

# 1 Totally used references

Row	Article	Author	Year	Journal	Page Count
1	Deep Visual-Semantic Alignments for Generating Image Descriptions	Li Fei-Fei	2015	CVPR	13
2	Deep Fragment Embeddings for Bidirectional Image Sentence Mapping	Li Fei-Fei	2014	Advances in neural information processing systems	9
3	What do we perceive in a glance of a real-world scene?	Li Fei-Fei	2007	Journal of Vision	29
4	A Sentence is Worth a Thousand Pixels	Urtasan	2013	CVPR	8
5	What, where and who? classifying events by scene and object recognition.	Li Fei-Fei	2007	ICCV	8
6	Towards total scene understanding: Classification, annotation and segmentation in an automatic framework	Li Fei-Fei	2009	CVPR	8
7	Generating Text with Recurrent Neural Networks	Hinton	2011	ICML (International Conference on Machine Learning)	8
8	Decomposing a scene into Geometric and Semantically Consistent Regions	Koller	2009	Computer Vision	8
9	Every Picture Tells a Story: Generating Sentences from Images	Farhadi, Forsyth	2010	Computer Vision (ECCV)	14
10	Show, Attend and Tell: Neural Image Caption Generation with Visual Attention	Bengio	2015	arXive	10
11	From Captions to Visual Concepts and Back	Zitnick	2015	CVPR	10

## 2 Partially used references

Row	Article	Author	Year	Journal	Page Count
1	Using multiple segmentations to discover objects and their extent in image collections	Efros, Freeman et al.	2006	CVPR	5
2	A bayesian hierarchical model for learning natural scene categories.	Li Fei-Fei	2005	CVPR	8
3	Optimol: automatic online picture collection via incremental model learning.	Li Fei-Fei	2009	CVPR	22
4	Deep Captioning with Multimodal Recurrent Neural Networks (m-RNN)	Bengio	2014	arXive	17
5	Describing Objects by Their Attributes	Farhadi, Forsyth	2009	CVPR	8
6	Scene Discovery by Matrix Factorization	Farhadi	2008	Computer Vision (ECCV)	14
7	Collective Generation of Natural Image Description	-	2012	Annual Meeting of the Association for Computational Linguistics	10
8	Framing Image Description as a Ranking Task: Data, Models and Evaluation Metrics	Hodoosh, Hockenmaier	2013	JAIR	20
9	From Image Annotation to Image Description	-	2012	Neural Information Processing	8
10	Generating Text with Recurrent Neural Networks	Bengio	2011	ICML	8
11	A Fuzzy Ontology and its Application to News Summarization	-	2005	Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions	22
12	A Method for Automatic Evaluation of Machine Translation (BLEU)	Microsoft	2002	Annual meeting on association for computational linguistics	10