



Department of Computer Engineering & Information Technology

Amirkabir University of Technology

Research & Technical Presentation in Engineering

Visual Question Answering

Assignment 7

References Engineering

Provided by

Ali Gholami

Student Number

9531504

Advisor

Dr. Reza Safabakhsh

Computer Vision and Image Understanding Style

- Das, A., Agrawal, H., Zitnick, L., Parikh, D., & Batra, D. (2017). Human Attention in Visual Question Answering: Do Humans and Deep Networks Look at the Same Regions? *Computer Vision and Image Understanding*. doi:10.1016/j.cviu.2017.10.001
- Das, A., Kottur, S., Moura, J. M. F., Lee, S., & Batra, D. (2017). Learning Cooperative Visual Dialog Agents with Deep Reinforcement Learning. *Proceedings of the IEEE International Conference on Computer Vision, 2017–Octob*, 2970–2979. doi:10.1109/ICCV.2017.321
- Kafle, K., & Kanan, C. (2017). Visual question answering: Datasets, algorithms, and future challenges. *Computer Vision and Image Understanding*, 163, 3–20. doi:10.1016/j.cviu.2017.06.005
- Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. doi:10.1109/TPAMI.2016.2577031
- Saito, K., Shin, A., Ushiku, Y., & Harada, T. (2017). DualNet: Domain-invariant network for visual question answering. *Proceedings - IEEE International Conference on Multimedia and Expo*, 829–834. doi:10.1109/ICME.2017.8019436
- Wang, P., Wu, Q., Shen, C., Dick, A., & Hengel, A. van den. (2017). FVQA: Fact-based Visual Question Answering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–16. doi:10.1109/TPAMI.2017.2754246
- Wu, Q., Teney, D., Wang, P., Shen, C., Dick, A., & van den Hengel, A. (2017). Visual question answering: A survey of methods and datasets. *Computer Vision and Image Understanding*, 163, 21–40. doi:10.1016/j.cviu.2017.05.001
- Yu, D., Fu, J., Mei, T., & Rui, Y. (2017). Multi-level Attention Networks for Visual Question Answering. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4187–4195. doi:10.1109/CVPR.2017.446
- Yu, Z., Yu, J., Fan, J., & Tao, D. (2017). Multi-modal Factorized Bilinear Pooling with Co-attention Learning for Visual Question Answering. *Proceedings of the IEEE International Conference on Computer Vision, 2017–Octob*, 1839–1848. doi:10.1109/ICCV.2017.202
- Zhu, Y., Lim, J. J., & Fei-Fei, L. (2017). Knowledge Acquisition for Visual Question Answering via Iterative Querying. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6146–6155. doi:10.1109/CVPR.2017.651
- (Das, Agrawal, et al. 2017; Das, Kottur, et al. 2017; Kafle and Kanan 2017; Ren et al. 2017; Saito et al. 2017; Wang et al. 2017; Wu et al. 2017; D. Yu et al. 2017; Z. Yu et al. 2017; Zhu et al. 2017)