Daftar Pustaka

[1] P. Anderson, X. He, C. Buehler, D. Teney, M. Johnson, S. Gould, and L. Zhang. Bottom-up and top-down attention for image captioning and VQA. CoRR, abs/1707.07998, 2017.

[2] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. pages 1–9, 2014.

[3] B. Dai, S. Fidler, and D. Lin. A Neural Compositional Paradigm for Image Captioning. (NeurIPS):1–11, 2018.

[4] B. Dai, S. Fidler, R. Urtasun, and D. Lin. Towards Diverse and Natural Image Descriptions via a Conditional GAN. Proceedings of the IEEE International Conference on Computer Vision, 2017-Octob:2989–2998, 2017.

[5] B. Dai and D. Lin. Contrastive Learning for Image Captioning. (CI), 2017.

[6] A. Das, Subrata; Jain, Lalit; Das. Deep Learning for Military Image Captioning. 2018 21st International Conference on Information Fusion (FUSION), pages 2165–2171, 2018.

[7] A. Deshpande, J. Aneja, L. Wang, A. Schwing, and D. A. Forsyth. Fast, Diverse and Accurate Image Captioning Guided By Part-of-Speech. 2018.

[8] F. Fang, Q. Li, H. Wang, P. Tang, and C. Science. Image Captioning with Word Level Attention. 2018 25th IEEE International Conference on Image Processing (ICIP), pages 1278–1282, 2018.

[9] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem:770–778, 2016.

[10] S. Hochreiter. The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem So-lutions. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 06(02):107–116, 2003.

[11] S. Hochreiter and J. Urgen Schmidhuber. Ltsm. Neural Computation, 9(8):1735–1780, 1997.

[12] R. Kiros, R. Salakhutdinov, and R. Zemel. Multimodal neural language models. 31st International Confe-rence on Machine Learning, ICML 2014, 3:2012–2025, 2014.

[13] T.-y. Lin, C. L. Zitnick, and P. Doll. Microsoft COCO : Common Objects in Context. pages 1–15, 2014.

[14] J. Lu, C. Xiong, D. Parikh, and R. Socher. Knowing when to look: Adaptive attention via a visual sentinel for image captioning. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua:3242–3250, 2017.

[15] P. Mathur, A. Gill, A. Yadav, A. Mishra, and N. K. Bansode. Camera2Caption: A real-Time image cap-tion generator. ICCIDS 2017 - International Conference on Computational Intelligence in Data Science, Proceedings, 2018-Janua(2015):1–6, 2018.

[16] A. A. Nugraha, A. Arifianto, and Suyanto. Generating image description on Indonesian language using convolutional neural network and gated recurrent unit. 2019 7th International Conference on Information and Communication Technology, ICoICT 2019, 2019.

[17] K. Papineni, S. Roukos, T. Ward, and W.-j. Zhu. B LEU : a Method for Automatic Evaluation of Machine Translation. (July):311–318, 2002.

[18] B. A. Plummer, L. Wang, C. M. Cervantes, J. C. Caicedo, J. Hockenmaier, and S. Lazebnik. Flickr30k entities: Collecting region-to-phrase correspondences for richer image-to-sentence models, 2015.

[19] B. Shi, L. Ji, Y. Liang, N. Duan, P. Chen, Z. Niu, and M. Zhou. Dense Procedure Captioning in Narrated Instructional Videos. pages 6382–6391, 2019.

[20] H. Shi, B. Wang, and Z. Wang. Image Captioning based on Deep Reinforcement Learning. Icimcs, pages 1–5, 2018.

[21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the Inception Architecture for Computer Vision. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-December:2818–2826, 2016.

[22] V. Tech, C. L. Zitnick, and D. Parikh. CIDEr : Consensus-based Image Description Evaluation Ramakrishna Vedantam. 2014.

[23] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan. Show and tell: A neural image caption generator. Pro-ceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June:3156–3164, 2015.

[24] Q. Wu, C. Shen, P. Wang, A. Dick, and A. Van Den Hengel. Image Captioning and Visual Question Ans-wering Based on Attributes and External Knowledge. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(6):1367–1381, 2018.

[25] K. Xu, J. L. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhutdinov, R. S. Zemel, and Y. Bengio. Show, attend and tell: Neural image caption generation with visual attention. 32nd International Conference on Machine Learning, ICML 2015, 3:2048–2057, 2015.

[26] Z. Xu, J. Hu, and W. Deng. Recurrent convolutional neural network for video classification. Proceedings - IEEE International Conference on Multimedia and Expo, 2016-Augus:756–765, 2016.