

# Recopilación de referencias

September 3, 2020

## 1 Papers individuales

### 1.1 Show, Attend & Tell (Xu et al., 2015)

### 1.2 Automatic Generation of Medical Imaging Reports (Jing, Xie, and Xing, 2018)

- Image captioning with deep learning: (Vinyals et al., 2014)(Fang et al., 2014)(Karpathy and Fei-Fei, 2014)(Xu et al., 2015)(You et al., 2016)(Krause et al., 2016)(Liang et al., 2017)

### 1.3 Feature Difference Makes Sense (Park et al., 2020)

- Image captioning with deep learning: (Vinyals et al., 2014; Xu et al., 2015; Karpathy and Fei-Fei, 2014; You et al., 2016)(Zhou et al., 2017)(Anderson et al., 2017)

## References

- Anderson, Peter, Xiaodong He, Chris Buehler, Damien Teney, Mark Johnson, Stephen Gould, and Lei Zhang (2017). *Bottom-Up and Top-Down Attention for Image Captioning and Visual Question Answering*. arXiv: 1707.07998 [cs.CV].
- Fang, Hao, Saurabh Gupta, Forrest Iandola, Rupesh Srivastava, Li Deng, Piotr Dollár, Jianfeng Gao, Xiaodong He, Margaret Mitchell, John C. Platt, C. Lawrence Zitnick, and Geoffrey Zweig (2014). *From Captions to Visual Concepts and Back*. arXiv: 1411.4952 [cs.CV].
- Jing, Baoyu, Pengtao Xie, and Eric Xing (July 2018). “On the Automatic Generation of Medical Imaging Reports”. In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Melbourne, Australia: Association for Computational Linguistics, pp. 2577–2586. DOI: 10.18653/v1/P18-1240. URL: <https://www.aclweb.org/anthology/P18-1240>.
- Karpathy, Andrej and Li Fei-Fei (2014). *Deep Visual-Semantic Alignments for Generating Image Descriptions*. arXiv: 1412.2306 [cs.CV].
- Krause, Jonathan, Justin Johnson, Ranjay Krishna, and Li Fei-Fei (2016). *A Hierarchical Approach for Generating Descriptive Image Paragraphs*. arXiv: 1611.06607 [cs.CV].
- Liang, Xiaodan, Zhiting Hu, Hao Zhang, Chuang Gan, and Eric P. Xing (2017). *Recurrent Topic-Transition GAN for Visual Paragraph Generation*. arXiv: 1703.07022 [cs.CV].
- Park, Hyeryun, Kyungmo Kim, Jooyoung Yoon, Seongkeun Park, and Jinwook Choi (July 2020). “Feature Difference Makes Sense: A medical image captioning model exploiting feature difference and tag information”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*. Online: Association for Computational Linguistics, pp. 95–102. DOI: 10.18653/v1/2020.acl-srw.14. URL: <https://www.aclweb.org/anthology/2020.acl-srw.14>.
- Vinyals, Oriol, Alexander Toshev, Samy Bengio, and Dumitru Erhan (2014). *Show and Tell: A Neural Image Caption Generator*. arXiv: 1411.4555 [cs.CV].
- Xu, Kelvin, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhutdinov, Richard Zemel, and Yoshua Bengio (2015). *Show, Attend and Tell: Neural Image Caption Generation with Visual Attention*. arXiv: 1502.03044 [cs.LG].
- You, Quanzeng, Hailin Jin, Zhaowen Wang, Chen Fang, and Jiebo Luo (June 2016). “Image Captioning With Semantic Attention”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Zhou, Luowei, Chenliang Xu, Parker Koch, and Jason J. Corso (2017). “Watch What You Just Said: Image Captioning with Text-Conditional Attention”. In: *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*. Thematic Workshops '17. Mountain View, California, USA: Association for Computing Machinery, pp. 305–313. ISBN: 9781450354165. DOI: 10.1145/3126686.3126717. URL: <https://doi.org/10.1145/3126686.3126717>.