

BestMasters

Andreas Folkers

Steuerung eines autonomen Fahrzeugs durch Deep Reinforcement Learning

EXTRAS ONLINE



Springer Spektrum

BestMasters

Mit „**BestMasters**“ zeichnet Springer die besten Masterarbeiten aus, die an renommierten Hochschulen in Deutschland, Österreich und der Schweiz entstanden sind. Die mit Höchstnote ausgezeichneten Arbeiten wurden durch Gutachter zur Veröffentlichung empfohlen und behandeln aktuelle Themen aus unterschiedlichen Fachgebieten der Naturwissenschaften, Psychologie, Technik und Wirtschaftswissenschaften. Die Reihe wendet sich an Praktiker und Wissenschaftler gleichermaßen und soll insbesondere auch Nachwuchswissenschaftlern Orientierung geben.

Springer awards “**BestMasters**” to the best master’s theses which have been completed at renowned Universities in Germany, Austria, and Switzerland. The studies received highest marks and were recommended for publication by supervisors. They address current issues from various fields of research in natural sciences, psychology, technology, and economics. The series addresses practitioners as well as scientists and, in particular, offers guidance for early stage researchers.

Weitere Bände in der Reihe <http://www.springer.com/series/13198>

Andreas Folkers

Steuerung eines autonomen Fahrzeugs durch Deep Reinforcement Learning

 Springer Spektrum

Andreas Folkers
Bremen, Deutschland

1. Gutachter: Dr. Matthias Knauer, Universität Bremen
2. Gutachter: Prof. Dr. Christof Büskens, Universität Bremen

ISSN 2625-3577

ISSN 2625-3615 (electronic)

BestMasters

ISBN 978-3-658-28885-3

ISBN 978-3-658-28886-0 (eBook)

<https://doi.org/10.1007/978-3-658-28886-0>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2019

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Springer Spektrum ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Inhaltsverzeichnis

Abbildungsverzeichnis	vii
Tabellenverzeichnis	ix
Kurzzusammenfassung	xi
1 Einführung	1
2 Grundlagen des Deep Learning	3
2.1 Künstliche Neuronale Netze	3
2.1.1 Definition und Approximationseigenschaft Neuronaler Netze	4
2.1.2 Faltungsbasierte Netze	6
2.1.3 Lokale Translationsinvarianz	8
2.2 Training Neuronaler Netze	8
2.2.1 Stochastische Parameteroptimierung	9
2.2.2 Backpropagation	11
3 Deep Reinforcement Learning	15
3.1 Charakteristiken des Reinforcement Learning	16
3.1.1 Annahmen an die Umwelt	17
3.1.2 Beschreibung des Agenten	18
3.2 Lernverfahren	21
3.2.1 Actor-Critic-Methoden	21
3.2.2 Verallgemeinerte Approximation der Advantage	22
3.2.3 Monotone Policy Verbesserung	24
3.2.4 Proximal-Policy-Optimierung	28

4	Deep Controller für autonomes Fahren	33
4.1	Problemstellung	33
4.2	Regelungsverfahren	35
4.2.1	Modellbasierte Ansätze	35
4.2.2	Deep Controller	36
4.3	Umsetzung	37
4.3.1	Die Fahrzeugbewegung als Einspurmodell	37
4.3.2	Simulierte Regelung auf einem Parkplatz	39
4.3.3	Definition des Agenten	41
4.3.4	Definition der Belohnungsfunktion	43
5	Training und Auswertung	47
5.1	Auswertung des Trainings	47
5.1.1	Verlauf des Trainings für verschiedene Probleme	48
5.1.2	Variation von Hyperparametern des Lernverfahrens	50
5.2	Evaluierung in der Simulation	51
5.2.1	Numerische Stabilitätsanalyse der Koordinaten	53
5.2.2	Numerische Stabilitätsanalyse der Wahrnehmungskarte	55
5.2.3	Aufmerksamkeit in der Wahrnehmungskarte	57
5.2.4	Leistungsfähigkeit von trainierten Agenten	58
5.3	Evaluierung mit einem Testfahrzeug	61
5.3.1	Der Deep Controller im Projekt AO-Car	62
5.3.2	Der Versuchsparkplatz	63
5.3.3	Analyse der Umsetzung	64
6	Resümee	71
6.1	Zusammenfassung	71
6.2	Diskussion	73
	Literaturverzeichnis	77
A	Verwendete Hyperparameter	87

Abbildungsverzeichnis

2.1	Neuronales Netz mit vier Schichten	4
2.2	Visualisierung der diskreten Faltung	7
2.3	Beispiel für das Max-Pooling	9
2.4	Gradientenberechnung durch Backpropagation	13
3.1	Wechselwirkung der drei Hauptelemente des Reinforcement Learning	16
3.2	Inverses Pendel als Beispiel eines Reinforcement Learning Problems	17
3.3	Actor-Critic-Lernarchitektur	22
3.4	Maximierungsterm der Proximal-Policy-Optimierung, abhängig von der Advantage-Funktion	29
3.5	Generierung von Trajektorien für die Proximal-Policy-Optimierung	30
4.1	Sensorik des Forschungsfahrzeugs im Projekt AO-Car	33
4.2	Beispielsituation während der Parkplatzexploration	34
4.3	Vereinfachung des Einspurmodells und Koordinaten des resultierenden Fahrzeugzustandes	38
4.4	Zufällige Beispiele für simulierte Regelungsaufgaben	40
4.5	Weltwahrnehmung und Zielkoordinaten als diskrete Wahrnehmungskarte	41
4.6	Topologie der verwendeten Neuronalen Netze	43
5.1	Beispielsituation zur Darstellung des normalen und vereinfachten Trainings	48
5.2	Exemplarischer Trainingsverlauf des reinen Gewinns für FAHRER- und STOPPER-Agenten	49
5.3	Entwicklung des gemittelten reinen Gewinns während des Trainings für unterschiedliche Parametrisierungen des Lernverfahrens	50
5.4	Lösung einer simulierten Regelungsaufgabe durch einen fertig trainierten Agenten	52
5.5	Entwicklung von Geschwindigkeit, Orientierung, Lenkwinkel, Beschleunigung und Lenkwinkeländerung während einer Regelaufgabe	53
5.6	Einfluss von Änderungen der Fahrzeug- und Zielgeschwindigkeit sowie des Lenkwinkels auf die Steuerungen	54

5.7	Einfluss von Änderungen der Fahrzeug- und Zielorientierung auf die Steuerungen	55
5.8	Numerische Stabilitätsanalyse bei einer Punktstörung in der Wahrnehmungskarte	56
5.9	Numerische Stabilitätsanalyse bei einer Störung durch Linien in der Wahrnehmungskarte	57
5.10	Beispiele für Aufmerksamkeitskarten zu verschiedenen Situationen	59
5.11	Aufmerksamkeitskarten eines Agenten aus dem vereinfachten Training	61
5.12	Vereinfachte Darstellung der Kommunikation von Modulen im AO-Car Projekt	62
5.13	Luftaufnahme und resultierende Fahrbahnbegrenzungen des zu explorierenden Parkplatzes	63
5.14	Wahrnehmungskarten des Realfahrzeugs bei der Exploration des Testparkplatzes	64
5.15	Beispiel für den Verlauf der geschätzten Fahrzeugpositionen bei der Parkplatzeexploration mit dem Deep Controller	65
5.16	Steuerungen sowie geschätzte Geschwindigkeiten, Lenkwinkel und Orientierungen bei der Regelung mit dem Deep Controller	66
5.17	Ergebnisse bei der Anwendung des Deep Controller für drei Beispielsituationen	69
6.1	Beispiel für einen gekrümmten Parkplatz	73
6.2	Reduktion der Weltwahrnehmung auf relevante Informationen durch einen Polygonzug	74

Tabellenverzeichnis

4.1	Eigenschaften von Regelungskonzepten	36
4.2	Detaillierte Übersicht aller Schichten der verwendeten Neuronalen Netze . .	42
4.3	Qualitative Auflistung von Komponenten der Belohnungsfunktion	44
5.1	Vergleich von Agenten aus verschiedenen Trainingsvarianten	60