

DAFTAR PUSTAKA

- Alpheccar, A. 2017, *Signal processing with Haskell : Part 2*, <http://www.alpheccar.org/content/103.html>, diakses 5 April 2017.
- Bengio, Y., Simard, P., dan Frasconi, P. 1994, Learning long-term dependencies with gradient descent is difficult, *IEEE transactions on neural networks* 5.2, pp. 157–166.
- Brugnara, F., Falavigna, D., dan Omologo, M. 1993, Automatic segmentation and labeling of speech based on Hidden Markov Models, *Speech Communication* 12.4, pp. 357–370.
- Cho, K., Merriënboer, B. van, Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., dan Bengio, Y. 2014, Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1724–1734.
- Chodroff, E. 2017, *Corpus Phonetics Tutorial*, <http://pages.jh.edu/~echodro1/tutorial/intro.html>, diakses 5 April 2017.
- Chung, J., Gulcehre, C., Cho, K., dan Bengio, Y. 2014, Empirical evaluation of gated recurrent neural networks on sequence modeling, *arXiv preprint arXiv:1412.3555*.
- Davis, S. dan Mermelstein, P. 1980, Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences, *IEEE transactions on acoustics, speech, and signal processing* 28.4, pp. 357–366.
- Duchi, J., Hazan, E., dan Singer, Y. 2011, Adaptive subgradient methods for online learning and stochastic optimization, *Journal of Machine Learning Research* 12.Jul, pp. 2121–2159.
- Elisse, E. 2016, *Voice Control is the Most Significant Tech Story of 2016*, <http://www.triplepundit.com/2016/09/why-voice-control-is-the-most-significant-tech-story-of-2016/>, diakses 15 September 2016.
- Eyben, F., Weninger, F., Squartini, S., dan Schuller, B. 2013, Real-life voice activity detection with LSTM Recurrent Neural Networks and an application to Hollywo-

- od movies, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 1, 3, IEEE, pp. 483–487.
- Fausett, L., 1994, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- Ferroni, G., Bonfigli, R., Principi, E., Squartini, S., dan Piazza, F. 2015, A Deep Neural Network approach for Voice Activity Detection in multi-room domestic scenarios, *2015 International Joint Conference on Neural Networks (IJCNN)*, IEEE, pp. 1–8.
- Glorot, X. dan Bengio, Y. 2010, Understanding the difficulty of training deep feedforward neural networks. *Aistats*, vol. 9, pp. 249–256.
- Goodfellow, I., Bengio, Y., dan Courville, A., 2016, *Deep Learning*, MIT Press.
- Hermansky, H. dan Morgan, N. 1994, *RASTA Processing of Speech*.
- Hinton, G., Srivastava, N., dan Swersky, K. 2012, Neural Networks for Machine Learning Lecture 6a Overview of mini-batch gradient descent,
- Hirsch, H.-G. 2017, *Filtering and Noise Adding Tool*, <http://aurora.hsnr.de/download.html>, diakses 11 Januari 2017.
- Hochreiter, S. dan Schmidhuber, J. 1997, Long short-term memory. *Neural computation* 9.8, pp. 1735–80.
- Hughes, T. dan Mierle, K. 2013, Recurrent neural networks for voice activity detection, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, pp. 7378–7382.
- Jozefowicz, R., Zaremba, W., dan Sutskever, I. 2015, An empirical exploration of recurrent network architectures, *Journal of Machine Learning Research*.
- Kraljevski, I., Tan, Z.-H., dan Bissiri, M. P. 2015, Comparison of forced-alignment speech recognition and humans for generating reference VAD, *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, International Speech and Communication Association, pp. 2937–2941.
- LeCun, Y. A., Bottou, L., Orr, G. B., dan Müller, K.-R. 2012, Efficient backprop, *Neural networks: Tricks of the trade*, Springer, pp. 9–48.

- Leglaive, S., Hennequin, R., dan Badeau, R. 2015, Singing voice detection with deep recurrent neural networks, *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*, IEEE, pp. 121–125.
- Lusted, L. B. 1971, Signal Detectability and Medical Decision-Making, *Science* 171.3977, pp. 1217–1219.
- Mohammed, A.-R. 2014, Deep Neural Network acoustic models for ASR, *Disertasi*, Graduate Department of Computer Science, University of Toronto, Toronto.
- Olah, C. 2016, *Understanding LSTM Networks*, <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>, diakses 15 September 2016.
- Panayotov, V., Chen, G., Povey, D., dan Khudanpur, S. 2015, Librispeech: an ASR corpus based on public domain audio books, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, pp. 5206–5210.
- Povey, D., Ghoshal, A., Boulianne, G., Burget, L., Glembek, O., Goel, N., Hannemann, M., Motlicek, P., Qian, Y., Schwarz, P., et al. 2011, The Kaldi speech recognition toolkit, *IEEE 2011 workshop on automatic speech recognition and understanding*, EPFL-CONF-192584, IEEE Signal Processing Society.
- Ramirez, J., Górriz, J. M., dan Segura, J. C., 2007, *Voice activity detection. fundamentals and speech recognition system robustness*, INTECH Open Access Publisher, New York.
- Snyder, D., Chen, G., dan Povey, D. 2015, *MUSAN: A Music, Speech, and Noise Corpus*, arXiv:1510.08484v1.
- Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., dan Salakhutdinov, R. 2014, Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research* 15.1, pp. 1929–1958.
- Tong, S., Gu, H., dan Yu, K. 2016, A comparative study of robustness of deep learning approaches for VAD, *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5695–5699.
- Young, S. J., Kershaw, D., Odell, J., Ollason, D., Valtchev, V., dan Woodland, P., 2006, *The HTK Book Version 3.4*, Cambridge University Press.

Zhang, Z., Pinto, J., Plahl, C., Schuller, B., dan Willett, D. 2014, Channel mapping using bidirectional long short-term memory for dereverberation in hands-free voice controlled devices, *IEEE Transactions on Consumer Electronics* 60.3, pp. 525–533.