Deep Imitation Learning for Complex Manipulation Task from Virtual Reality Teleoperation

Navneet Singh Arora

***Abstract*— The paradigm of mastering with the aid of using imitation has won reputation because it allows coaching complex duties with minimum professional know-how of those duties. It is a powerful technique for robot skill acquisition. However, to effectively and robustly learn models, specialized algorithms are needed as learning by imitation poses its own set of challenges [2]. Obtaining demonstrations can be challenging such that they are suitable for learning a policy that maps from raw pixels to actions [1].**

In this paper, we describe the use of consumer-grade Virtual Reality headsets and hand tracking hardware that can naturally teleoperate robots and perform complex tasks [1].

Through imitation learning, deep neural network policies are learned including mapping pixels to actions for demonstrating the acquired skills. We will also briefly evaluate the imitation learning approach through contextual raw video and imitation learning to play table tennis using image dataset, comparing them with teleoperation and proposing a new approach for skill acquisition.

I. Introduction

Imitation Learning is a technique to acquire skills through demonstrations. These demonstrations can be in the form of raw images, videos, live gestures, and now even teleoperation. Already, it has been applied in a wide range of domains in robotics like autonomous driving, gesturing and object manipulation among others.

For imitation learning to succeed, high-quality demonstrations are needed which may require a lot of data collection using a bunch of sensors and sometimes even requiring expensive hardware to proceed with the task. Demonstrations could instead be collected from running trajectory optimization or reinforcement learning, but these methods require well-shaped, carefully designed reward functions, access to dynamics models, and substantial robot interaction time. Since these requirements are challenging to meet even for robotics experts, generating high-quality demonstrations programmatically for a wide range of manipulation tasks remains impractical for most situations [1]. Another challenge encountered with imitation learning is to handle the robotic manoeuvres which

*This work was supported by Shuang Li and University Of Hamburg

Fig. 1. Virtual Reality Teleoperation for Imitation Learning (i) with PR2 Robot mounted with RGB-D camera, (ii) human operator operating using VR headset and two VR hand controllers [1].

in most cases needs to be handled explicitly depending on the shape and body of the robot. In contrast, kinesthetic teaching has also been found unsuitable, where the human operator guides the robot by force to gather demonstrations.

More than any other challenge, the biggest difference while using the demonstrations has been in the observation space. As in the case of kinesthetic teaching, there is a lot of obstruction [13] [14] for the human operator because of the unwanted appearance of robotic arms and other parts. This is how teleoperation systems were designed for high-quality demonstrations to be easily collected without any visual obstructions. Therefore, these systems are often expensive and require specialized hardware.[1].

This research paper sets to find out if such systems and learning models can be made using inexpensive or consumer-grade hardware and if practical amounts of data collection and demonstrations are sufficient enough for successfully acquiring the skills or not.

II. Related Work

A lot of work has been done before this, and two main approaches have come up for skill acquisition within imitation learning. Behavioral learning and inverse reinforcement learning are these two approaches. On one hand, behavioural learning is based on supervised learning to perform actions from observations and on the other hand, inverse reinforcement learning is where a reward function is estimated to bring near-optimal behaviour through demonstrations.

Reinforcement Learning provides a way to acquire skills through trial and error and it has proved to be a successful way of mapping pixels to actions and learning deep neural net policies. But the amount of data and exploration required is huge and is considered to be an impractical way of learning multiple tasks at once. Besides, reinforcement learning algorithms require a reward function, which can be difficult to specify in practice [3].

Behavioural Cloning on the other hand tends to learn policies directly through the raw pixels which is a better approach as state information is often not available or could not be extracted on the go. Therefore in much more natural environments, this approach has been successful such as driving [4], [5] or simulated environments [6].

Since collecting demonstrations (third-person demonstrations via raw videos as an example) for real-world manipulation is a difficult and not a suitable task, and kinesthetic teaching is not intuitive and can result in visual obstructions, teleoperation via motion capturing devices can help to solve this problem. But it poses another challenge where the human sees the object or space from a different angle than what is visible to the robot.

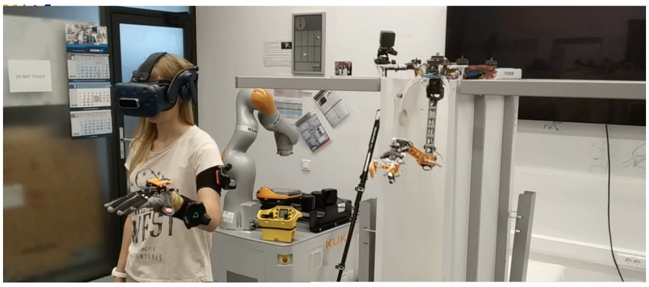

Fig. 2. Virtual Reality Teleoperation Manipulation in a controlled environment [10].

Virtual Reality Teleoperation helps resolve this issue by directly mapping the observations and actions between the human and the robot. It does not result in any issues [7] like visual obstruction, having different observational space between the human and the robot. It also leverages the natural human instinct which is not the case with any other imitation learning approach. Use of virtual reality not only helps in controlling humanoid-robots [9] but also for the communication of human intent [8].

III. Scenario: Hardware and Environment

To explore the feasibility of imitation learning through an inexpensive teleoperation system and to analyze the collection of high-quality demonstrations but limiting

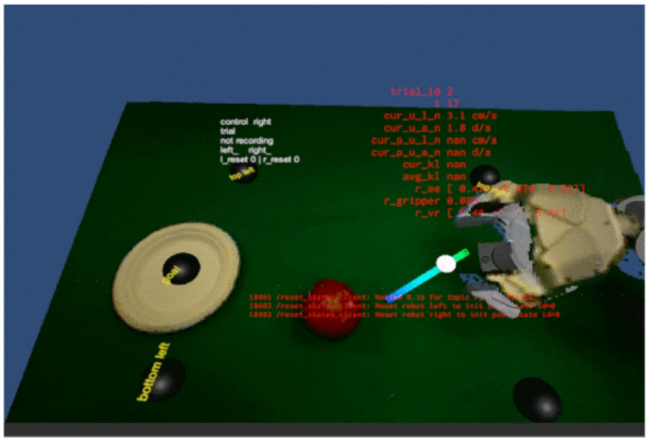

Fig. 3. First-person view from inside the VR teleoperation system during a demonstration. It shows Mapped Visual Scene through VR headset with markers and textual information to help the human operator. [1]

them to a practical amount of data for succeeding, a system is built using a consumer-grade Virtual Reality (VR) device to teleoperate a PR2 Robot [1]. The teleoperation system is based on the Vive VR teleoperation platform which provides a consumer-grade VR device (a headset for head-mounted display) and a PR2 Robot. The robot's head is mounted with a low-cost 3D camera (Primesense Carmine 1.09) which provides not only RGB images but also depth images. The teleoperation system is built on top of Unity, a 3D game engine with extensive support for VR devices.

A. Virtual Reality Teleoperation System

A set of images presenting a scene are captured through the RGB-D camera mounted on the head of the PR2 robot with per-pixel depth values. These corresponding coloured images are rendered and processed to reduce any gaps between consecutive images. Thus, instantly updating the view for the human operator to avoid any motion sickness which is observed due to 2 degrees of freedom for robot's head movement as compared to 6 degrees of freedom for the human counterpart. For precise mapping of movements and poses, 3d visualizations are overlaid (Figure-3) to assist the human operator, using markers instructing the operator to initialize certain objects during training and using arrows for indicating human control and similar other textual display.

B. Control Interface

Apart from the virtual mapping of scenes, motion-tracked VR controllers are also used to leverage the natural manipulation instincts and movements of the operator. These hand controllers also have 6 degrees of freedom with the trigger button on the controller, which is used to signal the robot gripper for an open or close state. This control interface setup helps map the human operator and the robot to a unified coordinate frame in the virtual environment where the controllers' movement through the operator are tracked to interpret the target pose of the robot's gripper [1]. Not only the movements but the controllers also help to apply and control the force precisely. Thus, allowing to dynamically vary the force as needed for example, during insertion and pushing [1].

IV. Learning Approach

The learning algorithm is based on a behavioural cloning algorithm. The algorithm learns deep visuomotor policies, directly mapping pixels to actions. The single and common neural network architecture maps raw colour and depth pixels to actions augmented with auxiliary prediction connections to accelerate learning. It uses stochastic gradient descent to train the neural network policies with batches of randomly sampled collected demonstrations to be used as an input D_{task}.

A. Overall Architecture

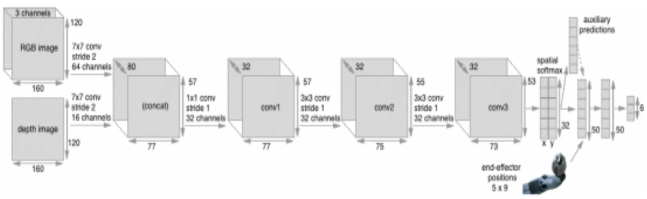

Fig. 4. Architecture of the Neural Network Policies, specifying various inputs and information to the model and the final output of the control unit [1].

The learning model involves collecting a data-set

$$D_{task} = \{(o_t^{(i)}, u_t^{(i)})\} \tag{1}$$

[1] that consists of example pairs of observation and corresponding controls. This set of information is collected through multiple demonstrations for a specific task. Along with these pair of examples, further information is captured including the depth images and auxiliary information to better learn and predict the outcome. This collected data-set, along with the additional information is then fed to the neural network as input which outputs the final control which includes the final movement, pose, and gripper position.

B. Neural Network and Policy Modules

The first step involves providing the inputs at a specific time interval to the neural network which includes the current RGB Image I_t and current Depth Image D_t. These images are captured by the RGB-D camera mounted on the head of the PR2 Robot. Apart from this, information of 5 most recent arm poses $p_{t-4:t}$ are used, collected through the end effector of the arms. Including the short history of the end effector points allows the robot to infer velocity and acceleration from the kinematic state.

$$o_t = (I_t, D_t, p_{t-4:t}) \tag{2}$$

This neural network architecture is further decomposed into three different modules:

$$\theta = (\theta_{vision}, \theta_{aux}, \theta_{control}) \tag{3}$$

1) Vision Module: This module uses a Convolutional Neural Network (CNN) extracting spatial feature points from the images of the observation o_t input.

$$f_t = CNN(I_t, D_t; \theta_{vision}) \tag{4}$$

The spatial features not only include the magnitude of the movement but also the direction of the movement of the end effector. Thus introducing directional alignment (Eq. 5) between the demonstrated controls and the network output.

$$\mathcal{L}_c = \arccos\left(\frac{u_t^T \pi_\theta(o_t)}{\|u_t\| \|\pi_\theta(o_t)\|}\right) \tag{5}$$

This however is slightly different from the standard behavioral cloning where the loss function only takes care of the magnitude and not the direction of the movement and typically uses l1 (Eq. 6) and l2 (Eq. 7) losses to fit the training [11].

$$\mathcal{L}_{l1} = \|\pi_\theta(o_t) - u_t\|_1 \tag{6}$$

$$\mathcal{L}_{l2} = \|\pi_\theta(o_t) - u_t\|_2^2 \tag{7}$$

2) Auxiliary Module: The Neural Network (NN) in this module follows the CNN of vision module for auxiliary prediction.

$$s_t = NN(f_t; \theta_{aux}) \tag{8}$$

This module uses the collection of the dataset that consists of an example of both observations and its corresponding control as a pair. The main aim of including this Auxiliary Prediction Task is to accelerate the learning process as similar approaches that leverage self-supervisory signals have been shown [12] to improve data efficiency and robustness.

Auxiliary Prediction Task leverages Self-Supervision Task which takes the spatially extracted features f_t and infers the labels from the readily available data-set D_t including the current gripper pose p_t and final gripper pose p_T.

The Loss Function which is added on top of the spatial layer, does the actual auxiliary prediction and then uses the labels s_t, fetched from the pre-trained process at a particular time frame. Thus, the Auxiliary loss function takes the extracted spatial features, uses the spatial features to predict the auxiliary movement, and then applies the pre-learning s_t for more efficiency and robustness.

$$\mathcal{L}_{aux}^{(a)} = \|NN(f_t; \theta_{aux}^{(a)}) - s_t^{(a)}\|_2^2 \tag{9}$$

All the training done through this approach is concurrent and does not involve any additional data-set for pre-training of CNN [11] as in most other visuomotor policies approaches.

3) Control Module: The Neural Network in this module is the final layer of the architecture which outputs the control. This output consists of angular velocity w_t and linear velocity v_t.

$$u_t = NN(p_{t-4:t}, f_t, s_t; \theta_{control}) \tag{10}$$

The desired gripper state for the specific task which involves grasping, the final layer outputs a gripper open/-close prediction $g_t \in 0, 1$, which is trained using sigmoid cross entropy loss [1].

$$\mathcal{L}_g = g_t \log(\sigma(\hat{g}_t)) - (1 - g_t)\log(1 - \sigma(\hat{g}_t)) \tag{11}$$

4) Loss Function: The overall loss function is a weighted combination of standard loss functions of all the three modules

$$\mathcal{L}(\theta) = \lambda_{l2}\mathcal{L}_{l2} + \lambda_{l1}\mathcal{L}_{l1} + \lambda_c\mathcal{L}_c + \lambda_{aux}\sum\nolimits_a \mathcal{L}_{aux}^{(a)} \tag{12}$$

For tasks involving grasping, the overall loss function is described as:

$$\mathcal{L}(\theta) = \lambda_{l2}\mathcal{L}_{l2} + \lambda_{l1}\mathcal{L}_{l1} + \lambda_c\mathcal{L}_c + \lambda_g\mathcal{L}_g + \lambda_{aux}\sum\nolimits_a \mathcal{L}_{aux}^{(a)} \tag{13}$$

V. Experiment

The experiment setup was done choosing in total 10 challenging manipulation tasks where the robot must (a) reach a bottle, (b) grasp a tool, (c) push a toy block, (d) attach wheels to a toy plane, (e) insert a block onto a shape-sorting cube, (f) align a tool with a nail, (g) grasp and place a toy fruit onto a plate, (h) grasp and drop a toy fruit into a bowl and push the bowl, (i) perform grasp-and-place in sequence for two toy fruits, (j) pick up a piece of dishevelled cloth (Fig. 5) [1].

These tasks were chosen as they have been considered as a benchmark for other state-of-the-art algorithms

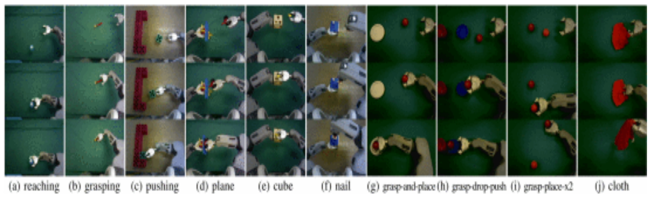

Fig. 5. Set of 10 manipulation tasks being performed during evaluation. [1].

but also as they require different aspects of robotic manipulation [11] involving object localization (a, b, c, g, h, i), high-precision control (a, f, e), managing simple deformable objects (j), and handling contact (c, d, e, f, h, i), all on top of good generalization.

The experiment was conducted in a controlled environment and was not performed iteratively but all the interactions with the robot for a specific task were performed in a single session. Similarly, no amount of desired noise was injected which is usually done in methods like GPS [11] and natural human error and variations were assumed to be sufficient while performing the demonstrations (Fig. 6).

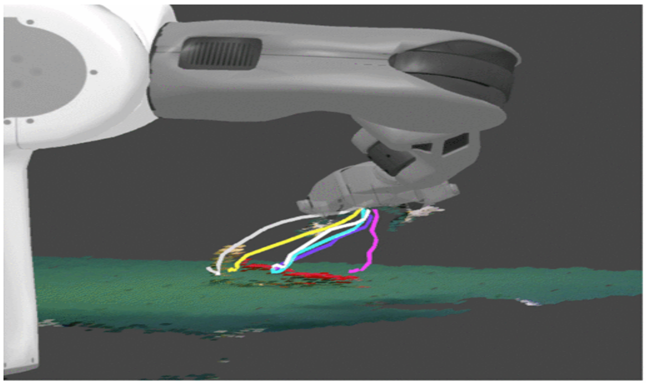

Fig. 6. This shows six different demonstration trajectories followed by the human operator while doing the demonstration from the same initial state for a specific grasping task [1]. This natural variation helps avoid injecting the noise.

VI. Results

As the aim of this research was to find out the feasibility of teleoperation, therefore a small number of demonstrations were used with all demonstrations under 30 minutes for every task to explore the effectiveness.

task	reaching	grasping	pushing	plane	cube	nail	grasp-and-place	grasp-drop-push	grasp-place-x2	cloth
test	91.6%	97.2%	98.9%	87.5%	85.7%	87.5%	96.0%	83.3%	80%	97.4%
demo time (min)	13.7	11.1	16.9	25.0	12.7	13.6	12.3	14.5	11.6	10.1
avg length (at 10 Hz)	41	37	58	47	37	38	68	87	116	60
# demo	200	180	175	319	206	215	109	100	60	100

Fig. 7. Table showing demonstration time, the total number of demonstrations and the success rate for all 10 tasks [1].

For such a small amount of demonstrations (below 30 minutes) for each task, the model was able to achieve high success rates. Not only were the tasks completed but also portrayed a good command of the acquired skill. For example, the robot was able to balance the block to maintain the correct direction using a single point of contact. Another important aspect noted has been the transition from one skill to the other and therefore performing a complete sequence of manoeuvres which has not been the case with any prior state-of-the-art algorithms.

It was worth noting that the increase in the total number of demonstrations and the duration of demonstrations lead to a much-improved success rate. With only 5 minutes of demonstration time leading to 50% of success rate and going up to 90% with 12 minutes of demonstration for nail task. As a slight variation to

task: nail		
number of demonstrations	demonstration time (estimated) (min)	success rates
193	12.2	88.9%
115	7.3	77.8%
67	4.2	50%

Fig. 8. Success rates of policies trained using different numbers of demonstrations for the nail task [1].

a normal behavioural cloning algorithm, the auxiliary prediction was involved as it proved to be successful in attaining data efficiency [12]. For grasp and place tasks and with the same amount of demonstrations, a significant amount of improvement was observed (Table in Fig. 9).

task: grasp-and-place		
number of demonstrations	success rates (with)	success rates (without)
109	96%	80%
55	53%	26%
11	28%	20%

Fig. 9. Comparison of a policy when trained with and without auxiliary prediction loss on the grasp-and-place task. [1].

A. *Suboptimal Behaviors*

While achieving success in most of the scenarios and metrics, the model was often suboptimal compared to human demonstrations. Taking a longer path for doing the task as compared to the human counterpart, slow movements, or pausing entirely in-between a task before continuing to execute the task, and accidentally nudging the object while attempting to grasp are a few of the suboptimal scenarios.

B. *Failure Cases*

For each task, multiple failure behaviours were reported as well: (a) Not able to operate the gripper with perfection (gripper state change did not happen), (b) Refusing to move after altering the state of the object, (c) Distance limitations ($<$ 3mm) for holding or keeping the object.

C. *Extrapolation*

On further evaluation, the policy for the nail task could generalize to new hammer orientations and positions [1]. However, in the reaching task, the policy rejected a previously unseen green bottle when it was present along with the training bottle.

VII. Evaluation

The experiment conducted turns out to be very effective in terms of learning sequence of manipulation tasks with limited amount of data made available for learning. The success rate is quite high and with increasing the number of demonstrations, it can be further improved. Apart from this, the auxiliary signal helps improve the model efficiency significantly.

Overall, despite having failure and sub-optimal scenarios, the result is sufficient enough to suggest the feasibility of this model and approach.

VIII. Other Imitation Learning Models

A. *Imitation from Observation: Learning to Imitate Behaviors from Raw Video via Context Translation*

The goal in imitation-from-observation is to learn policies only from a sequence of observations (which can be extremely high dimensional, such as camera images) of the desired behavior, with each sequence obtained under different context including changes in the environment, changes in the objects being manipulated, and changes in viewpoint, while observations might consist of sequences of images [15]. The approach is based on deep reinforcement learning and does not involve learning spatial features as done in our main approach using teleoperation but instead the translation (mapping pixels to action) is based on the reward functions. Similar to the input data-set D_{task} , this model also uses the pair of input samples, the source context image w_i and target context image w_j. On the contrary, the reward function corresponds to the squared Euclidean Distance between the observation and the demonstrations and its objective is to minimise it.

The method was successfully able to learn the tasks and outperformed all other baseline methods including the ones using pre-trained visual features. Despite the success, it requires huge data-set with 3000 - 5000 videos per experiment. The depth images are not available and magnitude of the force to be applied could not be translated to the final outcome. The major aspect which differs in this approach is the non availability of skill transition from one task to another.

B. Imitation for a robot to play Table Tennis

This model is completely different from the other two models presented so far. The model is heavily based on the image processing aspect for feature extraction using Canny Edge Detection and RANSAC [16]. Video demonstrations are fed as an input to the system, thus calculating the 3D coordinate system for the position and center of the racquet and the videos are processed frame by frame. These per frame images are processed extracting the racquet and using the synchronised image and camera coordinate system, the trajectories are calculated to mimic the exact behaviour.

With this approach, the racquet trajectories were mapped precisely with no difference between the original and fitted trajectories. Even after such computation, only the trajectories were mapped and the model was not able to mimic the entire gameplay. In order to achieve that, a lot more computations are needed to be done including ball trajectories, human movements, direction of movement and even the force to be applied to play a shot.

IX. Discussion

In the current approach, teleoperation through a VR system (Imitation from Action) is used to learn the manipulation of complex tasks. This approach uses the pre-trained Convolutional Neural Network (CNN) for gripper information. The RGB-D captured information is virtually mapped with various textual and visual indications for instructing the human operator and this collective information is then processed through Neural Network policies. Finally, overall spatial information, state information, and current information are fetched through the control interface for the final execution of the task.

A. Combined Approach

Here we propose a new approach for imitating a human to play table tennis. In this combined approach, we suggest making use of the current image processing method, where the racquet trajectories have already been calculated. We can use these pre-calculated trajectories to train our Convolutional Neural Network (CNN) for racquet information as well as using the raw images showing various states (angles, position, backside/ frontside) of the racquet, which will in turn speed up the whole training process similar to the auxiliary learning in this approach. Through the same RGB-D captured images, we can process the live imagery of the ball, allowing common observation space to be enabled via a VR headset for the human operator/player. Then finally through the control interface, taking into account the human instinct, real-time arm movement can be done enabling proper gameplay. Therefore, we will have the required auxiliary information and spatial information including the angle of movement, angular and linear velocity, along with the velocity of the ball which can then be applied to play the shot.

B. Advantages and Disadvantages

This approach has its advantages and disadvantages. It will require head movement, leg movement, and torso movement to reach the required degrees of freedom similar to the human operator which adds a lot of complexity. Therefore, a lot of pre-training needs to be done to get the spatial movement information for different parts of the robot and not just the arm. 3D visualization, various markers, textual and visual information can be overlaid on the scene for the human operator, achieving a similar level of preciseness but requiring more computational power as the object in table tennis is a moving ball instead of a static object. Therefore, a lot more sensors, capturing equipment will be needed which will make it quite expensive. On the other hand, it takes away the need to calculate the trajectories for the ball and spin on the ball which otherwise would have taken a similar kind of computation for model learning and this is already expensive and involving large duration of the learning period. Even though it will be expensive with much more hardware, as seen with the current research, the demonstration time and learning period will be relatively small.

C. Limitations

Despite having different good models, there are certain unknowns and limitations which are needed to be handled and their feasibility needs to be assessed before moving on with the suggested approach. Pre-training of CNN will change depending on the task to be accomplished, in this case, sport to be played. Also, a sport involves a lot more information apart from just hitting the ball, thus capturing this information and training the model could turn out to be difficult to handle. Not only this, pre-training will be needed based on the rules of the games as well. The controlling interface will also have to be modified and improved. As in the case of manipulation, the controllers had a button for sending out the signal to the robot to access the gripper state but such information is not required to play the sport. Moreover, a lot more controllers and sensors would be needed to precisely train the robotic movements.

Overall, the core elements of this approach remain the same. Also as in sports, the transition from one movement to another is an important aspect. Therefore, this approach of teleoperation through a control interface can prove to be fruitful while playing table tennis as transitioning from one shot to another quickly can bring it closer to how humans play. Thus, this approach can be looked into for feasibility to enable entire gameplay through Imitation Learning via Teleoperation.

X. Conclusion

In this paper, an approach was presented for imitation learning using VR Teleoperation. In a controlled environment, this approach did wonders considering the number of demonstrations and the duration of demonstrations done to achieve the outcome successfully. It leads to natural transition from one manoeuvre to another, similar to a humans instinct. Moreover, it required no task specific variations from the initial state of the model and was able to perform with single neural network architecture using same hyper parameters. We gathered enough evidence and further presented other two successful approaches of imitation learning and thus presented a new model approach, which brings together the advantages of every single model but most importantly the control interface characteristics to incorporate not only the magnitude and direction but also the natural instinct and transition from one point to another which is the most important aspect of gameplay in any sport.

It promises to bring much required improvement in the field of robot sports and can also helps achieve it at a faster pace.

References

[1] Zhang, Tianhao, et al. "Deep imitation learning for complex manipulation tasks from virtual reality teleoperation." 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018.

[2] Hussein, Ahmed, et al. "Imitation learning: A survey of learning methods." ACM Computing Surveys (CSUR) 50.2 (2017): 1-35.

[3] Ng, Andrew Y., and Stuart J. Russell. "Algorithms for inverse reinforcement learning." Icml. Vol. 1. 2000.

[4] Pomerleau, D. "An autonomous land vehicle in a neural network." Advances in Neural Information Processing Systems; Morgan Kaufmann Publishers Inc.: Burlington, MA, USA (1998).

[5] Bojarski, Mariusz, et al. "End to end learning for self-driving cars." arXiv preprint arXiv:1604.07316 (2016).

[6] S. Ross, G. J. Gordon and D. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning", AISTATS, vol. 1, no. 2, pp. 6, 2011.

[7] Argall, Brenna D., et al. "A survey of robot learning from demonstration." Robotics and autonomous systems 57.5 (2009): 469-483.

[8] Rosen, Eric, et al. "Communicating robot arm motion intent through mixed reality head-mounted displays." Robotics research. Springer, Cham, 2020. 301-316.

[9] Stanton, Christopher, Anton Bogdanovych, and Edward Ratanasena. "Teleoperation of a humanoid robot using full-body motion capture, example movements, and machine learning." Proc. Australasian Conference on Robotics and Automation. Vol. 8. 2012.

[10] Yashin, Grigoriy A., et al. "Aerovr: Virtual reality-based teleoperation with tactile feedback for aerial manipulation." 2019 19th International Conference on Advanced Robotics (ICAR). IEEE, 2019.

[11] Levine, Sergey, et al. "End-to-end training of deep visuomotor policies." The Journal of Machine Learning Research 17.1 (2016): 1334-1373.

[12] aderberg, Max, et al. "Reinforcement learning with unsupervised auxiliary tasks." arXiv preprint arXiv:1611.05397 (2016).

[13] Akgun, Baris, et al. "Keyframe-based learning from demonstration." International Journal of Social Robotics 4.4 (2012): 343-355.

[14] Dragan, Anca D., Siddhartha S. Srinivasa, and Kenton CT Lee. "Teleoperation with intelligent and customizable interfaces." Journal of Human-Robot Interaction 2.2 (2013): 33-57.

[15] Liu, YuXuan, et al. "Imitation from observation: Learning to imitate behaviors from raw video via context translation." 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2018.

[16] Wang, Qizhi, Xianqiang Zhang, and De Xu. "Human behavior imitation for a robot to play table tennis." 2012 24th Chinese Control and Decision Conference (CCDC). IEEE, 2012.

Eidesstattliche Erklärung

Hiermit versichere ich, Navneet Singh Arora, an Eides statt, dass ich die vorliegende Seminararbeit mit dem Titel *Deep Imitation Learning for Complex Manipulation Task from Virtual Reality Teleoperation*, sowie die Präsentationsfolien zu dem dazugehörigen mündlichen Vortrag ohne fremde Hilfe angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Alle Teile, die wörtlich oder sinngemäß einer Veröffentlichung entstammen sind als solche kenntlich gemacht.

Die Arbeit wurde in dieser oder ähnlicher Form noch nicht veröffentlicht, einer anderen Prüfungsbehörde vorgelegt oder als Studien- oder Prüfungsleistung eingereicht.

Declaration of an Oath

Hereby I, Navneet Singh Arora, declare that I have authored this thesis, titled *Deep Imitation Learning for Complex Manipulation Task from Virtual Reality Teleoperation*, and the presentation slides for the associated oral presentation independently and unaided. Furthermore, I confirm that I have not used other than the declared sources / resources.

I have explicitly marked all material which has been quoted either literally or by content from the used sources.

This thesis, in same or similar form, has not been published, presented to an examination board or submitted as an exam or course achievement.

Hamburg, February 17, 2021

Navneet Singh Arora