

DAFTAR PUSTAKA

- Berg, T.L., Berg, A.C. and Shih, J., 2010, September. Automatic attribute discovery and characterization from noisy web data. *In European Conference on Computer Vision* (pp. 663-676). Springer, Berlin, Heidelberg.
- Britz, D., 2016, Recurrent Neural Networks Tutorial, Part 1-Introduction to RNNs, <http://www.wildml.com/2015/09/recurrent-neural-networks-tutorial-part-1-introduction-to-rnns/>, diakses tanggal 04 April 2017.
- Chen, G., Ye, D., Xing, Z., Chen, J. and Cambria, E., 2017, May. Ensemble application of convolutional and recurrent neural networks for multi-label text categorization. *In Neural Networks (IJCNN), 2017 International Joint Conference on* (pp. 2377-2383). IEEE.
- Chen, X. and Lawrence Zitnick, C., 2015. Mind's eye: A recurrent visual representation for image caption generation. *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2422-2431).
- Chowdhury, G.G., 2003. Natural language processing. *Annual review of information science and technology*, 37(1), pp.51-89.
- Cho, K., B. van Merriënboer, D. Bahdanau, and Y. Bengio, 2014. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint arXiv:1409.1259*.
- Deng, L. dan Yu, D., 2014, Deep Learning : Methods and Applications, *Foundations and Trends in Signal Processing*, 7(3-4), pp.197 - 387.
- Deng, L. dan Yu, D., 2014, Deep Learning : Methods and Applications, *Foundations and Trends in Signal Processing*, 7(3-4), pp.197 - 387.
- Elliott, D. and Keller, F., 2013. Image description using visual dependency representations. *In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (pp. 1292-1302).
- Faucett, L., 1994. Fundamentals of neural networks. Architectures, Algorithms, and Applications, Prentice Hall-Inc., Canada.

- Felzenszwalb, P.F., Girshick, R.B., McAllester, D. and Ramanan, D., 2010. Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*, 32(9), pp.1627-1645.
- Godbole, G., 2015. *Natural Language Generation*.
- Goldberg, Y., 2017. Neural network methods for natural language processing. *Synthesis Lectures on Human Language Technologies*, 10(1), pp.1-309.
- Gupta, A. and Mannem, P., 2012, November. From image annotation to image description. In *International Conference on Neural Information Processing* (pp. 196-204). Springer, Berlin, Heidelberg.
- Gupta, A., Verma, Y. and Jawahar, C.V., 2012, July. Choosing Linguistics over Vision to Describe Images. In *AAAI* (p. 1).
- Haykin, S.S., Haykin, S.S., Haykin, S.S. and Haykin, S.S., 2009. *Neural networks and learning machines* (Vol. 3). Upper Saddle River, NJ, USA:: Pearson.
- Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural computation*, 9(8), pp.1735-1780.
- Jaimes, A. and Chang, S.F., 1999, December. Conceptual framework for indexing visual information at multiple levels. In *Internet Imaging* (Vol. 3964, pp. 2-16). International Society for Optics and Photonics.
- Jurafsky, D., 2000. *Speech & language processing*. Pearson Education India.
- Karpathy, A. and Fei-Fei, L., 2015. Deep visual-semantic alignments for generating image descriptions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3128-3137).
- Kiros, R., Salakhutdinov, R. and Zemel, R.S., 2014. Unifying visual-semantic embeddings with multimodal neural language models. *arXiv preprint arXiv:1411.2539*.
- Kiros, R., Salakhutdinov, R. and Zemel, R.S., 2014. Unifying visual-semantic embeddings with multimodal neural language models. *arXiv preprint arXiv:1411.2539*.
- Kulkarni, G., Premraj, V., Ordonez, V., Dhar, S., Li, S., Choi, Y., Berg, A.C. and Berg, T.L., 2013. Babytalk: Understanding and generating simple image descriptions.

- IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12), pp.2891-2903.
- Lampert, C.H., Nickisch, H. and Harmeling, S., 2009, June. Learning to detect unseen object classes by between-class attribute transfer. *In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (pp. 951-958). IEEE.
- Lazebnik, S., Schmid, C. and Ponce, J., 2006. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. *In Computer vision and pattern recognition, 2006 IEEE computer society conference on* (Vol. 2, pp. 2169-2178). IEEE.
- Li, S., Kulkarni, G., Berg, T. L., Berg, A. C., and Choi, Y., 2011. Composing simple image descriptions using webscale n-grams. *In CoNLL*, pages 220–228.
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2), pp.91-110.
- Mao, J., Xu, W., Yang, Y., Wang, J. and Yuille, A.L., 2014. Explain images with multimodal recurrent neural networks. *arXiv preprint arXiv:1410.1090*.
- Mou, L., Yan, R., Li, G., Zhang, L. and Jin, Z., 2015. Backward and Forward Language Modeling for Constrained Sentence Generation. *arXiv preprint arXiv:1512.06612*.
- Mikolov, T., Karafiat, M., Burget, L., Cernocký, J. and Khudanpur, S., 2010. Recurrent neural network based language model. *In Eleventh Annual Conference of the International Speech Communication Association*.
- Mikolov, T., Chen, K., Corrado, G., Dean, J., Sutskever, L. and Zweig, G., 2013. word2vec. URL <https://code.google.com/p/word2vec>.
- Mikolov, T., Yih, W.T. and Zweig, G., 2013. Linguistic regularities in continuous space word representations. *In Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 746-751).
- Oliva, A. and Torralba, A., 2001. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, 42(3), pp.145-175

- Ordonez, V., Kulkarni, G. and Berg, T.L., 2011. Im2text: Describing images using 1 million captioned photographs. *In Advances in neural information processing systems* (pp. 1143-1151).
- Pennington, J., Socher, R. and Manning, C., 2014. Glove: Global vectors for word representation. *In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543).
- Raj, D., 2017, Metrics for NLG Evaluation, <https://medium.com/explorations-in-language-and-learning/metrics-for-nlg-evaluation-c89b6a781054>, diakses tanggal 04 April 2017.
- Rajagede, A. R., 2016, Deep Learning untuk Pengenalan Huruf Hijaiyah Berharakat, *Skripsi*, Departemen Ilmu Komputer dan Matematika, FMIPA, Universitas Gadjah Mada.
- Rashtchian, C., Young, P., Hodosh, M. and Hockenmaier, J., 2010, June. Collecting image annotations using Amazon's Mechanical Turk. *In Proceedings of the NAA-CL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk* (pp. 139-147). Association for Computational Linguistics.
- Reiter, E. and Dale, R., 2006. *Building Natural Language Generation Systems*.
- Socher, R., Bengio, Y. and Manning, C., 2012. *Deep learning for NLP*. Tutorial at ACL.
- Sutskever, I., Martens, J. and Hinton, G.E., 2011. Generating text with recurrent neural networks. *In Proceedings of the 28th International Conference on Machine Learning (ICML-11)* (pp. 1017-1024).
- Syahra, S. G. S., 2018, Implementasi Attentive Recurrent Neural Network dalam Pembuatan Headline dengan Pendekatan Abstraktif, *Skripsi*, Departemen Ilmu Komputer dan Matematika, FMIPA, Universitas Gadjah Mada.
- Yang, Y., Teo, C.L., Daumé III, H. and Aloimonos, Y., 2011, July. Corpus-guided sentence generation of natural images. *In Proceedings of the Conference on Empirical Methods in Natural Language Processing* (pp. 444-454). Association for Computational Linguistics.

Yao, B. and Fei-Fei, L., 2010, June. Grouplet: A structured image representation for recognizing human and object interactions. *In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on* (pp. 9-16). IEEE.