

cases	doc_1		doc_2		decision	id
	authors	<ul style="list-style-type: none"><li>Ahmed Sabir</li><li>Francesc Moreno-Noguer</li><li>Llu��s Padr��3</li></ul>	authors	<ul style="list-style-type: none"><li>Moreno-Noguer, Francesc</li><li>Padr��3, Llu��s</li><li>Sabir, Ahmed</li></ul>	DUPLICATES	102
	title	Visual Semantic Relatedness Dataset for Image Captioning	title	Visual Semantic Relatedness Dataset for Image Captioning		
	publication_date	2023-01-20 20:04:35+00:00	publication_date	2023-01-20 00:00:00		
	source	SupportedSources.ARXIV	source	SupportedSources.CORE		
	journal	None	journal			
	volume		volume			
	doi		doi	None		
	urls	<ul style="list-style-type: none"><li>http://arxiv.org/pdf/2301.08784v1</li><li>http://arxiv.org/abs/2301.08784v1</li><li>http://arxiv.org/pdf/2301.08784v1</li></ul>	urls	<ul style="list-style-type: none"><li>http://arxiv.org/abs/2301.08784</li></ul>		
	id	id5420861525919663414	id	id257141867278510667		
	abstract	Modern image captioning system relies heavily on extracting knowledge from images to capture the concept of a static story. In this paper, we propose a textual visual context dataset for captioning, in which the publicly available dataset COCO Captions (Lin et al., 2014) has been extended with information about the scene (such as objects in the image). Since this information has a textual form, it can be used to leverage any NLP task, such as text similarity or semantic relation methods, into captioning systems, either as an end-to-end training strategy or a post-processing based approach.	abstract	Modern image captioning system relies heavily on extracting knowledge from images to capture the concept of a static story. In this paper, we propose a textual visual context dataset for captioning, in which the publicly available dataset COCO Captions (Lin et al., 2014) has been extended with information about the scene (such as objects in the image). Since this information has a textual form, it can be used to leverage any NLP task, such as text similarity or semantic relation methods, into captioning systems, either as an end-to-end training strategy or a post-processing based approach.Comment: Project Page: bit.ly/3Zq6AT		
	versions		versions			