









- [7] Kexin Feng and Theodora Chaspari. 2020. A Review of Generalizable Transfer Learning in Automatic Emotion Recognition. *Frontiers in Computer Science* 2 (2020), 9.
- [8] Jort F Gemmeke, Daniel PW Ellis, Dylan Freedman, Aren Jansen, Wade Lawrence, R Channing Moore, Manoj Plakal, and Marvin Ritter. 2017. Audio set: An ontology and human-labeled dataset for audio events. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 776–780.
- [9] Shawn Hershey, Sourish Chaudhuri, Daniel PW Ellis, Jort F Gemmeke, Aren Jansen, R Channing Moore, Manoj Plakal, Devin Platt, Rif A Saurous, Bryan Seybold, et al. 2017. CNN architectures for large-scale audio classification. In *2017 IEEE international conference on acoustics, speech and signal processing (icassp)*. IEEE, 131–135.
- [10] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
- [11] Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861* (2017).
- [12] Jeremy Howard and Sebastian Ruder. 2018. Universal language model fine-tuning for text classification. *arXiv preprint arXiv:1801.06146* (2018).
- [13] Andrej Karpathy, George Toderici, Sanketh Shetty, Thomas Leung, Rahul Sukthankar, and Li Fei-Fei. 2014. Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 1725–1732.
- [14] Hong-Wei Ng, Viet Dung Nguyen, Vassilios Vonikakis, and Stefan Winkler. 2015. Deep learning for emotion recognition on small datasets using transfer learning. In *Proceedings of the 2015 ACM on international conference on multimodal interaction*. 443–449.
- [15] Charles C. Onu, Jonathan Lebensold, William L. Hamilton, and Doina Precup. 2019. Neural Transfer Learning for Cry-Based Diagnosis of Perinatal Asphyxia. In *Interspeech 2019, 20th Annual Conference of the International Speech Communication Association, Graz, Austria, 15-19 September 2019*, Gernot Kubin and Zdravko Kacic (Eds.). ISCA, 3053–3057. <https://doi.org/10.21437/Interspeech.2019-2340>
- [16] Maxime Oquab, Leon Bottou, Ivan Laptev, and Josef Sivic. 2014. Learning and transferring mid-level image representations using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1717–1724.
- [17] Sandeep Kumar Pandey, HS Shekhawat, and SRM Prasanna. 2019. Deep Learning Techniques for Speech Emotion Recognition: A Review. In *2019 29th International Conference Radioelektronika (RADIOELEKTRONIKA)*. IEEE, 1–6.
- [18] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. 2019. PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*. 8024–8035.
- [19] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research* 12 (2011), 2825–2830.
- [20] Soujanya Poria, Erik Cambria, Rajiv Bajpai, and Amir Hussain. 2017. A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion* 37 (2017), 98–125.
- [21] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. 3973–3983.
- [22] Nils Reimers and Iryna Gurevych. 2020. Making Monolingual Sentence Embeddings Multilingual using Knowledge Distillation. *arXiv preprint arXiv:2004.09813* (04 2020). <http://arxiv.org/abs/2004.09813>
- [23] Julian Risch, Anke Stoll, Marc Ziegele, and Ralf Krestel. 2019. hpiDEDIS at GemEval 2019: Offensive Language Identification using a German BERT model. In *Preliminary proceedings of the 15th Conference on Natural Language Processing (KONVENS 2019)*. Erlangen, Germany: German Society for Computational Linguistics & Language Technology, 403–408.
- [24] Wendy A Rogers and Tracy L Mitzner. 2017. Envisioning the future for older adults: Autonomy, health, well-being, and social connectedness with technology support. *Futures* 87 (2017), 133–139.
- [25] Sourav Sahoo, Puneet Kumar, Balasubramanian Raman, and Partha Pratim Roy. 2019. A Segment Level Approach to Speech Emotion Recognition Using Transfer Learning. In *Asian Conference on Pattern Recognition*. Springer, 435–448.
- [26] Bjorn W. Schuller, Anton Batliner, Christian Bergler, Eva-Maria Messner, Antonia Hamilton, Shahin Amiriparian, Alice Baird, Georgios Rizos, Maximilian Schmitt, Lukas Stappen, Harald Baumeister, Alexis Deighton MacIntyre, and Simone Hantke. 2020. The INTERSPEECH 2020 Computational Paralinguistics Challenge: Elderly Emotion, Breathing & Masks. In *Proceedings of Interspeech*. Shanghai, China, 5 pages. to appear.
- [27] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. 2014. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* 15, 1 (2014), 1929–1958.
- [28] Juanjuan Wang, Mantao Xu, Hui Wang, and Jiwu Zhang. 2006. Classification of imbalanced data by using the SMOTE algorithm and locally linear embedding. In *2006 8th international Conference on Signal Processing*, Vol. 3. IEEE.
- [29] Zixiaofan Yang and Julia Hirschberg. 2018. Predicting Arousal and Valence from Waveforms and Spectrograms Using Deep Neural Networks. In *Interspeech*. 3092–3096.