















- Air Traffic Management Research and Development Seminar (ATM2017), Seattle, WA, USA, 2017.
- [7] The project MALORCA (Machine Learning of Speech Recognition Models for Controller Assistance) <http://www.malorca-project.de>, n.d.
- [8] The project HAAWAI (Highly Automated Air Traffic Controller Workstations With Artificial Intelligence Integration) <https://www.haawaii.de>, n.d.
- [9] PJ.10-96-W2: SESAR2020 funded industrial research projects under the European Union's grant agreement 874470, [https://cordis.europa.eu/programme/id/H2020\\_SESAR-IR-VLD-WAVE2-10-2019/de](https://cordis.europa.eu/programme/id/H2020_SESAR-IR-VLD-WAVE2-10-2019/de), n.d.
- [10] PJ.05-97-W2 SESAR2020 funded industrial research projects under the European Union's grant agreement 874464, see for further information [https://www.remote-tower.eu/wp/?page\\_id=888](https://www.remote-tower.eu/wp/?page_id=888), and <https://www.remote-tower.eu/wp/?p=824> and <https://www.sesarju.eu/index.php/projects/DTT>, n.d.
- [11] M. Kleinert, H. Helmke, S. Moos, P. Hlousek, C. Windisch, O. Ohneiser, H. Ehr, and A. Labreuil, "Reducing Controller Workload by Automatic Speech Recognition Assisted Radar Label Maintenance," 9th SESAR Innovation Days, Athens, Greece, 2019.
- [12] Mozilla Corporation, "DeepSpeech Model," 2021, <https://deepspeech.readthedocs.io/en/v0.9.3> (visited on 08/31/2021).
- [13] A. Hannun, C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen, R. Prenger, S. Satheesh, S. Sengupta, A. Coates, and A. Y. Ng, "Deep Speech: Scaling up end-to-end speech recognition," ArXiv 1412.5567, 2014.
- [14] T. Lohrenz, P. Schwarz, Z. Li, and T. Fingscheidt, "Relaxed Attention: A Simple Method to Boost Performance of End-to-End Automatic Speech Recognition," submitted to IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), Cartagena, Colombia, 2021.
- [15] Kaldi Toolkit <http://kaldi-asr.org/doc/about.html>, n.d.
- [16] M. Mohri, F. Pereira, and M. Riley, "Weighted finite-state transducers in speech recognition," Computer Speech & Language, 2002.
- [17] H. Helmke, M. Kleinert, J. Rataj, et al, "Cost Reductions Enabled by Machine Learning in ATM - How can Automatic Speech Recognition enrich human operators performance?," 13th USA/Europe Air Traffic Management Research and Development Seminar, Vienna, Austria, 2019.
- [18] T. Pellegrini, J. Farinas, E. Delpech, and F. Lancelot, "The Airbus Air Traffic Control speech recognition 2018 challenge: Towards ATC automatic transcription and call sign detection," in 20th Annual Conference of the International Speech Communication Association (INTERSPEECH 2019), Austria, Sep 2019.
- [19] J. Zuluaga-Gomez, P. Motlicek, Q. Zhan, et al, "Automatic Speech Recognition Benchmark for Air-Traffic Communications," In: Proceedings of Interspeech 2020. Shanghai: International Speech Communication Association, 2020, pp. 2297-2301.
- [20] X. Huang, A. Acero, H.-W. Hon, and R. Reddy, "Spoken Language Processing: A Guide to Theory, Algorithm, and System Development," 1st Edition, Prentice Hall PTR, USA, 2001.
- [21] S. Davis and P. Mermelstein. "Comparison of parametric representations for mono-syllabic word recognition in continuously spoken sentences," IEEE Transactions on Acoustics, Speech, and Signal Processing 28.4, pp. 357-366, 1980.
- [22] H. Hermansky, "Perceptual linear predictive (PLP) analysis of speech," Journal of the Acoustical Society of America 87.4, pp. 1738-1752, 1990.
- [23] S. Bhatt, A. Jain, and A. Dev, "Acoustic Modeling in Speech Recognition: A Systematic Review," International Journal of Advanced Computer Science and Applications 11.4, 2020.
- [24] J. P. H. van Santen, "Handbook of Standards and Resources for Spoken Language Systems," Computational Linguistics 24.3, 1998.
- [25] D. Wang, X. Wang, and S. Lv, "An Overview of End-to-End Automatic Speech Recognition," Symmetry 11.8, 2019
- [26] T.-S. Nguyen, S. Stüker, and A. Waibel, "Toward Cross-Domain Speech Recognition with End-to-End Models," in Proceedings of the Life Long Learning for Spoken Language Systems Workshop colocated with ASRU Singapore, 2019.
- [27] A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, "Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks", Proceedings of the 23rd International Conference on Machine Learning. ICML, USA, 2006.
- [28] D. Amodei, R. Anubhai, E. Battenberg, and C. Case, "Deep Speech 2: End-to-End Speech Recognition in English and Mandarin," in Proceedings of the 33rd International Conference on Machine Learning - Volume 48, 2016.
- [29] R. Morais, "A Journey to <10% Word Error Rate," 2017 <https://hacks.mozilla.org/2017/11/a-journey-to-10-word-error-rate> (visited on 08/31/2021).
- [30] Mozilla Corporation, "DeepSpeech Model," 2020. url: <https://deepspeech.readthedocs.io/en/v0.9.3/DeepSpeech.html> (visited on 08/31/2021).
- [31] Z. C. Lipton, "A Critical Review of Recurrent Neural Networks for Sequence Learning," arXiv:1506.00019, 2015.
- [32] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and LongShort-Term Memory (LSTM) Network," Physica D: Nonlinear Phenomena Volume 404, March 2020.
- [33] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory", Neural Computation, pp. 1735-1780, 1997.
- [34] A. Hannun, "Sequence Modeling with CTC," Distill, 2017. url: <https://distill.pub/2017/ctc/> (visited on 08/31/2021).
- [35] P. Werbos, "Backpropagation through time: what it does and how to do it," Proceedings of the IEEE 78.10, pp. 1550-1560, 1990.
- [36] D. P. Kingma and J. Ba. "Adam: A Method for Stochastic Optimization", 3rd International Conference on Learning Representations ICLR, San Diego, CA, USA, 2015.
- [37] Mozilla Corporation, "CTC beam search decoder," 2020. url: <https://deepspeech.readthedocs.io/en/v0.9.3/Decoder.html> (visited on 08/31/2021).
- [38] K. Heafield, I. Pouzyrevsky, J. H. Clark, and P. Koehn, "Scalable Modified Kneser-Ney Language Model Estimation", Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, Sofia, Bulgaria, pp. 690-696, 2013.
- [39] Mozilla Corporation, "Training Your Own Model," 2020. url: <https://deepspeech.readthedocs.io/en/v0.9.3/TRAINING.html#training-a-model> (visited on 08/31/2021).
- [40] Mozilla Corporation, "Transfer-Learning (new alphabet)," 2020. url: <https://deepspeech.readthedocs.io/en/v0.9.3/TRAINING.html#transfer-learning-new-alphabet> (visited on 08/31/2021).
- [41] O. Eberhard and T. Zesch, "Effects of Layer Freezing when Transferring DeepSpeech to New Languages", arXiv:2102.04097, 2021.
- [42] H. Helmke, M. Slotty, M. Poiger, D.F. Herrer, O. Ohneiser et al., "Ontology for transcription of ATC speech commands of SESAR 2020 solution PJ.16-04," IEEE/AIAA 37th Digital Avionics Systems Conference (DASC), London, UK, 2018.
- [43] M. Kleinert, H. Helmke: S. Shetty, O. Ohneiser, H. Ehr, A. Prasad, P. Motlicek, and J. Harfmann "Automated Interpretation of Air Traffic Control Communication: The Journey from Spoken Words to a Deeper Understanding of the Meaning," IEEE/AIAA 40th Digital Avionics Systems Conference (DASC), virtual, 2021.
- [44] K. Hofbauer, S. Petrik, and H. Hering, "The ATCOSIM Corpus of Non-Prompted Clean Air Traffic Control Speech," Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), Marrakech, Morocco, 2008.
- [45] H. Helmke, S. Shetty, M. Kleinert, O. Ohneiser, H. Ehr, A. Prasad, P. Motlicek, A. Cerna, and C. Windisch, "How to Measure Speech Recognition Performance in the Air Traffic Control domain? The Word Error Rate is only half of the truth!," Satellite Workshop of Interspeech, Brno, Czech Republic, 2021.
- [46] M. Kleinert, H. Helmke, H. Ehr, Chr. Kern, D. Klakow, P. Motlicek, M. Singh, and G. Siol, "Building Blocks of Assistant Based Speech Recognition for Air Traffic Management Applications," 8th SESAR Innovation Days, Salzburg, Austria, 2018.