Literatur

[Amo+15] Dario Amodei, Rishita Anubhai, Eric Battenberg, Carl Case, Jared Casper, Bryan Catanzaro, Jingdong Chen, Mike Chrzanowski, Adam Coates, Greg Diamos, Erich Elsen, Jesse Engel, Linxi Fan, Christopher Fougner, Tony Han, Awni Hannun, Billy Jun, Patrick LeGresley, Libby Lin, Sharan Narang, Andrew Ng, Sherjil Ozair, Ryan Prenger, Jonathan Raiman, Sanjeev Satheesh, David Seetapun, Shubho Sengupta, Yi Wang, Zhiqian Wang, Chong Wang, Bo Xiao, Dani Yogatama, Jun Zhan and Zhenyao Zhu. *Deep Speech 2: End-to-End Speech Recognition in English and Mandarin*. 2015. arXiv: `1512.02595 [cs.CL]`.

[AZ19] Aashish Agarwal and Torsten Zesch. „German End-to-end Speech Recognition based on DeepSpeech“. In: *KONVENS*. 2019.

[Bae+20] Alexei Baevski, Henry Zhou, Abdelrahman Mohamed and Michael Auli. *wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations*. 2020. arXiv: `2006.11477 [cs.CL]`.

[Gai+14] C. Gaida, P. Lange, Rico Petrick, P. Proba, Ahmed Malatawy and David Suendermann-Oeft. „Comparing Open-Source Speech Recognition Toolkits“. In: 2014.

[Kah+19] Jacob Kahn, Morgane Rivière, Weiyi Zheng, Eugene Kharitonov, Qiantong Xu, Pierre-Emmanuel Mazaré, Julien Karadayi, Vitaliy Liptchinsky, Ronan Collobert, Christian Fuegen, Tatiana Likhomanenko, Gabriel Synnaeve, Armand Joulin, Abdelrahman Mohamed and Emmanuel Dupoux. *Libri-Light: A Benchmark for ASR with Limited or No Supervision*. Dez. 2019.

[MK18] Benjamin Milde and Arne Köhn. *Open Source Automatic Speech Recognition for German*. 2018. arXiv: `1807.10311 [cs.CL]`.

[Moz21] Mozilla. *Common Voice by Mozilla*. 2021. URL: `https://commonvoice.mozilla.org/en/datasets`.

[Par+19] Daniel S. Park, William Chan, Yu Zhang, Chung-Cheng Chiu, Barret Zoph, Ekin D. Cubuk and Quoc V. Le. „SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition“. In: *Proc. Interspeech 2019*. 2019, S. 2613–2617. DOI: `10.21437/Interspeech.2019-2680`. URL: `http://dx.doi.org/10.21437/Interspeech.2019-2680`.

[sil19] silenterus. *Multi-Language Dataset Cleaner/Combiner for Mozilla's DeepSpeech Framework*. `https://github.com/silenterus/deepspeech-cleaner`. 2019.

[Syn+20] Gabriel Synnaeve, Qiantong Xu, Jacob Kahn, Tatiana Likhomanenko, Edouard Grave, Vineel Pratap, Anuroop Sriram, Vitaliy Liptchinsky and Ronan Collobert. *End-to-end ASR: from Supervised to Semi-Supervised Learning with Modern Architectures*. 2020. arXiv: 1911.08460 [cs.CL].

[Tea21] Silero Team. *Silero Models: pre-trained enterprise-grade STT / TTS models and benchmarks*. https://github.com/snakers4/silero-models. 2021.

[Wan+20] Changhan Wang, Yun Tang, Xutai Ma, Anne Wu, Dmytro Okhonko and Juan Pino. *fairseq S2T: Fast Speech-to-Text Modeling with fairseq*. 2020. arXiv: 2010.05171 [cs.CL].

[Xu+20] Jiahua Xu, Kaveen Matta, Shaiful Islam and Andreas Nürnberger. „German Speech Recognition System Using DeepSpeech“. In: *Proceedings of the 4th International Conference on Natural Language Processing and Information Retrieval*. NLPIR 2020. Seoul, Republic of Korea: Association for Computing Machinery, 2020, S. 102–106. ISBN: 9781450377607. DOI: 10.1145/3443279.3443313. URL: https://doi.org/10.1145/3443279.3443313.