

cases	doc_1		doc_2		decision	id
	authors	<ul style="list-style-type: none"><li>• Murat Sensoy</li><li>• Lance Kaplan</li><li>• Federico Cerutti</li><li>• Maryam Saleki</li></ul>	authors	<ul style="list-style-type: none"><li>• Cerutti, Federico</li><li>• Kaplan, Lance</li><li>• Saleki, Maryam</li><li>• Sensoy, Murat</li></ul>	DUPLICATES	140
	title	Uncertainty-Aware Deep Classifiers using Generative Models	title	Uncertainty-Aware Deep Classifiers using Generative Models		
	publication_date	2020-06-07 00:00:00	publication_date	2020-01-01 00:00:00		
	source	SupportedSources.INTERNET_ARCHIVE	source	SupportedSources.CORE		
	journal		journal			
	volume		volume			
	doi		doi	10.1609/aaai.v34i04.6015		
	urls	<ul style="list-style-type: none"><li>• <a href="https://web.archive.org/web/20200610070344/https://arxiv.org/pdf/2006.04183v1.pdf">https://web.archive.org/web/20200610070344/https://arxiv.org/pdf/2006.04183v1.pdf</a></li></ul>	urls	<ul style="list-style-type: none"><li>• <a href="http://arxiv.org/abs/2006.04183">http://arxiv.org/abs/2006.04183</a></li></ul>		
	id	id-3208312883179761422	id	id1293373950842840618		
	abstract	Deep neural networks are often ignorant about what they do not know and overconfident when they make uninformed predictions. Some recent approaches quantify classification uncertainty directly by training the model to output high uncertainty for the data samples close to class boundaries or from the outside of the training distribution. These approaches use an auxiliary data set during training to represent out-of-distribution samples. However, selection or creation of such an auxiliary data set is non-trivial, especially for high dimensional data such as images. In this work we develop a novel neural network model that is able to express both aleatoric and epistemic uncertainty to distinguish decision boundary and out-of-distribution regions of the feature space. To this end, variational autoencoders and generative adversarial networks are incorporated to automatically generate out-of-distribution exemplars for training. Through extensive analysis, we demonstrate that the proposed approach provides better estimates of uncertainty for in- and out-of-distribution samples, and adversarial examples on well-known data sets against state-of-the-art approaches including recent Bayesian approaches for neural networks and anomaly detection methods.	abstract	Deep neural networks are often ignorant about what they do not know and overconfident when they make uninformed predictions. Some recent approaches quantify classification uncertainty directly by training the model to output high uncertainty for the data samples close to class boundaries or from the outside of the training distribution. These approaches use an auxiliary data set during training to represent out-of-distribution samples. However, selection or creation of such an auxiliary data set is non-trivial, especially for high dimensional data such as images. In this work we develop a novel neural network model that is able to express both aleatoric and epistemic uncertainty to distinguish decision boundary and out-of-distribution regions of the feature space. To this end, variational autoencoders and generative adversarial networks are incorporated to automatically generate out-of-distribution exemplars for training. Through extensive analysis, we demonstrate that the proposed approach provides better estimates of uncertainty for in- and out-of-distribution samples, and adversarial examples on well-known data sets against state-of-the-art approaches including recent Bayesian approaches for neural networks and anomaly detection methods.Comment: This is a post-referred version of a conference paper published in AAAI 202		
	versions		versions			