cases	doc_1		doc_2		decision	id
			authors	• Yunchao Gong • Qifa Ke • Michael Isard • Svetlana Lazebnik	DUPLICATES	377
	authors	• Yunchao Gong • Qifa Ke • M. Isard • S. Lazebnik	**title**	A Multi-View Embedding Space for Modeling Internet Images, Tags, and their Semantics		
	title	A Multi-View Embedding Space for Modeling Internet Images, Tags, and Their Semantics	**publication_date**	2012-12-18 22:02:43+00:00		
	publication_date	2012-12-18 00:00:00	**source**	SupportedSources.ARXIV		
	source	SupportedSources.SEMANTIC_SCHOLAR	**journal**	None		
	journal	International Journal of Computer Vision	**volume**			
	volume	106	**doi**			
	doi	10.1007/s11263-013-0658-4	**urls**	• http://arxiv.org/pdf/1212.4522v2 • http://arxiv.org/abs/1212.4522v2 • http://arxiv.org/pdf/1212.4522v2		
	urls	• https://www.semanticscholar.org/paper/7fceccc7a1046caa4936b14eeacb71ccf4d6be10	**id**	id-8979179367337538343		
	id	id-4794057921900810675	**abstract**	This paper investigates the problem of modeling Internet images and associated text or tags for tasks such as image-to-image search, tag-to-image search, and image-to-tag search (image annotation). We start with canonical correlation analysis (CCA), a popular and successful approach for mapping visual and textual features to the same latent space, and incorporate a third view capturing high-level image semantics, represented either by a single category or multiple non-mutually-exclusive concepts. We present two ways to train the three-view embedding: supervised, with the third view coming from ground-truth labels or search keywords; and unsupervised, with semantic themes automatically obtained by clustering the tags. To ensure high accuracy for retrieval tasks while keeping the learning process scalable, we combine multiple strong visual features and use explicit nonlinear kernel mappings to efficiently approximate kernel CCA. To perform retrieval, we use a specially designed similarity function in the embedded space, which substantially outperforms the Euclidean distance. The resulting system produces compelling qualitative results and outperforms a number of two-view baselines on retrieval tasks on three large-scale Internet image datasets.		
	abstract	None	**versions**			
	versions					