From Perception to Decision: A Data-driven Approach to End-to-end Motion Planning for Autonomous Ground Robots

Mark Pfeiffer, Michael Schaeuble, Juan Nieto, Roland Siegwart and Cesar Cadena

***Abstract*— Learning from demonstration for motion planning is an ongoing research topic. In this paper we present a model that is able to learn the complex mapping from raw 2D-laser range findings and a target position to the required steering commands for the robot. To our best knowledge, this work presents the first approach that learns a target-oriented end-to-end navigation model for a robotic platform. The supervised model training is based on expert demonstrations generated in simulation with an existing motion planner. We demonstrate that the learned navigation model is directly transferable to previously unseen virtual and, more interestingly, real-world environments. It can safely navigate the robot through obstacle-cluttered environments to reach the provided targets. We present an extensive qualitative and quantitative evaluation of the neural network-based motion planner, and compare it to a grid-based global approach, both in simulation and in real-world experiments.**

I. Introduction

One of the major challenges in robotics is to make robots perform as desired by human operators. Regarding ground robot navigation, this problem is defined as getting from the current position to a target position, fulfilling the desired navigation policy. Although objectives like, e.g. short path or a safe distance to obstacles are perfectly clear to the human operator, it typically requires time-consuming hand tuning such that the robot moves as desired and required. Additionally, classical motion planning solutions require several steps of data preprocessing that typically are decoupled [1]. A map of the environment has to be provided, the sensor data has to be preprocessed and potential objects have to be detected such that the planner can react accordingly in a later stage.

With the aim of reducing the amount of hand-tuning parameters of several processes in order to achieve the desired navigation performance, in this work we present an approach that goes vice-versa: *A data-driven end-to-end motion planner*. The robot is provided with expert demonstrations of how to navigate in a given virtual training environment. Like this, a robot operator can show the desired behavior and navigation strategy to the robot. During navigation, the goal is not only to replicate the provided expert demonstrations in one specific scenario as in teach-and-repeat approaches [2], but rather to be able to learn collision avoidance strategies and transfer them to previously unseen real-world environments.

This work has partially received funding from the European Union Seventh Framework Programme FP7/2007-2013, Challenge 2, Cognitive Systems, Interaction, Robotics, under grant agreement No. 610603, EUROPA2.

The authors are with the ETH Zurich, Zurich, Switzerland. {pfmark, schamich, nietoj, rsiegwart, cesarc}@ethz.ch.

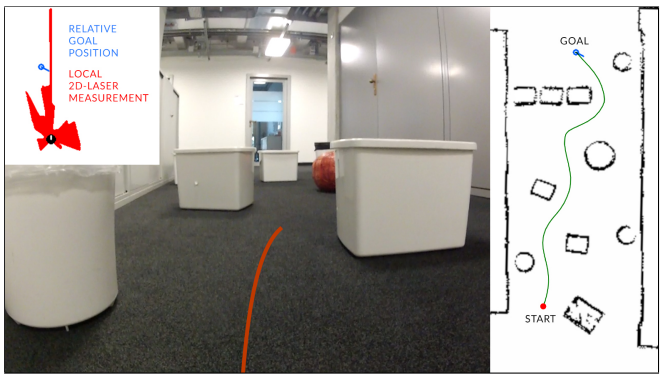


Fig. 1: The robotic platform is able to safely navigate through a maze-like environment only using local laser and target information (upper left). The final traversed trajectory is visualized on the map (right). The robot view image is only shown for visualization purposes but not used as an input to the end-to-end algorithm.

In comparison to multi-layer map based approaches as presented in [1], our approach does not require a global map to navigate. Given the sensor data and a relative target position, the robot is able to navigate to the desired target while avoiding collisions with surrounding obstacles. The approach is constructed under the hypothesis that by learning a physical understanding of the environment and the navigation characteristics of an expert operator, our machine learning-based approach is able to perform in a similar way, even in previously unseen scenarios. By design, the approach is not limited to any kind of environment. However in this work, our analysis is focused on the navigation in static environments.

Computing the motion commands directly from the laser data can be an arbitrarily complex task which requires a model that is able to capture the relevant characteristics for local motion planning. Among various machine learning approaches, deep neural network (DNN)-based ones have the largest potential to model complex dependencies. They have shown their potential for complex physical scene understanding and feature extraction in various applications like [3], [4], [5], to name a few. To train an end-to-end motion planner, we make use of an existing global path planning approach that provides complete trajectories from start to goal positions. Since for spatial scene understanding and collision avoidance the distances to surrounding objects are of special interest, we use a front facing 270° laser range finder as the only sensor of the robot. The model is trained on simulation data generated with a global path planner as the expert operator.

The performance of the learned end-to-end motion model is tested both in simulation and on a real robotic platform. In order to analyze the generalization error of the

model for local motion planning, it is especially important to conduct experiments in previously unseen environments. Therefore, we deployed the robot for evaluations in unknown environments which additionally were significantly more challenging than the one used for training.

The main contributions of this work are:

- A data-driven end-to-end motion planner from laser range findings to motion commands.
- Deployment and tests on a real robotic platform in unknown environments.
- Extensive evaluation and comparison of the presented local approach with respect to a motion planner with global map information.

The remainder of this document is structured as follows: In Section II we present an overview of the related work. Section III formulates the problem and presents our approach. The experiments and their results are shown in Section IV before we discuss the results in Section V and draw a conclusion in Section VI.

II. Related work

Data driven end-to-end motion planning covers various research areas, both from the perception and the motion planning side. On the perception side, the scene understanding part is especially important. The input data has to be processed to extract relevant information.

Since collisions with surrounding objects need to be avoided and the target has to be reached, in our work especially the physical or more precisely the spatial scene understanding has to be given. Fragkiadaki *et al.* [3] showed that it is possible to learn a model for ball-ball and ball-wall dynamics based on demonstrations shown to the model during training time. Without any prior knowledge about ball kinematics and collisions, this model is able to predict the motion of several billiard balls in previously unseen configurations. It shows that it is possible to model physical/spatial interactions using DNNs. Chen *et al.* [6] presented an approach that extracts spatial information from image data and uses the extracted features for motion planning of an autonomous car. The approach is not end-to-end since it consists of multiple processing layers, yet it already shows that the extracted information can be used directly by a motion planner. Ondurska *et al.* [7] presented an end-to-end application of neural networks for dynamic object tracking using laser data. Their work shows that neural networks can be used to extract important spatial information from two-dimensional (2D) laser data.

On the motion planning side, previous work already showed the performance gain by using learning-based motion models for navigation. Abbeel *et al.* [8] showed the application of apprenticeship learning for learning human's navigation strategies on a parking lot. The application was able to significantly reduce the amount of hand-tuning for motion planning. However, knowledge about the map and road network is required which makes the approach specific for a single environment. The approaches by Kuderer *et al.* [9], Pfeiffer *et al.* [10] and Kretzschmar *et al.* [11] use the maximum entropy inverse reinforcement learning techniques to interaction-aware motion models. The learned interaction models for pedestrians are based on demonstrated behaviour in occupied environments. These applications also show that the amount of hand-tuning of the motion- and interaction models can be reduced by applying machine learning techniques and that the learned motion models can outperform hand-engineered ones.

For applications that require a close coordination between perception and control, like path planning for a robotic arm for closing a bottle, Levine *et al.* [12] showed that end-to-end learning approaches can be beneficial and outperform multi-layered ones. A policy search problem is transformed into a deep reinforcement learning problem that uses a convolutional neural network (CNN) for the complex mapping between states and actions. The learned model can successfully plan motor torques of a robotic arm, given raw image data. Another image-based deep reinforcement learning approach for motion planning/decision making was presented by Mnih *et al.* [13]. They showed that it is possible to learn to play various computer games based on screen pixel data and even to outperform human players.

Regarding mobile robotic applications of end-to-end learning, Ross *et al.* [14] presented an approach that learns a left/right controller for an unmanned aerial vehicle (UAV) based on image data. The UAV was able to autonomously navigate through a forest while successfully avoiding collisions with trees in the majority of the cases. However, only the left/right motion has to be controlled while the forward motion command is still selected by a human operator. Kim *et al.* [15] extended this approach to a hallway application where they learn translational and rotational velocities. Yet the approach was only tested in empty hallways without any objects blocking the UAV's path. Another image-based end-to-end collision avoidance approach is presented by Muller *et al.* [16]. They focus on the image feature extraction and the transferability among different environmental conditions. It is shown that the collision avoidance works, yet the navigation performance of the robot is not analyzed.

Sergeant *et al.* [17] presented a laser-based and data-driven end-to-end motion planning approach based on deep auto-encoders. The collision avoidance capabilities of this approach are shown in simulation and on a robotic platform. Unlike our framework, no robot target position is taken into account and therefore it is not applicable as a local motion planning technique if a target has to be reached. With this approach the robot can drive reasonable paths, however no specific goal can be reached.

III. Approach

This section describes the underlying problem and our approach to solve it.

A. Problem formulation

Humans have outstanding capabilities in perceiving the environment, extracting important information and taking rational decisions based on this information. However, to take

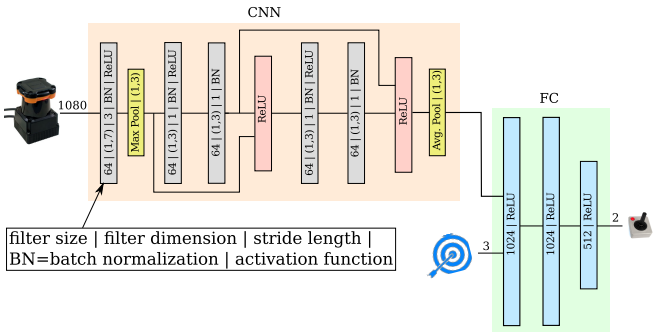


Fig. 2: Structure of the CNN. The laser data is processed by the convolutional part which consists of two residual building blocks as presented in [18]. The FC part of the network fuses the extracted features and the target information. The input/output dimensions of the overall model are shown on the connections. L_1 regularization is applied to all model parameters.

decisions, they can rely on a large knowledge base gained throughout their entire lives.

In order to take the right decisions — in our application this is to move a robot to the desired target position (including heading) — robots have to overcome several challenges. First, the relevant information has to be extracted from the sensor data. Second, using this information, a model has to be found that describes the relationship between the observations and the actions to take. Third, during deployment, this model has to be used to take the right decisions as soon as new observations are available.

Whereas many approaches solve these tasks independently, we present an approach that directly computes suitable steering commands based on sensor and target data. Given expert demonstrations, we try to find a function

$$\mathbf{u} = \mathcal{F}_{\boldsymbol{\theta}}(\mathbf{y}, \mathbf{g}) \tag{1}$$

that directly maps a vector of sensor data $\mathbf{y}$ and goal information $\mathbf{g}$ to desired steering commands $\mathbf{u}$. This function is parametrized by a parameter vector $\boldsymbol{\theta}$. During supervised training we find the function parameters $\boldsymbol{\theta}$ that best explain a set of training data. The optimization criterion is based on $|\mathcal{F}_{\boldsymbol{\theta}}(\mathbf{y}, \mathbf{g}) - \mathbf{u}_{\text{exp}}|$, the difference between the predicted steering commands and the ones provided by the expert operator. During deployment, the model parameters $\boldsymbol{\theta}$ are given and the steering commands can be computed given the input data $\mathbf{y}$ and $\mathbf{g}$.

B. End-to-end model

The end-to-end relationship between input data and steering commands can result in an arbitrarily complex model. Among different machine learning approaches, DNN/CNNs are well known for their capabilities as a hyper-parametric function approximation to model complex and highly non-linear dependencies. In this work, we will use a CNN for the representation of the function $\mathcal{F}_{\boldsymbol{\theta}}$. The entire processing pipeline of extracting information and finding the right steering commands has to be covered by a single model.

As mentioned above, the inputs are given by the measurements of the 2D laser range finder and the relative target position which means the position of the target (polar coordinates) in a robot-centric coordinate system. In order to retrieve spatial scene understanding features, the laser data is processed by a CNN before the outputs of that sub-network are fused with the target information and processed by the FC layers of the model. The structure of the neural network model is shown in Figure 2. The network consists of two residual building blocks including shortcut connections as suggested in [18], where it was shown that the training complexity can be reduced by using residual networks, compared to stacked convolutions.

Throughout this paper, two versions of the model will be investigated. For the first version (*CNN_smallFC*), the three FC layer dimensions are (256, 256, 256) while for the second version (*CNN_bigFC*), their dimensions are increased to (1024, 1024, 512). The convolutional part of both networks remains unchanged, as shown in Figure 2. Our neural network model implementation is based on Google's TensorFlow framework [19].

C. Model training

The ultimate goal for the presented approach is to be able to learn a driving characteristic demonstrated to a robot by an expert operator in a supervised manner. To avoid the burden of human driving data collection at big scale, we resort to simulation where a global motion planner is used as an expert. This is a very valuable feature in robotic applications where data collection is expensive. For each time step i, the data tuple $\gamma_i = (\mathbf{y}_i, \mathbf{g}_i, \mathbf{u}_{\text{exp},i})$ consists of laser measurements, target information and an expert velocity command $\mathbf{u}_{\text{exp},i}$. Here, the velocity command $\mathbf{u} = (v, \omega)$ includes translational and rotational velocity. In order to reduce temporal correlations in the training data to a minimum, the tuples are randomized before used for training. The optimization is conducted using the *Adam* optimizer [20] with mini-batch training. The loss function for each supervised learning step k is given by

$$J_k(\Gamma_B) = \frac{1}{N_B} \cdot \sum_{j=i}^{i+N_B} |\mathcal{F}_{\boldsymbol{\theta}_k}(\mathbf{y}_j, \mathbf{g}_j) - \mathbf{u}_{\text{exp},j}|, \tag{2}$$

where the mini-batch $\Gamma_B = [\gamma_i, \ldots, \gamma_{i+N_B}]$ is comprised of multiple samples of the training data tuples. $\mathcal{F}_{\boldsymbol{\theta}_k}$ represents the model at the current training step. The gradient of this cost function w.r.t. the model parameters can be computed using backpropagation.

D. Motion planner deployment

One advantage of neural network models compared to other approaches is their predictable query time during testing. Whereas the complexity of multi-stage approaches might increase if the environment becomes more complicated, the complexity of a DNN is unaffected by the robot's environment and the query remains unaltered. Since no external preprocessing of the laser data is required, the computational complexity and therefore also the query time for a steering command only depends on the complexity of the model, which is constant once it is trained.

The presented neural network model computes the steering commands frame-by-frame. No internal or external memory is used to take into account previous in- and outputs.

IV. Experiments

This section covers the conducted experiments and their evaluation. First, the robotic platform is introduced. It is used throughout all experiments, both as a model in simulation and for the real-world tests. Second, the training data generation is explained before four experiments (two in simulation, two on the real platform) for evaluation are presented. In the following, the neural network based motion planner will also be referred to as the deep planner.

A. Robotic platform

We use a Kobuki based TurtleBot as a robotic platform. We added a front-facing Hokuyo UTM[1] laser range finder (see video[2]) with a field of view (FOV) of 270° and a maximum scanning range of 30 m to the differential drive robot. The angular resolution of the laser sensor is 0.25° which leads to 1080 measurements. We use an Intel® NUC with an i7-5557U processor with 3.10 GHz running Ubuntu 14.04 as an onboard computer and the Robot Operating System (ROS) [21] as a middleware.

B. Training data

As in [17], we use the ROS 2D navigation stack for the expert two-level motion planner (global and local). On the global level we use a grid-based Dijkstra [1] planner while a dynamic window approach (DWA) [22] planner is used on the local level. Stage 2D is used as a dynamic simulator.

During training data generation, the robot drives to randomly selected target positions on the 10 m × 10 m training map (*train*). The simulated sensor data, relative target positions and expert steering commands are recorded. The selected target positions are guaranteed to be collision free, e.g. lie outside of possible obstacles. The original *train* map is shown in the particular leftmost column of Figure 3 and Figure 4, respectively. For our real-world tests, we used fused training data both from the *train* and the $eval_2$ map. The latter comprises clutter and other objects the robot might face during real-world tests. We generated 6000 trajectories in the *train* map and 4000 trajectories in the $eval_2$ map with 2.1M and 2.2M input/output tuples, respectively. Training the model on a Nvidia GeForce GTX 980 Ti GPU[3] roughly takes 8 h with the conducted 2M training steps.

There were four reasons that caused the decision to only use simulation data (generated with an existing planner) for training: (i) The data generation is deterministic as the ROS planner is deterministic, (ii) for basic experiments we can eliminate noisy sensor data as a potential reason for failure, (iii) it is faster and easier to generate data and (iv) we can test the robustness and the generalization capabilities of the approach if we apply a model trained in simulation on a real robot.

[1]https://www.hokuyo-aut.jp/02sensor/07scanner/utm_30ln.html
[2]https://youtu.be/ZedKmXzwdgI
[3]http://www.geforce.com/hardware/geforce-gtx-980-ti/buy-gpu

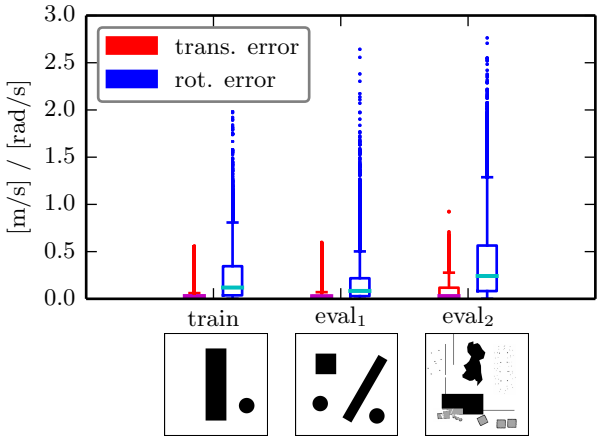


Fig. 3: Error statistics of the frame-by-frame error between the ROS (expert) and the deep planner. The evaluation data was not used for training before. The three small figures visualize the maps on which the evaluation was conducted. Maps are better viewed by zooming in on a computer screen.

C. Frame-by-frame evaluation

The experiment focuses on the evaluation error of the CNN model regarding the computed steering commands. As mentioned in Section IV-B, the model was trained with the samples from the *train* map only. This experiment is conducted using the *CNN_smallFC* model. In order to be independent of a GPU during testing, the query has to be done using a CPU only. On an Intel® i7-4810MQ with 2.8 GHz the average model query time is 4.3 ms.

For the evaluation, we generated input/output tuples for the three maps — comprising *train*, $eval_1$ and $eval_2$ — by driving to 30 random target positions each, using the expert motion planner. Given this data (unseen during training), the error between the translational and rotational steering commands of the deep and the expert motion planner is computed for each input/output tuple.

As the error statistics in Figure 3 suggest, the smallest evaluation error can be observed on the *train* map. We identified that the large outliers of the rotational velocity command typically occur when turning on the spot. For example turning 180° either to the right or left has a large impact on the rotational velocity error but only a minor impact on the actual robot behavior. The $eval_1$ map has a different structure compared to the *train* map, yet the obstacles have a similar shape. This causes an increase of the evaluation error, especially the rotational velocity command. Moving on to the $eval_2$ map, both the translational and the rotational part of the evaluation error increase further.

This result confirms our expectations. The model is able to transfer the scene understanding and navigation knowledge gained in one environment to another one. However, if the deviation in the structure of the environment is large — as e.g. between the *train* and the $eval_2$ environment — the proper scene understanding might not be given which leads to the larger deviation between expert- and deep planner steering commands.

D. Trajectory comparison in a simulated environment

The previous experiment showed that it is possible to transfer learned knowledge from one map to another, yet the

results were only analyzed frame-by-frame. In this section we want to analyze the performance of the navigation model when it is deployed as a motion planner on our robotic platform. This experiment shows, whether the navigation characteristics were learned from the expert or whether there was overfitting on a specific map. The CNN model and training data are the same as in the previous experiment.

Both for the *train* and the $eval_1$ environment, missions with fixed target positions are created. While the ROS planner has global knowledge about the map as during training, the deep planner only receives the relative target (red dots, Figure 4) and the laser range findings at each timestep as an input. Since the simulation is deterministic, only one mission per planner is evaluated. The reproducability of the approach will be tested in the succeeding real-world experiment.

Since the planning structure used for training is a layered motion planning approach, it cannot be described in a single cost function. Therefore, in addition to the visual inspection, the trajectories are also evaluated based on the following metrics inspired by [23]:

- d_{goal}: distance of the final trajecory position to the given goal point summed over all trajectories
- E_{trans}: integrated absolute value of translational acceleration over all trajectories
- E_{rot}: integrated absolute value of rotational acceleration over all trajectories
- $dist$: overall travelled distance for the mission
- $time$: overall travel time for the mission

Figure 4(top) shows an example of the final executed trajectories. The relative errors between both planners for each of the selected metrics are shown in Figure 4(bottom). Those performance metrics are independent of the cost function of the map-based ROS planner but should provide an estimate of the trajectory characteristics of both planners.

The results show that our end-to-end approach is able to capture the navigation policy of the expert operator. Especially on the *train* map, the trajectories driven with both planners are congruent for a majority of the cases, although only local information is used for the deep planner. This is somehow expected, given that the training data was recorded on this map. The deep planner slightly outperforms the expert planner in terms of the given metrics. Since neither the ROS nor the deep planner were trained or tuned for those metrics, this has nothing to do with overfitting. It just rates the performance based on the independent metrics.

Transferring the deep planner to the $eval_1$ environment shows how well the knowledge gained from one map is transferable to the other. Although at many positions there are different topologies on the map, e.g. whether an obstacle should be passed on the left or right side, the deep planner takes similar actions and routes as the expert planner that has perfect knowledge of the full map. Regarding the positioning of the robot at the goal positions and the rotational energy, the deep planner now is inferior to the expert one with respect to the given evaluation metrics. The increase in the rotational energy was already indicated by the rotational

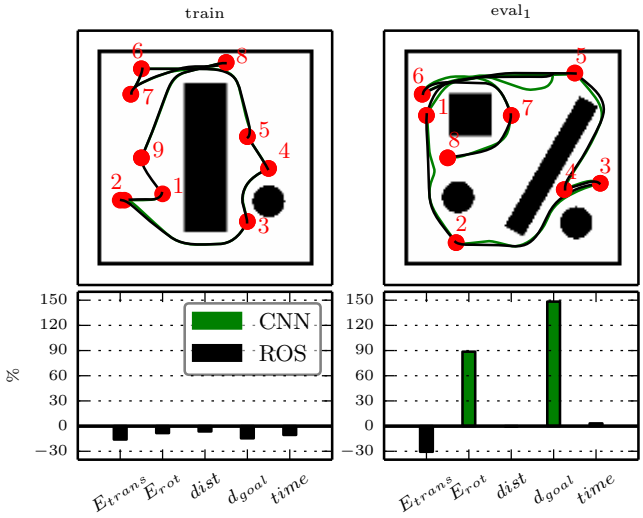


Fig. 4: Performance comparison between the ROS (expert) and the deep planner. Testing results of the trained navigation model on the *train* (left) and $eval_1$ map (right). Top: comparison on a trajectory level with target locations marked in red. Bottom: relative error in [%] of the ROS/deep planner for the evaluation metrics: final distance to goal (d_{goal}), translational energy (E_{trans}), travelled distance ($dist$), rotational energy (E_{rot}) and the travel time ($time$). Green bars indicate a relative error of the deep planner with respect to the ROS expert, black bars the opposite. Therefore, positive (green) error means that the ROS planner is better, negative (black) error means that the deep planner is better.

velocity error increase in Figure 3. Furthermore, this is indicated by the trajectories of the deep planner on the $eval_1$ map in Figure 4(top-right), where it shows deviations from the expert trajectories at several positions. In any of the deviations, the deep planner swerves back to the path of the expert planner, yet the correction causes the extra amount of rotational energy. Inspite the small deviations the travel distance and time are still similar to the expert planner since as Figure 4(top-right) shows, the deep planner has a slight tendency to cut edges sharper than the ROS planner.

The figures in the right column of Figure 4 clearly show that the performance of the deep planner is worse than on the *train* map. Yet they also confirm that the CNN model is able to learn a given navigation characteristic of an expert operator and transfer it to a previously unseen environment and not only to replicate expert demonstrations.

E. Real-world navigation

The following experiment is similar to the one presented in Section IV-D, yet now the driving tests are conducted using the real robot. In order to be able to localize and to navigate with the expert planner, a map of the environment was recorded beforehand. As up to now, the navigation with the deep planner is only based on local laser and target information during test time. During the experiment, the robot has to traverse a maze like area with many obstacles, a table with chairs blocking one half of a corridor, an area with a lot of clutter and also long corridors. Snapshots of the environment are shown together with the results of this experiment in Figure 5 and in the video attachment.

As mentioned above, for the real-world experiments we trained the model based on the *train* and the $eval_2$ environment in order to provide it with the basic navigation

TABLE I: Amount of manual joystick control needed for successful navigation (in [%] of the total distance traveled.)

Target	1	2	3	4	5	6	7	8	9	10	11	12	13
CNN_smallFC	1	0	1	0	4	0	0	0	2	4	6	0	0
CNN_bigFC	0	0	1	1	0	1	0	0	0	0	4	1	0

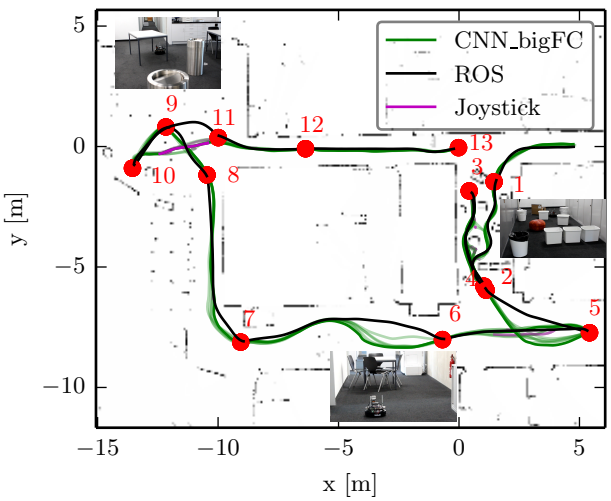


Fig. 5: Comparison between the driven trajectories of the ROS (expert) and the deep planner (*CNN_bigFC*) on a real robot. For the expert planner, one experiment is shown while for the deep planner 6 experiments are shown. The three small pictures give an impression of the actual environment.

principles but also with object shapes that could potentially be observed in reality. We also found that increasing the size of the FC layers of the network improves the navigation performance in the real-world while it was not beneficial during the simulation tests.

The deep planner is able to drive the majority of the missions fully autonomous. However, in some positions, the human operator briefly had to "help" the robot when it got stuck. The joystick interventions are marked in Figure 5. Table I shows that the amount of joytick interventions was reduced significantly by increasing the size of the three FC layers. Since the environment is richer and more complex, also the extracted features from the CNN layers might be more diverse. The FC layers of the *CNN_smallFC* model most likely are to small to deal with the increased amount of information.

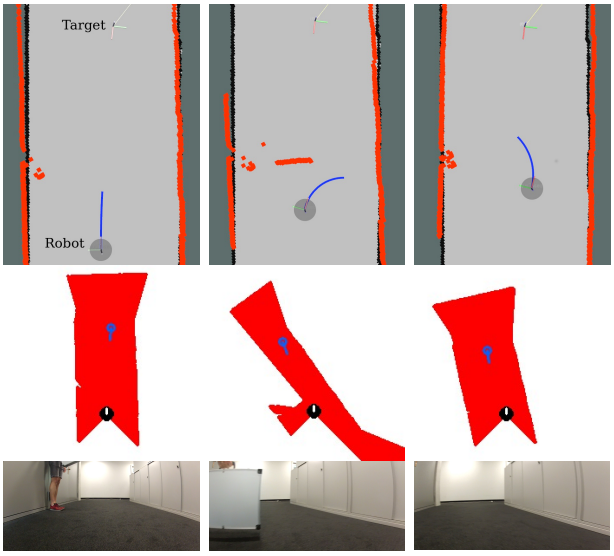
In order to test the repeatability of the deep planner results, we conducted six drives with the *CNN_bigFC*. All of those trajectories are shown in Figure 5. As in simulation, the navigation characteristics of the deep planner are similar to the one of the expert planner. Here, the deep planner reacts differently in a few areas, however still consistent between the missions. For example between targets 4 and 5, the turn with the deep planner is wider and between 6 and 7 the deep planner reacts later to the tables than the ROS planner does. This is due to the fact that only local information is used while the expert planner has global knowledge of the environment. Since the laser beams point radially from the robot, the distance to the legs of the table and chairs needs to be sufficiently small such that they are understood as obstacles by the CNN model. Between targets 10 and 11 the required human interference was the highest for both models. Considering that the robot was only trained in closed environments with relatively well-arranged obstacles, it was not able to traverse this cluttered area fully autonomously in any of the experiments.

Although the joystick interventions were required to reach some target positions, we have to mention that the majority of the interventions were required to "unstuck" the robot and not to avoid an imminent collision. In general, throughout the whole experiment, no instable motion of the robot could be observed. Interestingly, situations the robot could not handle rather caused it to stop and stand still instead of driving unpredictable and instable paths.

F. Reaction to sudden changes

This last experiment qualitatively shows the performance of the deep planner when facing suddenly appearing and disappearing obstacles. The results for this experiment are shown in Figure 6 and in the attached video. Although the navigation model only computes a single steering command, the constant velocity path for this steering command is shown for visualization purposes.

The robot has to drive from one end of a corridor to the opposite end. While the path to the target is clear in the beginning (Figure 6a), the robot is faced with a suddenly appearing object blocking its path (Figure 6b). The deep planner clearly reacts to the object by swerving to the right. After removing the obstacle, the robot corrects its path in order to approach the target as fast as possible (Figure 6c).



(a) (b) (c)

Fig. 6: Reaction to unforeseen objects. While driving towards the target position, the path of the robot gets blocked by an obstacle and afterwards it is freed again. In the top row of this figure, the constant velocity path using the computed steering commands and the overall setup are visualized. The middle row visualizes the robot-centric laser range findings and the relative target position fed to the robot while robot-view images of the last row are only used for visualization purposes.

V. DISCUSSION

In this work we showed that it is possible to learn a navigation policy from an expert operator making use of a CNN architecture for the complex end-to-end mapping function from raw sensor data to steering commands. We showed in various experiments that the learned navigation model is able to transfer gained knowledge from training environments to unseen and complex environments. One important finding of this work is that it is possible to train such a model in simulation and deploy it on a real platform while still having satisfactory navigation performance. This in as extremely valuable feature in many robotic applications where data generation is expensive.

To our best knowledge, this is the first approach that is able to perform target oriented navigation and collision avoidance based on an end-to-end approach using neural networks. Our experiments showed that the model is capable of much more than solving trivial start to goal scenarios or simple collision avoidance tasks. It is even able to solve complex navigation tasks in maze-like environments in the majority of the cases, only using local information.

Although we compare our approach to a map-based motion planner, we are fully aware that it cannot completely replace a map-based path planner. If the environment becomes more complex, the deep planner is still limited to be a local motion planner that relies on a global path planner to provide targets. At this point it also has to be clarified that our goal is not to replicate an existing planner with a CNN. Ideally, the final training data would come from a human operator that can train or re-train the robot himself. Another option might be to use training data from an optimal global planner which however is not real-time feasible. Taking several thousand trajectories from a single human demonstrator might be infeasible. Therefore, further research will show how the training samples can be reduced and how simulated and real-world training data can play together.

During the real world experiments, we found that one limitation of the current approach is wide open spaces with a lot of glass and/or clutter around. This potentially results from the fact that the model was trained purely from perfect simulation data. Training or re-training a model using real sensor data might reduce this effect.

Furthermore, we observed that the deep planner is able to avoid small dead ends if it approaches them from the outside. Once the robot enters a convex dead-end region, it is not capable of freeing itself. In addition to that, the robot's heading sometimes fluctuates before avoiding an obstacle. This issue will be further analyzed and might be solved by using recurrent neural networks with internal memory.

VI. CONCLUSION

In this work we presented a data-driven end-to-end motion planning approach for a robotic platform. Given local laser range findings and a relative target position, our approach is able to compute the required steering commands for a differential drive platform.

The end-to-end model is based on a CNN and is able to learn navigation strategies from an expert operator and transfer this knowledge between different environments. We provided an extensive evaluation for simulation and real-world experiments and showed that it is possible to train a navigation model using simulation data and deploy it on a real robotic platform in an unseen environment.

REFERENCES

[1] S. M. LaValle, *Planning algorithms*. Cambridge university press, 2006.

[2] P. Furgale and T. D. Barfoot, "Visual teach and repeat for long-range rover autonomy," *Journal of Field Robotics*, vol. 27, no. 5, pp. 534–560, 2010.

[3] K. Fragkiadaki, *et al.*, "Learning visual predictive models of physics for playing billiards," *arXiv preprint arXiv:1511.07404*, 2015.

[4] A. Punjani and P. Abbeel, "Deep learning helicopter dynamics models," in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 3223–3230.

[5] C. Cadena, *et al.*, "Multi-modal Auto-Encoders as Joint Estimators for Robotics Scene Understanding," in *Proc. of Robotics: Science and Systems (RSS)*, 2016.

[6] C. Chen, *et al.*, "Deepdriving: Learning affordance for direct perception in autonomous driving," in *Proc. of the IEEE Int. Conf. on Computer Vision (ICCV)*, 2015, pp. 2722–2730.

[7] P. Ondruska and I. Posner, "Deep tracking: Seeing beyond seeing using recurrent neural networks," in *Thirtieth AAAI Conf. on Artificial Intelligence*, 2016.

[8] P. Abbeel, *et al.*, "Apprenticeship learning for motion planning with application to parking lot navigation," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Nice, France, Sept. 2008, pp. 1083–1090.

[9] M. Kuderer, *et al.*, "Feature-based prediction of trajectories for socially compliant navigation," in *Proc. of Robotics: Science and Systems (RSS)*, 2012, p. 193.

[10] M. Pfeiffer, *et al.*, "Predicting actions to act predictably: Cooperative partial motion planning with maximum entropy models," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*. IEEE, Oct. 2016.

[11] H. Kretzschmar, *et al.*, "Socially compliant mobile robot navigation via inverse reinforcement learning," *The Int. Journal of Robotics Research*, p. 0278364915619772, 2016.

[12] S. Levine, *et al.*, "End-to-end training of deep visuomotor policies," *Journal of Machine Learning Research*, vol. 17, no. 39, pp. 1–40, 2016.

[13] V. Mnih, *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.

[14] S. Ross, *et al.*, "Learning monocular reactive uav control in cluttered natural environments," in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), 2013*. IEEE, 2013, pp. 1765–1772.

[15] D. K. Kim and T. Chen, "Deep neural network for real-time autonomous indoor navigation," *arXiv preprint arXiv:1511.04668*, 2015.

[16] U. Muller, *et al.*, "Off-road obstacle avoidance through end-to-end learning," in *Advances in neural information processing systems*, 2005, pp. 739–746.

[17] J. Sergeant, *et al.*, "Multimodal deep autoencoders for control of a mobile robot," in *Proc. of Australasian Conf. for Robotics and Automation (ACRA)*, 2015.

[18] K. He, *et al.*, "Deep residual learning for image recognition," *arXiv preprint arXiv:1512.03385*, 2015.

[19] M. Abadi, *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: http://tensorflow.org/

[20] J. Ba and D. Kingma, "Adam: A method for stochastic optimization," in *Proc. of Int. Conf. on Learning Representations (ICLR)*, 2015.

[21] M. Quigley, *et al.*, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*. Kobe, Japan, 2009, p. 5.

[22] D. Fox, *et al.*, "The dynamic window approach to collision avoidance," *IEEE Robotics & Automation Magazine*, vol. 4, no. 1, pp. 23–33, 1997.

[23] W. Xu, *et al.*, "A real-time motion planner with trajectory optimization for autonomous vehicles," in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*. IEEE, 2012, pp. 2061–2067.