Vilnius University

An investigation of deep imitation learning for mobile robot navigation

Shubham Juneja
Supervisor: Dr. Virginijus Marcinkevičius
Semester 3

Vilnius University

Plan of studies & implementation summary

Study year	Exams		Conference participations		Publications	
	Planned	Completed	Planned	Completed	Planned	Completed
I (2020/2021)	2	2	1	1		
II (2021/2022)	2	2				
III (2022/2023)			0	0	1	0
IV (2023/2024)			1	0	1	0

Vilnius University

Report of activity plan

Exams		Conference Participation		Publications	
Planned	**Status**	**Planned**	**Status**	**Planned**	**Status**
Machine Learning	Passed with 9/10	All Sensors 2021, Nice, France	Paper accepted and presented at All sensors 2021 conference at Nice, France. On the 20th of July.	Idea paper with the title “Combining Multiple Modalities with Perceiver in Imitation-based Urban Driving”	Published
Research methods and methodology of informatics and computer engineering	Passed with 9/10	Planned participation at International conference on Robotics & AI, Singapore.	Research ongoing, Deadline 1st of June.		
Fundamentals of informatics	Passed with 7/10				
Optimisation	Passed with 7/10				

Workshops

Workshop	ECTS
MOKSLINIŲ REZULTATŲ PUBLIKAVIMAS PAGAL FORMALAUS VERTINIMO REIKALAVIMUS	0.1
MOKSLINĖS INFORMACIJOS IŠTEKLIAI, PAIEŠKA, IR ĮRANKIAI	0.1
MENDELEY PRAKTINIS UŽSIĖMIMAS	0.15
DeepLearn Summer School (Planned July ’22)	
Total:	0.35/3

Stages of research and dissertation preparation

Vilnius University

	Name of task	Duration	Notes
2.	Carrying out research: **2.1. Development of research methodology:** 1. Identification and specification of problems arising in currently available methods. 2. Specification of tasks to conduct which address to identified problems. (GET OBJECTIVES) 3. Specification of navigation environments which will be analysed further. 4. Selection of appropriate research methodology. 5. Planning of theoretical and empirical research.	September 2021 – October 2021	Problems in imitation learning identified. Navigation environments and methods identified.
	2.2. Theoretical research: 1. Analysis of reactive imitation learning methods for sensorimotor control and strategy functions, which utilize deep neural networks, such as behaviour cloning, inverse reinforcement learning, generative adversarial imitation learning, etc.	November 2021 – February 2022	Theoretical methods studied.
	2.2. Theoretical research: 2. Research on new reactive mobile robot navigation trajectory controller, based on learning from experience (e.g. imitation learning, reinforcement learning). 3. Research on hierarchical goal-directed visual navigation system for mobile robots, based on aforementioned reactive component.	March 2022 – May 2022	
3.	**2.3. Empirical Research:** 1. Implementation of results of 2.2.2 and 2.2.3 to improve the state-of-the-art trajectory controller and navigation methods.	May 2022 – August 2022	

Vilnius University

Research Object and Aim

Research object:

- Deep imitation learning methods.
- Application of deep imitation learning methods for mobile robot navigation.

Research aim:

- To develop, implement and research an autonomous navigation system for mobile robots based on imitation learning and deep neural networks

Vilnius University

Objectives of Research

1. To **develop and investigate** new sensorimotor reflex algorithms based on deep neural networks and various simulation learning paradigms (e.g. behaviour cloning, generative adversarial imitation learning) (e.g. trajectory following, obstacle avoidance, approach to a recognized object).
2. To **compose and implement** a new navigation system for mobile robots from the obtained sensorimotor reflexes.
3. To **compare** the obtained navigation system with alternative robot navigation algorithms.
4. To **prepare publicly available datasets** for the research of autonomous robot navigation algorithms based on the principles of deep neural networks and imitation training.

Vilnius University

What has been carried out so far

- Literature study from papers on imitation learning for mobile robot navigation
- Took courses:
    - Machine learning (at VU)
    - Research methodology (at VU)
    - Fundamentals of Informatics (at VU)
    - Optimisation (at VU)
    - Reinforcement learning (Online)
- Trying out Simulators (CARLA and OpenAI gym)
- Attempted to run state of the art methods in simulation
- Participation in an international conference

Literature Review (Continued)

Learning to imitate

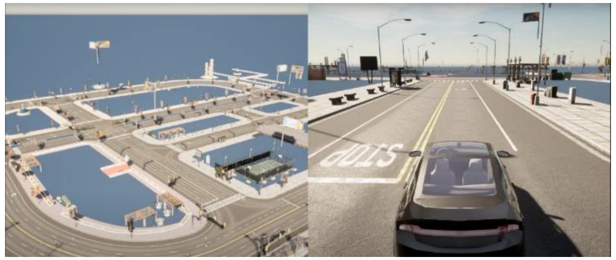
Vilnius University

- In imitation learning:
  - Given: Demonstrations
  - Goal: Train a policy (model) to mimic demonstrations
- Being a form of machine learning, data is collected, models are optimized, accuracies are evaluated.

Vilnius University

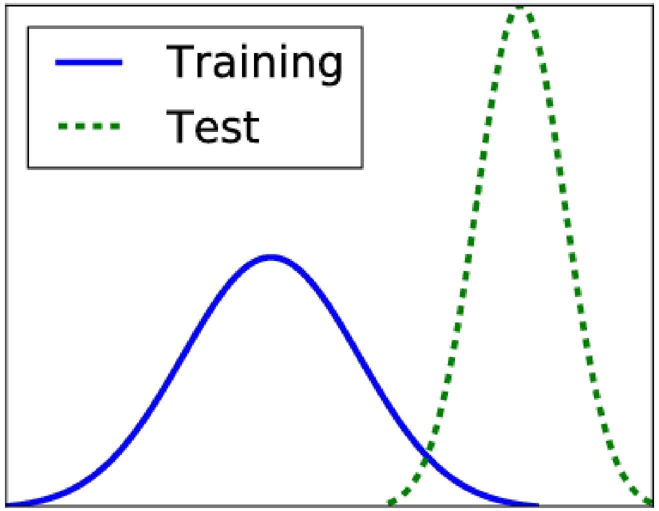
About the problem to solve

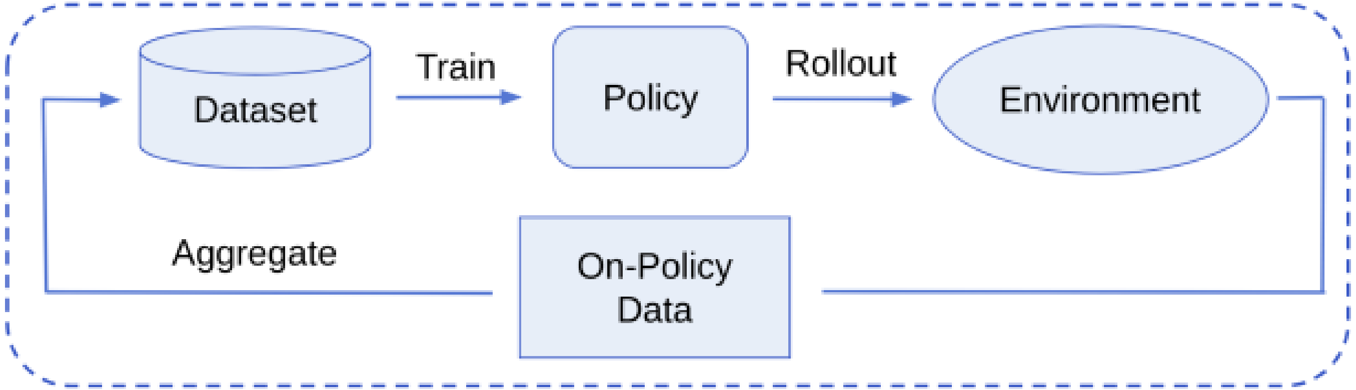
- Learning sensorimotor skills to drive and navigate based on visual input.
- It can be done with traditional methods such as SLAM, but it would require expensive sensors and extensive programming.
- The idea of imitation learning promises to solve this problem by learning from human demonstrations.
- Yet, it remains unsolved due the unpredictability of the real world causing the problem of covariate shift.
- To compare the ability between methods NoCrash benchmark has been established.
- NoCrash benchmark uses CARLA simulator to seed vehicles in different parts of a map and tests the ability of reaching from point A to B, under different sets of conditions.



DAgger or Data Aggregation

- Big issue in imitation learning is the problem of covariate shift.
- Data aggregation is a method for solving covariate shift.





Stephane Ross, Geoffrey J. Gordon, and Drew Bagnell. A ´ reduction of imitation learning and structured prediction to no-regret online learning. In Conference on Artificial Intelligence and Statistics (AISTATS), 2011. 1, 2, 5

Concepts for solving Covariate Shift

- Data augmentation:
  - Coming up with such augmentations which transform a given distribution into something closer to a distribution far from the given.
  - What has been done so far:
    - Camera and angle shifting
    - Blurring, cropping, changing brightness, adding noise to image (Standard methods)

Fig. 2. (a) An original and (b) a synthesized image from [5]. The synthesized image looks as if the car has drifted towards the center of the road.

Concepts for solving Covariate Shift

Vilnius University

- Data diversification:
  - During data collection / generation, the data can become very monotonous.
  - This has been attempted to solve by adding random noise during data collection.
  - This creates unexpected deviations during the expert has control and challenges the expert to tackle such deviations.

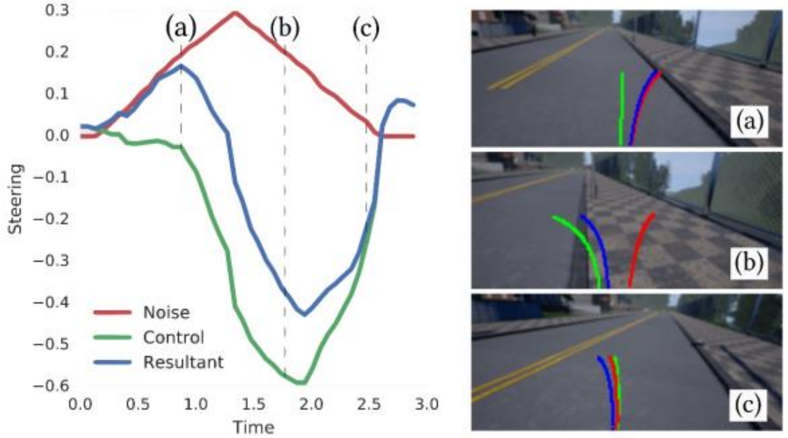


Fig. 3. Data diversification in [6]. Noise is injected during data collection

Concepts for solving Covariate Shift

- On-policy learning:
  - While application of policy the controller faces inputs which are out of distribution, this can be corrected and collected using dataset aggregation methods (DAgger).
  - Core DAgger methods require expert monitoring which demands high resources.
  - Methods like SafeDAgger tackle problem of removing the expert out of the equation.

Concepts for solving Covariate Shift

Vilnius University

- Dataset balancing:
  - Datasets are dominated by common behaviours and hence the trained algorithms suffer at scenarios where data is uncovered or less likely to be sampled.
  - To tackle this, methods such as weighted sampling are used.
  - Weighting down the datapoints which resemble going straight and up for turns.
  - Balancing has also been explored during mini-batch generation.
  - The idea is to see the hard-to-solve cases more often.

Reinforcement Learning for End-To-End Navigation

Vilnius University

- Reinforcement learning (RL) is a machine learning paradigm where a system learns to maximize the rewards given to it by acting in an environment.
- RL is known for being less data efficient, although the concept promises exploring data on its own and generating the required distribution.
- To solve navigation, the RL methods such as the following have been widely applied:
    - Policy gradients
    - Deep Q-learning
    - Proximal policy optimisation

RL (continued)

- To tackle the data efficiency problem, policies have been tried to be trained with behaviour cloning first and then moved into a RL setting.
- This approach promises a head start in learning.

Concepts for solving RL based learning

- Further areas where RL can be explored are:
  - Rewards:
  - Learning in real world
  - Learning in Simulation

Concepts for solving RL based learning

- Rewards:
  - Rewards shape the learned behaviours in RL based methods.
  - The simpler the rewards are, the easier the learning becomes.
  - Simple rewards can be:
    - High reward for being in the centre of the lane
    - High reward for driving on road
    - High penalty for coming in contact with side-walks
    - Very high penalty for crashes
    - High penalty for driving in the wrong lane
    - Some penalty getting close to other vehicles
  - Over all reward policy can also be learned using inverse RL and is usually a combination of the mentioned.

Concepts for solving RL based learning

- Learning in the real world:
  - Challenge is the not incur any damage in the real world while exploration.
  - Some approaches have make analytical rules of transferring the control to an expert.
  - While some approaches stop the experiments if defined rules are breached.

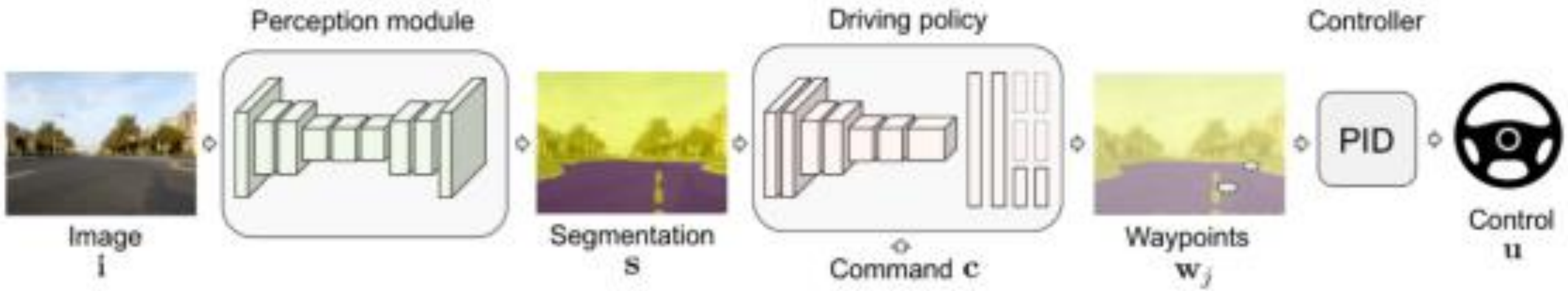
Concepts for solving RL based learning

- Learning in simulation:
  - This is the most efficient way of learning and exploring without causing real world damage.
  - There are simulators available which attempt to model the real world.
  - They are :
    - CARLA simulator
    - GTA 5 computer game
  - Although the problem arises when deploying the policy to real world.

Transfer from simulation to real world

Vilnius University

- Few methods to transfer to real world have been attempted:
  - In overall, the end-to-end methods form an intermediate representation for simulation and real world.
  - Once this mapping is done, the driving is then trained.
  - A well known way of mapping is with image segmentation.



Vilnius
University

Published work

On record:

- Conference: All sensors 2021
- Participation type: Idea paper

Combining Multiple Modalities with Perceiver in Imitation-based Urban Driving

Shubham Juneja
Institute of Data Science &
Digital Technologies
Vilnius University
Vilnius, Lithuania
Email: shubham.juneja@mif.stud.vu.lt

Virginijus Marcinkevičius
Institute of Data Science &
Digital Technologies
Vilnius University
Vilnius, Lithuania
Email: virginijus.marcinkevicius@mif.vu.lt

Povilas Daniušis
Department of Business
Technologies & Entrepreneurship
Vilnius Gediminas Technical University
Vilnius, Lithuania
Email: povilas.daniusis@vgtu.lt

Off record:

- Journal “Springer: Autonomous Robots”
- Impact factor: 3.6

SpringerLink

Published: 04 May 2021

Topological navigation graph framework

Povilas Daniušis, Shubham Juneja, Lukas Valatka & Linas Petkevičius

Autonomous Robots (2021) | Cite this article

Combining multiple modalities with Perceiver in IL based learning

- We present a study pointing out how end-to-end methods rely on a single modality while lacking the performance compared to traditional autonomous driving methods which take a modular approach.
- Therefore, we propose a method to enrol more than one modality in the learner.
- We propose the use of a perceiver architecture in the learner as this architecture shows capability of learning with varying number and types of modalities as input data.
- Since the published paper is a idea paper, no experiments were presented.

Work plan for semester 4

Vilnius University

Research:

- Theoretical research continuation.
- Implementation of state-of-the-art trajectory controller and navigation methods.

Summer school participation:

- DeepLearn summer school, July 2022, Spain

Conference Participation:

- International Conference of Robotics and AI, Singapore

Vilnius University

Thank you