WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

A Bayesian Approach to Reinforcement Learning of Vision-Based Vehicular Control

Zahra Gharaee, Karl Holmquist, Linbo He, Michael Felsberg

Computer Vision Laboratory (CVL), Department of Electrical Engineering, University of Linköping

{zahra.gharaee, karl.holmquist, linbo.he, michael.felsberg}@liu.se

Bayesian Approach to RL

Environment: CARLA Simulator

Fig.1

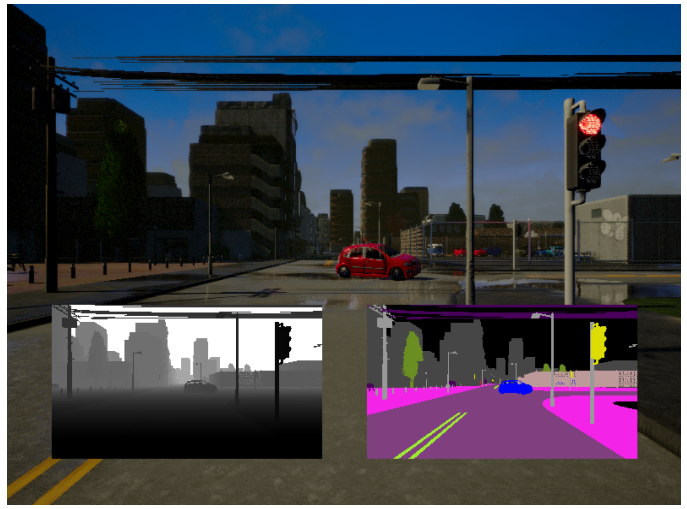


Fig.2

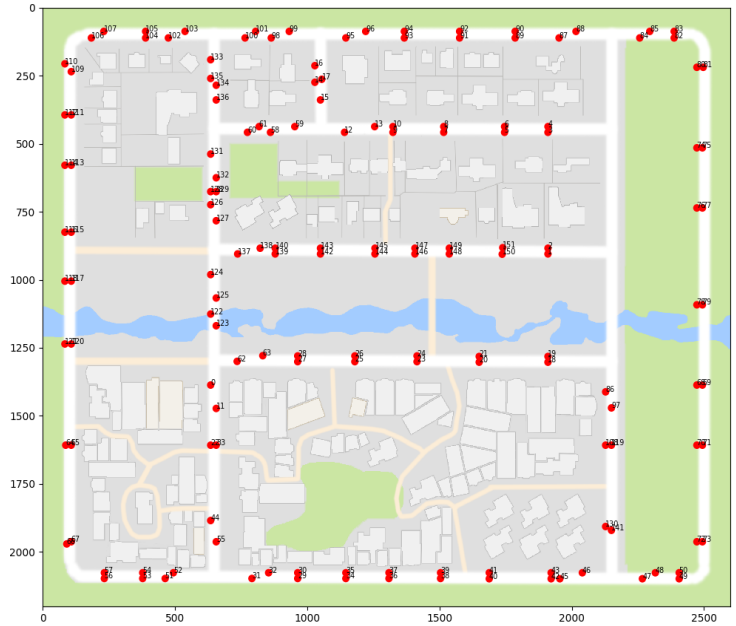


Fig.3

Images from ”CARLA: An Open Urban Driving Simulator”

Bayesian Approach to RL

- Input: Vehicle Front View Camera
- Decision Making: Epsilon Greedy Policy
- Learning Strategy: On-Policy Probabilistic Q-Learning
- Loss Function: Temporal Difference Error

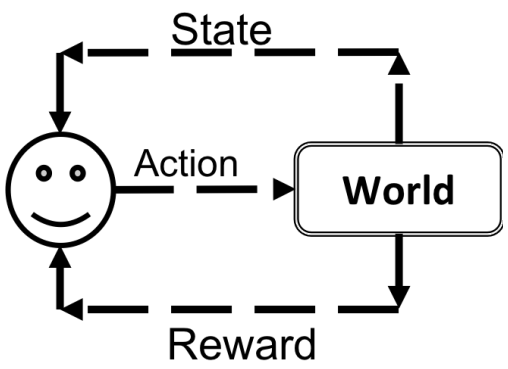


Fig.1

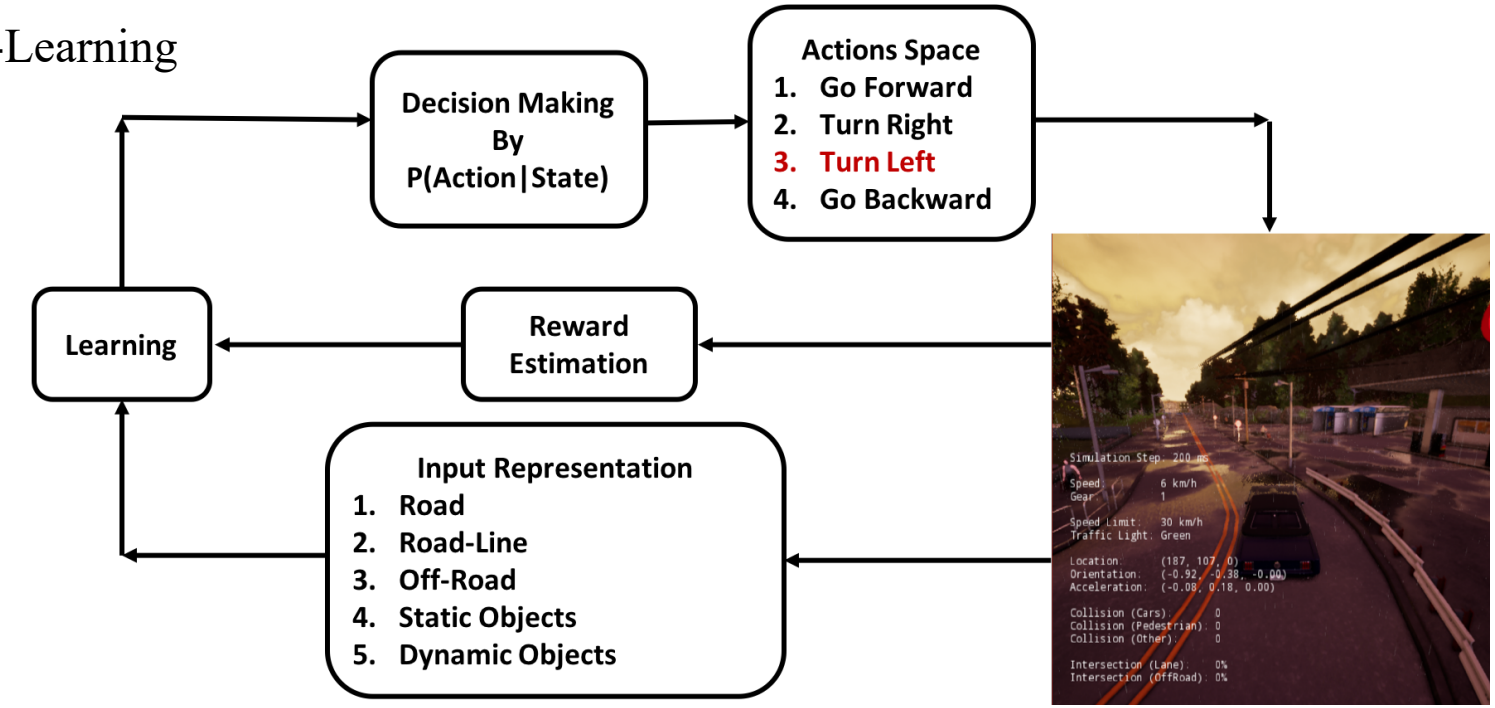
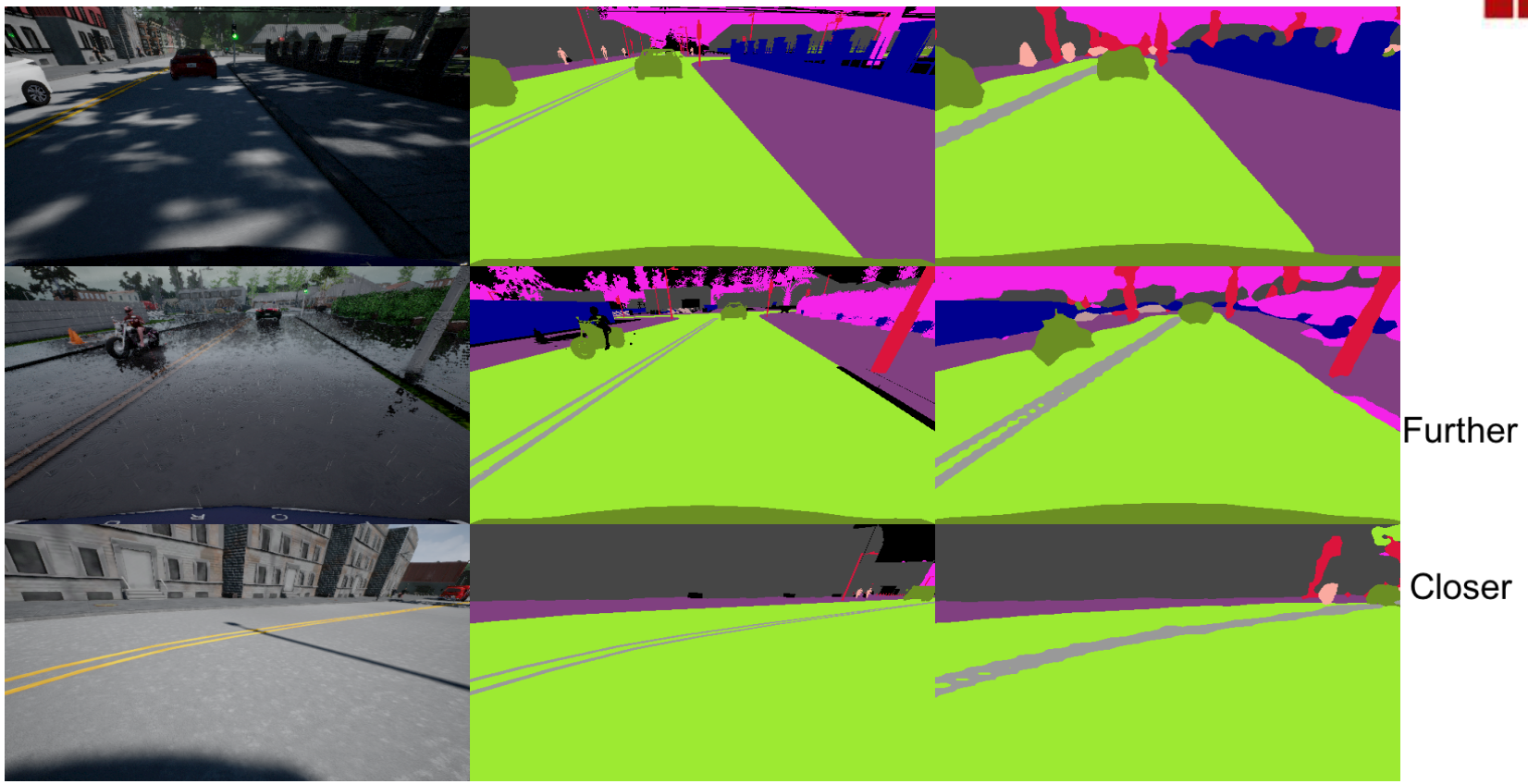


Fig.2

Bayesian Approach to RL

Input



RGB Ground-Truth Estimation

Fig.1

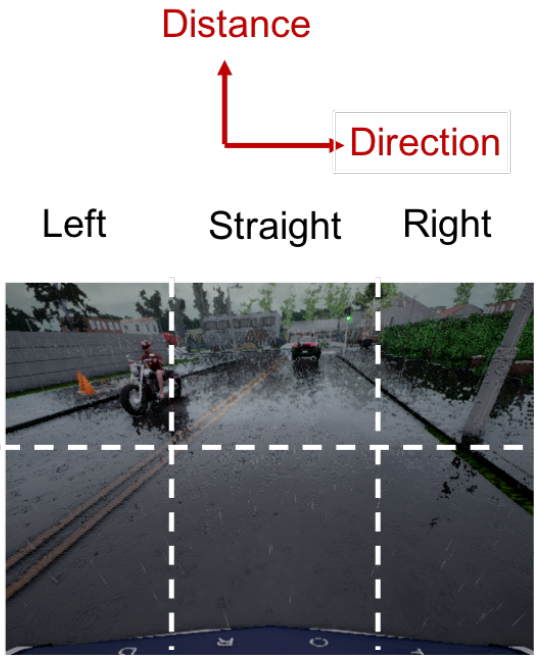


Fig.2

Bayesian Approach to RL Learning Paradigm

- Clustering input states through a GMM
- Calculating action probabilities based on Bayes rule and the component wise probabilities p(m|a)
- Applying standard discrete Q-Learning through state-action table

Learning Rate

State-Action Value Function

Temporal Difference Error

$$Q(m_t, a_t) \leftarrow Q(m_t, a_t) + \alpha\, TD_{error}$$

Equation (1)

$$TD_{Error} = Reward + \gamma Q(m_{t+1}, a_{t+1}) - Q(m_t, a_t)$$

Equation (2)

Forgetting Factor

Bayesian Approach to RL

Reward Signal

$$r = \begin{cases} -r_{k_1} & \text{if Collision} \\ -r_{k_2} \cdot r_o, & \text{else if Off-road} \\ -r_{k_3} \cdot r_l, & \text{else if Opposite-lane} \\ r_{\text{speed}}, & \text{else,} \end{cases}$$

$$r_t = r + r_{\text{road-view}}.$$

$$r_{\text{speed}} = \begin{cases} -r_{k_4} \cdot \left(\frac{v_t - v_{\text{target}}}{v_{\text{target}}}\right)^2, & \text{if } v_t < 0 \\ -r_{k_5} \cdot \left(\frac{v_t - v_{\text{target}}}{v_{\text{target}}}\right)^2, & \text{if } 0 < v_t < v_{\text{target}} \\ 0, & \text{else.} \end{cases}$$

Control Signal

From Simulator

1. Steer signal
2. Throttle signal
3. Brake signal
4. Reverse gear

Discrete Action Space

1. Drive Forward
2. Turn to Left
3. Turn to Right
4. Drive Backward

Bayesian Approach to RL

Metrics

- Being offroad
- Being in the meeting lane
- Being either offroad or in the meeting lane
- Accomplished tasks
- Tasks without collisions
- Total distance driven

Experiments

- TGDG: Training and deployment using GT
- TEDE: Training and deployment using Estimation
- TGDE: Training on GT but deployment using Estimation
- TEDG: Training on Estimation but deployment using GT

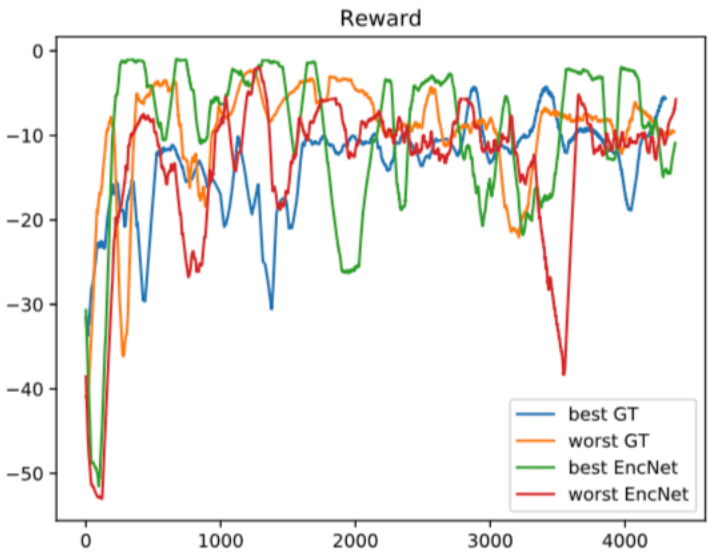


Fig.1

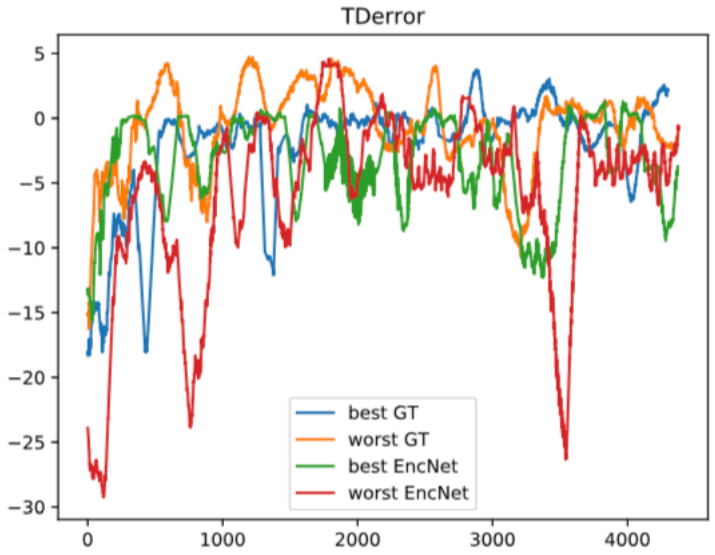


Fig.2

Bayesian Approach to RL

Results

Model		Offroad	Otherlane	Either	Success	No collision	Score	Dist[m]
TGDG	average	0.0%	11.1%	11.1%	96.8%	84.1%	0.90 (±0.07)	11132 (±703)
	best model	0.0%	2.4%	2.4%	100.0%	100.0%	0.99	11168
TGDE	average	0.1%	11.6%	11.7%	89.0%	80.8%	0.86 (±0.073)	11392 (±232)
	best model	0.0%	2.5%	2.5%	100.0%	100.0%	0.99	11119
TEDE	average	11.4%	29.8%	37.9%	34.5%	73.4%	0.57 (±0.3)	10209 (±1950)
	best model	0.0%	17.5%	17.5%	100.0%	100.0%	0.94	11227
TEDG	average	11.5%	25.8%	36.3%	33.3%	78.3%	0.58 (±0.3)	11437 (±2201)
	best model	0.0%	22.7%	22.7%	100.0%	100.0%	0.92	11238
RL		42.4%	21.0%	52.4%	41.7%	50.0%	0.46	8787
IL		8.3%	0.7%	8.4%	86.9%	90.5%	0.90	11293

tab.1

	New Town							New Town & new weather						
Infraction type	MP	IL	RL	TGDG	TEDE	TEDG	TGDE	MP	IL	RL	TGDG	TEDE	TEDG	TGDE
Opposite lane	0.45	1.12	0.23	2.14	0.18	0.24	**4.10**	0.40	0.78	0.21	2.13	0.18	0.25	**2.93**
Sidewalk	0.46	0.76	0.43	0.40	**10.24**	9.80	0.11	0.43	0.81	0.48	0.39	6.64	**9.80**	0.10
Collision-static	0.44	0.40	0.23	**2.52**	2.16	2.18	0.55	0.45	0.28	0.25	**3.55**	1.53	3.27	0.79
Collision-vehicle	0.51	**0.59**	0.41	0.34	0.22	0.27	0.38	**0.47**	0.44	0.37	0.33	0.26	0.24	0.34
Collision-pedestrian	1.40	1.88	**2.55**	1.4	1.20	1.35	0.53	1.46	1.41	**2.99**	1.39	0.69	1.13	1.31

tab.2

Conclusion and Future Plans

- Continuous action space
- Moving from virtual to the real-world
- Design and implementation of realistic road netwoks

References

- H. Firouzi, M. N. Ahmadabadi, B. N. Araabi, S. Amizadeh, M. S. Mirian, and R. Siegwart, “Interactive learning in continuous multimodal space: a bayesian approach to action-based soft partitioning and learn- ing,” IEEE Transactions on Autonomous Mental Development, vol. 4, no. 2, pp. 124–138, 2012.
- Z.Gharaee, A.Fatehi,M.S.Mirian,andM.N.Ahmadabadi,“Attention control learning in the decision space using state estimation,” International Journal of Systems Science, vol. 47, no. 7, pp. 1659–1674, 2016.
- A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, “CARLA: An open urban driving simulator,” in Proceedings of the 1st Annual Conference on Robot Learning, 2017, pp. 1–16.

Thank you for your attention!