to control a car in a self-supervised fashion without the need for an expert driver. This is achieved by using camera poses obtained through visual odometry methods.
2. We show how a combination of trajectory information from multiple runs leads to improved performance.
3. Using our proposed scalable solution, we demonstrate how our model can generalize even to variations in translation and rotation of the camera at inference time.

Note that under dynamic weather and lighting conditions, VO methods may fail to furnish the correct camera poses. Therefore, we additionally show how our approach can be extended to such circumstances where VO may fail.

Although there are many publicly available data sets [Geiger et al., 2012, Cordts et al., 2016, Maddern et al., 2017, Sun et al., 2020, Huang et al., 2018, Wenzel et al., 2020] focused on driving applications, they do not offer the possibility to test the online performance of self-driving cars necessary to evaluate our approach. As discussed in Section 3, this is because the online performance of embodied agents does not necessarily correlate with the offline error metrics [Anderson et al., 2018, Codevilla et al., 2018a, Ross et al., 2011]. To circumvent this problem, we report results tested on the CARLA simulator [Dosovitskiy et al., 2017]. It offers the facility to conduct an online evaluation of our method.

2 Method

In this section, we first describe the self-supervised control framework we use. Our approach focuses on low-level lateral control for vehicles. The problem we want to tackle is the following: given an image from a monocular camera at any instance in time, we would like to determine the immediate steering command where the car should maneuver next. This can be done by predicting a translation vector indicating the direction from where the car currently is to where it ought to be. But how do we get this translation vector without explicit supervision?

2.1 Camera Pose Estimation

The only requirement for our approach is to have camera poses estimated by visual odometry (VO). This can be done by placing a camera setup at the front of a car recording the images as the driver controls the car. A VO method run on a sequence of these images yields the corresponding camera poses.

The 6DoF pose can be estimated using a state-of-the-art VO method such as DSO [Engel et al., 2017] or ORB-SLAM [Mur-Artal and Tardós, 2017]. The camera poses are represented as transformation matrices $\mathbf{T}_i \in \mathbf{SE}(3)$, which are transforming a point from the world frame into the camera frame. To get the relative pose between two cameras at time i and $i+1$, we can extract out the rotation matrix \mathbf{R} and the translation vector \mathbf{t} as follows:

$$\begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} = \mathbf{T}_i \mathbf{T}_{i+1}^{-1}, \quad (1)$$

where the rotation matrix \mathbf{R} performs a rotation in Euclidean space which can be represented by yaw, pitch, and roll angles. The translation vector \mathbf{t} denotes a translation around the x, y, z axes. Since the camera is rigidly attached to the car, the relative pose between any 2 camera frames also gives the relative information between the car positions at these 2 instances. We further assume that the car has a planar motion with its state determined by x, y , and orientation angle in a global frame of reference. Throughout this paper, we will assume that we do not observe motion along the z direction of the car, as well as a constant roll and pitch.

The translation vector is defined as the relative vector between 2 camera poses. Note that in the global frame of reference, trajectories corresponding to the same steering commands may undesirably have different translation vectors. Therefore, to circumvent this issue, we redefine the frame of reference to a local coordinate system such that the forward motion of the car at any instance in time is always in the x' direction, whereas any lateral motion is always in the y' direction. The coordinates of the relative

position vector in the local frame of reference are labeled as dx and dy . Details of which can be found in the supplementary. Note that if the rate at which frames are received or the speed at which the ego-vehicle is moving is variable, then calculating the translation vector between 2 consecutive frames will yield inconsistent results. To overcome this, we select only those pair of frames that are a fixed distance dx apart up to a certain tolerance. Hence, dx can be considered to be a constant.

2.2 Bicycle Model for a Car

We model the dynamics of a 4 wheel front-wheel drive car using the bicycle model [Wang and Q, 2001] depicted in Figure 2. The model assumes planar motion with the spatial coordinates and orientation describing the state of the car. The 2 rear wheels of the car are represented by a single rear wheel (point A). Likewise, the 2 front wheels are represented by a single front wheel (point B). The lateral motion of the bicycle can be controlled by maneuvering the steering angle (δ) of the front wheel, whereas the longitudinal/forward motion of the vehicle is controlled by applying the throttle.

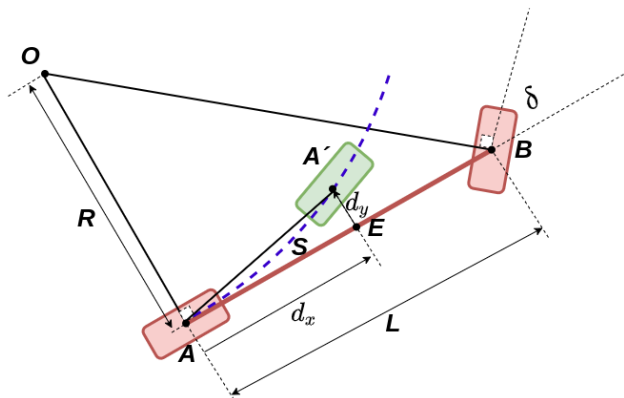


Figure 2: Depicts the bicycle model for a forward drive 4-wheeled car. L is the distance between the rear (point A) and front (point B) wheels. With a steering angle of δ , the car would traverse a curvature whose radius is given by R and the rear wheel would move from point A to point A' by covering a distance S . The longitudinal and lateral motion of the car in the car's local frame of reference is given by dx and dy respectively. O is the instantaneous center of rotation at the intersection of line segments drawn perpendicular to the front and rear wheels of the car. The dotted line represents the curvature of the trajectory.

During turns, the front and rear wheels will be oriented in different directions. Attempting to make sharp turns at high speeds would exert excessive lateral forces, causing the car to slip. An important assumption we make is that the vehicle does not experience any slip. This is true when both wheels are oriented in the same direction irrespective of the

speed. Whereas, while executing turns, this holds at moderately low speeds (5 m/s or ≈ 20 km/h) [Rajamani, 2012]. We can geometrically express the steering angle, δ in terms of the lateral translation vector determined by the method described in Section 2.1 as follows:

$$\angle OBA + 90 + \delta = 180 \quad (2)$$

$$\angle OBA = 90 - \delta \quad (3)$$

$$\angle AOB + 90 + \angle OBA = 180 \quad (4)$$

$$\angle AOB = \delta \quad (5)$$

$$\tan(\angle AOB) = \tan(\delta) = \frac{L}{R} \quad (6)$$

$$R = \frac{L}{\tan(\delta)} \quad (7)$$

If the frame rate at which the information of successive frames is high coupled with the assumption of the vehicle moving at moderately low speeds while executing the turn would yield the following implications:

1. The change in the orientation of the vehicle would be small giving $\Delta\phi \approx \angle A'AE \approx 0$.

$$\tan(\angle A'AE) \approx \tan(\Delta\phi) \approx \Delta\phi \approx \frac{dy}{dx}. \quad (8)$$

2. The distance S , traversed by the vehicle along the curvature is $\approx AA'$. Moreover, $\Delta\phi \approx 0$ implies $AA' \approx dx$.
3. The angular velocity of the car is given by $\frac{d\phi}{dt} \approx \frac{1}{R} \cdot \frac{dS}{dt}$. A high frame rate implies:

$$\Delta\phi \approx \frac{S}{R} \approx \frac{AA'}{R} \approx \frac{dx}{R} \quad (9)$$

Substituting Equation (7) in (9) results in

$$\Delta\phi = \frac{dx}{L} \tan(\delta) \quad (10)$$

Equating Equations (8) and (10) produces:

$$\frac{dy}{dx} = \frac{dx}{L} \tan(\delta) \quad (11)$$

$$\delta = \tan^{-1}\left(\frac{dy \cdot L}{dx \cdot dx}\right) \quad (12)$$

As described in Section 2.1, the image pairs are selected such that they are approximately a distance of dx apart. Hence, denominator of Equation (12) can be taken as constant. Since L is the length between the front and rear

wheels, it is fixed, hence the above equation can be rewritten as:

$$\delta = \tan^{-1}(dy \cdot \alpha) \quad (13)$$

Where α is a constant that can be tuned at run-time depending on the car. Therefore, we can determine the steering angle of the vehicle from the lateral component of the relative translation vector which in turn is determined from the method described in Section 2.1. For this, we train a neural network model which takes in an image as input and outputs the corresponding value of dy by minimizing the $L1$ loss between the value predicted by the model and value of dy . In the supplementary material, we demonstrate the implications on the performance of the trained models upon deviating from the assumptions given in this section.

2.3 Multi-Car Trajectories

One issue with training a self-supervised vehicle control model with only a single car trajectory is that there may not be many variations in the collected data. Hence, at inference time, the car may slightly deviate from the reference trajectory (being the only trajectory on which it was trained) and come across a scene of possibly a different trajectory which it had never seen in the training set. Hence, the model may not be able to figure out the correct action to take against the next sequence of images. This is depicted in Figure 3.

To solve for this, note that drivers have a unique driving preference. Some tend to drive in the center, while others close to the lane or sidewalks. In fact, the same driver might demonstrate different trajectories while driving along the same route multiple times. Note that these behaviors lead to a diverse set of trajectories along the same section of a road. Hence, a slight deviation for one driver may be the usual trajectory for a different driver. But how can we scale our training to also incorporate this vast amount of information being generated by multiple drivers/runs across the same road section? This is also depicted in Figure 3 by the orange curve showing the trajectory generated by a different run along the same route. For any image along this new trajectory, the model aims to predict the relative position vector to the next closest image pose of the original purple trajectory. The frames in this new trajectory can be relocalized against the reference using direct visual localization approaches such as GN-Net [von Stumberg et al., 2020] or by running the PnP algorithm in a RANSAC scheme using robust visual image descriptors, e.g. R2D2 [Revaud et al., 2019].

Reference Trajectory: For any section of the road having multiple trajectories, we define the reference trajectory as the one to which the relative vector from any of the many trajectories is mapped to. In Figure 3, the purple line is taken as the reference trajectory. For the case of multiple

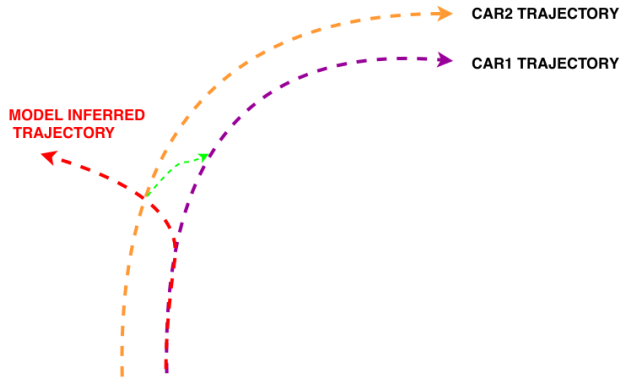


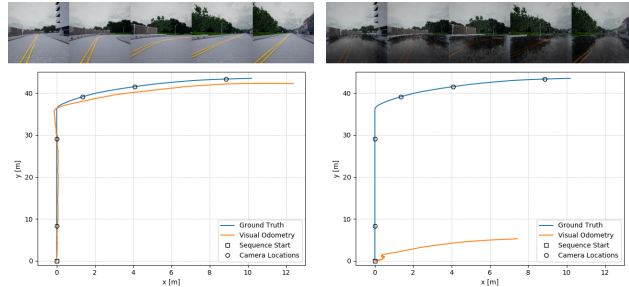
Figure 3: The purple line shows a section of the trajectory for *Car1* used to train the self-supervised model. At test time, the model might slightly deviate away from this learned trajectory and come across images it has never seen in the training set. Hence, it would not be able to take the correct action to return to the original path after this slight deviation. Eventually, this deviation would accumulate with the car possibly crashing into an obstacle. The orange line shows the trajectory of another different car (*Car2*). During training, images from this trajectory can also be used to map the relative vector from *Car2* trajectory to the next frame of *Car1*. Hence, at test time, even if the car deviates from the original *Car1* trajectory, it will still be able to make the correct maneuver to return to its original course.

trajectories, then for each, we calculate the lateral motion vector to every other trajectory and sum it up. The centermost will have the lowest sum and is taken as the reference.

2.4 Control Under Dynamic Weather/Lighting Conditions

In dynamic weather conditions, VO methods may fail to generate an accurate representation of the trajectory. Hence, it may not be possible to obtain self-supervised training data for controlling a car in such a scenario. This is depicted in Figure 4, where under the *CloudyNoon* condition the trajectory generated by VO matches closely with the ground truth trajectory. This is in contrast to the *HardRainNoon* condition, where visual odometry fails for the same sequence.

To overcome this issue, we adopt the teacher-student approach proposed by [Wenzel et al., 2018]. Note that our model can be divided into a feature extraction (FEM) and Steering Angle Prediction (SAP) module. The FEM extracts out features with an RGB image as input. These features are in turn fed to the SAP which predicts the coordinate for the lateral motion of the car. This model is trained with the visual odometry trajectories obtained under the *CloudyNoon* condition and is referred to as the teacher. This model will fail if tested on *HardRainNoon* condition.



(a) *CloudyNoon* Trajectory. (b) *HardRainNoon* Trajectory.

Figure 4: This figure shows the influence of inclement weather on a small sample stretch of a trajectory produced by VO on the CARLA simulator. The curve in blue is the ground truth trajectory, and the curve in orange is the one obtained with VO. VO fails on the *HardRainNoon* condition. Top: Camera snapshots from the respective locations on the sequence indicated by circles in the ground truth trajectory.

For this, we train another student model whose SAP module has exactly the same weights as that of the teacher model. Meanwhile, its FEM module is trained by feeding an image translated from *CloudyNoon* to *HardRainNoon* using an image-to-image translation (GAN) network [Anoosheh et al., 2018] trained in an unsupervised manner. The generator of the GAN takes in an image from domain A , I_A (*CloudyNoon*) and translates it to an image I_B belonging to domain B (*HardRainNoon*). The weights are updated by minimizing the mean square error between the features of the 2 models. The architecture is depicted in Figure 5.

3 Experiments

Coming up with good metrics for evaluating the task of visual navigation is rather challenging. In particular, for autonomous driving, it is not obvious how to measure the performance of the driving policy. The authors of [Anderson et al., 2018] present different problem scenarios for the analysis of embodied agent navigation and suggest some evaluation metrics for standard scenarios. Moreover, [Codevilla et al., 2018a] conduct further studies on online and offline metrics, especially for vision-based sensorimotor control. Specifically, in this work, the difference between metrics on static images compared to evaluating online in the simulation is analyzed. Their result showed that driving models can have similar offline metrics but drastically different driving performance. They concluded from their experiments that offline metrics are weakly correlated with actual driving quality. Therefore, it is not trivial to link from offline to online performance due to a low correlation between them.

Due to the aforementioned reasons, we interpret our con-

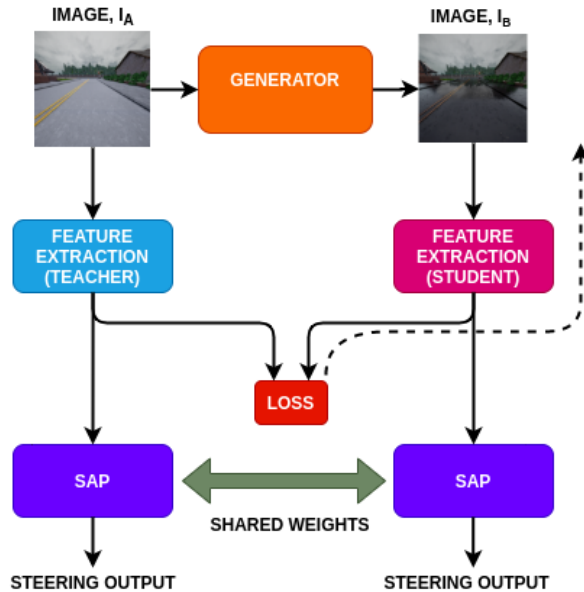


Figure 5: This figure shows the teacher-student architecture used to produce the steering commands independently of the image input domain. The generator is pre-trained using samples from both domains. The solid arrows represent forward propagation into the network, whereas the dotted line depicts back-propagation to update the weights of the student model.

clusions in an online setting as this is directly correlated to the actual driving performance. For this, we use the stable version v0.8.2 of CARLA [Dosovitskiy et al., 2017] simulator for training and evaluation of our models. CARLA provides 15 different weather conditions. We train our models only on the *CloudyNoon* condition but test our results on all other conditions. We compare the performance of our method with that of [Khan et al., 2019] which trains a supervised model that is robust across all the weather conditions. The results were reported for executing turns around corners at fixed speeds. The metric to measure the performance of the models was done by calculating the ratio of time the car remains within its own driving lane. We adapt the same for comparison in our experiments.

3.1 Models

For the purpose of evaluation, we evaluate 4 different models whose descriptions are given below. While in principle, our approach can be deployed in conjunction with any VO based method, we have used [Wang et al., 2017] for our experiments:

1. *Supervised model with ground truth steering angle:* This is the method of [Khan et al., 2019] trained in a supervised manner using the ground truth steering angles directly obtained from the CARLA simulator.

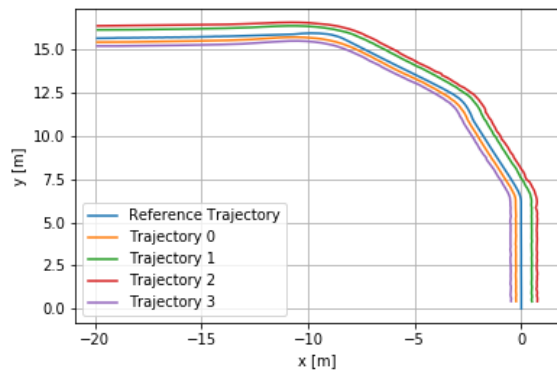


Figure 6: This figure shows the reference trajectory obtained by running the VO algorithm. 4 additional trajectories in relation to the reference are also depicted.

The network is capable of handling images from all 15 weather conditions. We may use this model trained with ground truth supervised steering angle labels as the Oracle to compare against the other models not trained with any ground truth data.

2. *Visual Odometry, with Multiple Trajectories:* Here we use the estimated camera poses obtained through VO to implicitly map to the steering angle of the car. Figure 6 shows the reference VO trajectory along with 4 other trajectories in relation to the reference.
3. *Visual Odometry, with Single Trajectory:* Similar to the previous model. The images used for training this model are only taken from the reference trajectory.
4. *Visual Odometry, with Multiple Trajectories, Robust:* Same as the model described previously using VO on multiple trajectories except this model is also made robust to all the remaining weather conditions in a completely unsupervised manner by the method described in Section 2.4. However, to cater for multiple weathers not only the generator for conversion between *CloudyNoon* to *HardRainNoon* is trained but also the generators for *HardRainSunset*, *WetCloudyNoon*, *CloudySunset*, *MidRainSunset*, *ClearSunset*, *ClearNoon*, *WetSunset*, *WetCloudySunset*, *WetNoon*. Images produced by all these generators are used for training the student model.

Figure 7 is a plot showing the performance of the aforementioned 4 models against all the 15 weather conditions.

4 Discussion

In the following, we will elaborate on important key points regarding the performance of the 4 models and conduct additional studies to further investigate the findings.

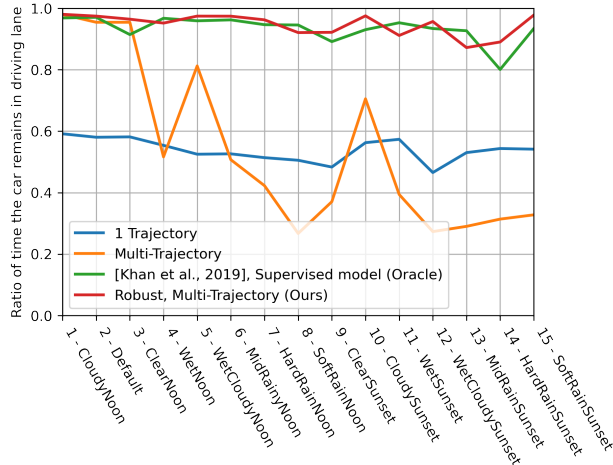


Figure 7: The plot exhibits the ratio of time the car remains within its own driving track for the 4 models across all the 15 different weather conditions. Higher is better.

1. **Single Trajectory Performance:** All models demonstrate good performance on the weather condition on which they were trained with the model trained on 1 trajectory being an exception. This is because, during the online evaluation, slight deviations from the reference trajectory might lead to a collapse of the model since it has no way of coming back to its original course. This is due to the fact that the model had never seen such scenes from different trajectories during training. Another reason that would explain this poor performance is that if the poses estimated by visual odometry used to train this model are by themselves not perfect, then the model is being trained to the imperfections of this imprecise dataset. Although these predictions may be marginally incorrect locally, however, in an online setting they tend to accumulate over the long run and thereby causing a collision with obstacles.
2. **Multi Trajectory Performance:** In contrast, the 2 models trained with VO on multiple trajectories exhibit superior online performance alluding to the utility of combining information from multiple trajectories. In fact, the performance of these 2 models is comparable with the Oracle that was trained in a supervised manner using ground truth labels.
3. **Method Scalability:** A plausible explanation for this is depicted in Figure 8 which demonstrates the scaling capability of our method. The blue curve shows that increasing the number of training trajectories improves the online performance. In fact, after a certain number of trajectories, the performance tends to saturate since any addition is already captured by the previous. Moreover, the figure also shows that our method is robust to changes in camera parameters at

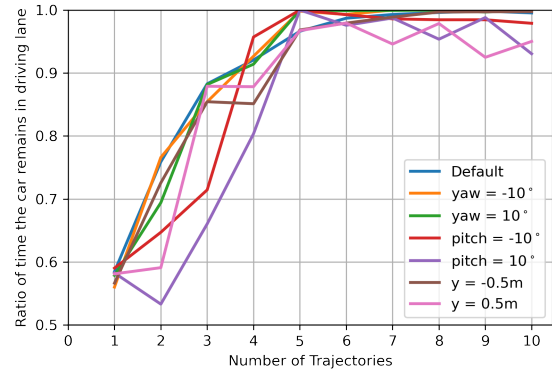


Figure 8: This figure shows the ratio of the time the car remains within its own driving track for different number of trajectories and changes in camera parameters from the default. The yaw and pitch values report the degree change in these parameters from the default. Whereas the relative movement of the camera along the width of the car (y) with respect to the default position is given in meters. Higher is better.

inference time. Hence the position/orientation of the camera can be adjusted to yield the best performance at inference.

4. **Performance under Dynamic Weather Conditions:** Note that both the models trained on multiple trajectories give good performance on the weathers that closely resemble the *CloudyNoon* condition. However, on all other conditions, only our model trained for domain adaptation in a completely unsupervised manner by the method described in Section 2.4 maintains equivalent performance. It is interesting to note that our approach even works for the *HardRainNoon* condition, where, VO had failed to produce the correct trajectory of the car as depicted in Figure 4.
5. **Robustness to Perturbations:** Figure 9 shows the implication of inducing varying degrees of perturbations to the final steering command. As the perturbations are progressively increased, the performance for the model trained with only one trajectory, tends to correspondingly deteriorate. The models trained with multiple trajectories are less influenced by these perturbations. This can be observed in Figure 10 wherein, not only the mean accuracy tends to improve but the deviation in performance across the perturbations also shows a diminishing trend with the increasing number of trajectories.
6. **Influence of Scalability on Perturbations:** It is interesting to note from Figure 8 that models which were trained with 5 or more trajectories maintain a performance ratio of remaining in the driving lane of greater

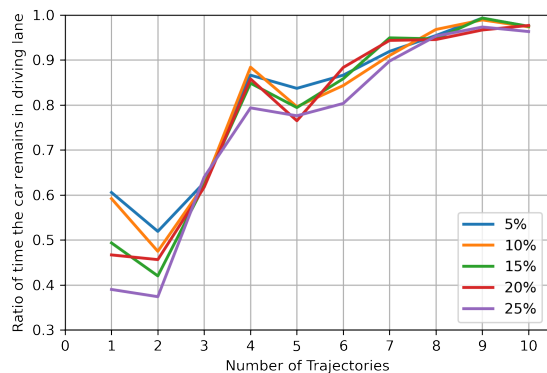


Figure 9: This figure shows the influence on the performance of the models trained with different number of trajectories as varying levels of perturbations (shown as a percentage) are introduced into the final steering command. For a specific percentage of perturbation, each line in the figure represents the mean ratio of the time the car remains within its own driving track across all the camera positions described in Figure 8.

than 0.9. However, the model trained with 5 trajectories is inferior in performance when perturbations are introduced as depicted in Figure 10. Here, a minimum of 8 trajectories was needed to achieve similar performance. This alludes that scaling our method across more trajectories tends to lead to better robustness to perturbations.

Videos depicting the online performance of the models along with additional details of the experiments can be found in the supplementary material.

5 Conclusion

In this work, we address a central challenge in the development of self-driving cars – namely the question of how a machine can learn suitable driving strategies merely by observing the driver’s behavior. We presented a novel approach for self-supervised vehicle control that combines the strength of any state-of-the-art visual odometry method with learning-based techniques. We achieve this in a framework that uses camera poses traced out along a vehicle’s trajectory. Hence training can be done in an entirely self-supervised fashion. Moreover, our method of using multiple trajectories allowed for the scaling of the model. The proposed framework facilitates replacing the time-consuming and expensive process of collecting expert-driving data. This approach could be particularly useful for deploying on public transportation vehicles that have different drivers traversing the same route multiple times.

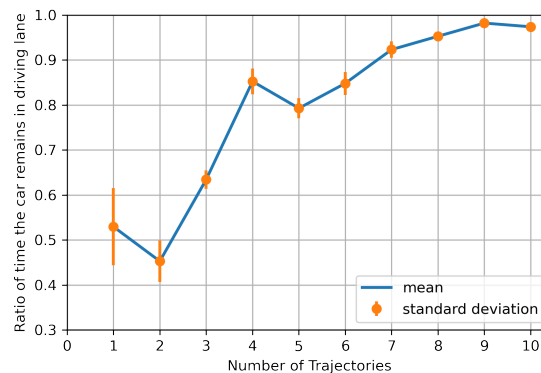


Figure 10: This figure shows the mean and standard deviation for the curves shown in Figure 9 across the different levels of perturbations.

References

- [Abbeel et al., 2007] Abbeel, P., Coates, A., Quigley, M., and Ng, A. Y. (2007). An application of reinforcement learning to aerobatic helicopter flight. In *Advances in neural information processing systems*, pages 1–8.
- [Abbeel and Ng, 2004] Abbeel, P. and Ng, A. Y. (2004). Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the twenty-first international conference on Machine learning*, page 1. ACM.
- [Agarwal et al., 2019] Agarwal, T., Arora, H., Parhar, T., Deshpande, S., and Schneider, J. G. (2019). Learning to drive using waypoints. In *Machine Learning for Autonomous Driving Workshop at the 33rd Conference on Neural Information Processing Systems, Vancouver, Canada*.
- [Anderson et al., 2018] Anderson, P., Chang, A., Chaplot, D. S., Dosovitskiy, A., Gupta, S., Koltun, V., Kosecka, J., Malik, J., Mottaghi, R., Savva, M., et al. (2018). On evaluation of embodied navigation agents. *arXiv preprint arXiv:1807.06757*.
- [Anoosheh et al., 2018] Anoosheh, A., Sattler, T., Timofte, R., Pollefeys, M., and Van Gool, L. (2018). Night-to-day image translation for retrieval-based localization. *arXiv preprint arXiv:1809.09767*.
- [Bansal et al., 2019] Bansal, S., Tolani, V., Gupta, S., Malik, J., and Tomlin, C. (2019). Combining optimal control and learning for visual navigation in novel environments. In *Conference on Robot Learning (CoRL)*.
- [Bojarski et al., 2016] Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., Jackel, L. D., Monfort, M., Muller, U., Zhang, J., et al. (2016). End to end learning for self-driving cars. *arXiv preprint arXiv:1604.07316*.

- [Carlson, 2010] Carlson, J. (2010). *Mapping Large, Urban Environments with GPS-Aided SLAM*. PhD thesis, Carnegie Mellon University.
- [Chen et al., 2015] Chen, C., Seff, A., Kornhauser, A., and Xiao, J. (2015). Deepdriving: Learning affordance for direct perception in autonomous driving. In *IEEE International Conference on Computer Vision (ICCV)*, pages 2722–2730.
- [Codevilla et al., 2018a] Codevilla, F., López, A. M., Koltun, V., and Dosovitskiy, A. (2018a). On offline evaluation of vision-based driving models. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 236–251.
- [Codevilla et al., 2018b] Codevilla, F., Müller, M., López, A., Koltun, V., and Dosovitskiy, A. (2018b). End-to-end driving via conditional imitation learning. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1–9. IEEE.
- [Cordts et al., 2016] Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., and Schiele, B. (2016). The cityscapes dataset for semantic urban scene understanding. In *Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [Dosovitskiy et al., 2017] Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., and Koltun, V. (2017). CARLA: An open urban driving simulator. In *Conference on Robot Learning (CoRL)*.
- [Engel et al., 2017] Engel, J., Koltun, V., and Cremers, D. (2017). Direct sparse odometry. *IEEE transactions on pattern analysis and machine intelligence*, 40(3):611–625.
- [Geiger et al., 2012] Geiger, A., Lenz, P., and Urtasun, R. (2012). Are we ready for autonomous driving? the kitti vision benchmark suite. In *Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [Huang et al., 2018] Huang, X., Cheng, X., Geng, Q., Cao, B., Zhou, D., Wang, P., Lin, Y., and Yang, R. (2018). The apolloscape dataset for autonomous driving. *CoRR*, abs/1803.06184.
- [Kaufmann et al., 2018] Kaufmann, E., Loquercio, A., Ranftl, R., Dosovitskiy, A., Koltun, V., and Scaramuzza, D. (2018). Deep drone racing: Learning agile flight in dynamic environments. In *CoRL*.
- [Kendall et al., 2019] Kendall, A., Hawke, J., Janz, D., Mazur, P., Reda, D., Allen, J.-M., Lam, V.-D., Bewley, A., and Shah, A. (2019). Learning to drive in a day. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 8248–8254. IEEE.
- [Khan et al., 2019] Khan, Q., Wenzel, P., Cremers, D., and Leal-Taixé, L. (2019). Towards generalizing sensorimotor control across weather conditions. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.
- [Lillicrap et al., 2015] Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., and Wierstra, D. (2015). Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971*.
- [Maddern et al., 2017] Maddern, W., Pascoe, G., Linegar, C., and Newman, P. (2017). 1 Year, 1000km: The Oxford RobotCar Dataset. *The International Journal of Robotics Research (IJRR)*, 36(1):3–15.
- [Muller et al., 2006] Muller, U., Ben, J., Cosatto, E., Flepp, B., and Cun, Y. L. (2006). Off-road obstacle avoidance through end-to-end learning. In *Advances in neural information processing systems*, pages 739–746.
- [Mur-Artal et al., 2015] Mur-Artal, R., Montiel, J. M. M., and Tardos, J. D. (2015). Orb-slam: a versatile and accurate monocular slam system. *IEEE transactions on robotics*, 31(5):1147–1163.
- [Mur-Artal and Tardós, 2017] Mur-Artal, R. and Tardós, J. D. (2017). Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE T-RO*, 33(5).
- [Qin et al., 2017] Qin, T., Li, P., and Shen, S. (2017). Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, PP.
- [Rajamani, 2012] Rajamani, R. (2012). Vehicle dynamics and control. In *Second Edition, Publisher: Springer*.
- [Revaud et al., 2019] Revaud, J., Weinzaepfel, P., de Souza, C. R., and Humenberger, M. (2019). R2D2: repeatable and reliable detector and descriptor. In *NeurIPS*.
- [Ross et al., 2011] Ross, S., Gordon, G., and Bagnell, D. (2011). A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 627–635.
- [Schleicher et al., 2009] Schleicher, D., Bergasa, L., Ocaña, M., Barea, R., and Guillén, M. (2009). Real-time hierarchical gps aided visual slam on urban environments. In *International Conference on Robotics and Automation (ICRA)*, volume 5717.
- [Sun et al., 2020] Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B., et al. (2020). Scalability in perception for autonomous driving: Waymo open dataset. In *Conference on Computer Vision and Pattern Recognition (CVPR)*.

- [von Stumberg et al., 2018] von Stumberg, L., Usenko, V., and Cremers, D. (2018). Direct sparse visual-inertial odometry using dynamic marginalization. In *International Conference on Robotics and Automation (ICRA)*.
- [von Stumberg et al., 2020] von Stumberg, L., Wenzel, P., Khan, Q., and Cremers, D. (2020). GN-Net: The gaussian loss for multi-weather relocalization. *IEEE Robotics and Automation Letters (RA-L)*, 5(2):890–897.
- [Wang and Q, 2001] Wang, D. and Q, F. (2001). Trajectory planning for a four-wheel-steering vehicle. In *IEEE International Conference on Robotics and Automation (ICRA)*.
- [Wang et al., 2017] Wang, R., Schwörer, M., and Cremers, D. (2017). Stereo DSO: Large-scale direct sparse visual odometry with stereo cameras. In *International Conference on Computer Vision (ICCV)*.
- [Weiss et al., 2011] Weiss, S., Scaramuzza, D., and Siegwart, R. (2011). Monocular-SLAM-based navigation for autonomous micro helicopters in GPS-denied environments. *Journal of Field Robotics*, 28(6):854–874.
- [Wenzel et al., 2018] Wenzel, P., Khan, Q., Cremers, D., and Leal-Taixé, L. (2018). Modular vehicle control for transferring semantic information between weather conditions using GANs. In *Conference on Robot Learning (CoRL)*.
- [Wenzel et al., 2020] Wenzel, P., Wang, R., Yang, N., Cheng, Q., Khan, Q., von Stumberg, L., Zeller, N., and Cremers, D. (2020). 4Seasons: A cross-season dataset for multi-weather SLAM in autonomous driving. In *German Conference on Pattern Recognition (GCPR)*.