cases	doc_1		doc_2		decision	id
	authors	• Sedlmeier, A. • Gabor, T. • Phan, T. • Belzner, L.	**authors**	• Andreas Sedlmeier • Thomas Gabor • Thomy Phan • Lenz Belzner • Claudia Linnhoff-Popien	DUPLICATES	150
	title	Uncertainty-Based Out-of-Distribution Detection in Deep Reinforcement Learning	**title**	Uncertainty-Based Out-of-Distribution Detection in Deep Reinforcement Learning		
	publication_date	2019-12-04 00:00:00	**publication_date**	2019-01-08 00:00:00		
	source	SupportedSources.CROSSREF	**source**	SupportedSources.INTERNET_ARCHIVE		
	journal		**journal**			
	volume		**volume**			
	doi	10.1007/s42354-019-0238-z	**doi**			
	urls	• http://link.springer.com/content/pdf/10.1007/s42354-019-0238-z.pdf • http://link.springer.com/article/10.1007/s42354-019-0238-z/fulltext.html • http://link.springer.com/content/pdf/10.1007/s42354-019-0238-z.pdf • http://dx.doi.org/10.1007/s42354-019-0238-z	**urls**	• https://web.archive.org/web/20191025111127/https://arxiv.org/pdf/1901.02219v1.pdf		
	id	id5621195001120869475	**id**	id-185732030543801788		
	abstract		**abstract**	We consider the problem of detecting out-of-distribution (OOD) samples in deep reinforcement learning. In a value based reinforcement learning setting, we propose to use uncertainty estimation techniques directly on the agent's value estimating neural network to detect OOD samples. The focus of our work lies in analyzing the suitability of approximate Bayesian inference methods and related ensembling techniques that generate uncertainty estimates. Although prior work has shown that dropout-based variational inference techniques and bootstrap-based approaches can be used to model epistemic uncertainty, the suitability for detecting OOD samples in deep reinforcement learning remains an open question. Our results show that uncertainty estimation can be used to differentiate in- from out-of-distribution samples. Over the complete training process of the reinforcement learning agents, bootstrap-based approaches tend to produce more reliable epistemic uncertainty estimates, when compared to dropout-based approaches.		
	versions		**versions**			