















- Orange Labs, July 2019.
- [12] J. Salamon and J. P. Bello, "Deep convolutional neural networks and data augmentation for environmental sound classification", in *IEEE Signal Processing Letters*, vol. 24, no. 3, pp. 279-283, March 2017.
- [13] D. S. Park, W. Chan, Y. Zhang, C.-C. Chiu, B. Zoph, E. D. Cubuk, and Q. V. Le, "SpecAugment: A simple data augmentation method for automatic speech recognition", in *Proc. INTERSPEECH 2019*, pp. 2613-2617, arXiv:1904.08779.
- [24] S. Laine and T. Aila, "Temporal Ensembling for Semi-Supervised Learning", in *ICLR 2017*, arXiv: 1610.02242, 2017.
- [14] J. Ebberts, and R. Häb-Umbach, "Convolutional recurrent neural network and data augmentation for audio tagging with noisy labels and minimal supervision", in *Proc. DCASE2019 Workshop*, New York University, NY, USA, October 2019, pp. 64-68.
- [15] W. Lim, "SpecAugment for sound event detection in domestic environments using ensemble of convolutional recurrent neural networks", in *Proc. DCASE2019 Workshop*, New York University, NY, USA, October 2019, pp. 129-133.
- [16] N. Takahashi, M. Gygli, B. Pfister, and L. Van Gool, "Deep convolutional neural networks and data augmentation for acoustic event recognition", in *Proc. INTERSPEECH 2016*, arXiv:1604.07160.
- [17] T. Inoue, P. Vinayavekhin, S. Wang, D. Wood, A. Munawar, B. J. Ko, N. Greco, and R. Tachibana, Ryuki, "Shuffling and mixing data augmentation for environmental sound classification", in *Proc. DCASE2019 Workshop*, New York University, NY, USA, October 2019, pp. 109-113.
- [18] Y. Tokozume, Y. Ushiku, and T. Harada, "Learning from between-class examples for deep sound recognition", in *Proc. ICLR 2018*, arXiv:1711.10282.
- [19] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "Mixup: Beyond empirical risk minimization", in *Proc. ICLR 2018*, arXiv:1710.09412.
- [20] P. Pratik, W. J. Jee, S. Nagisetty, R. Mars, and C. Lim, Chongsoon, "Sound event localization and detection using CRNN architecture with mixup for model generalization", in *Proc. DCASE2019 Workshop*, New York University, NY, USA, October 2019, pp. 99-203.
- [21] L. Delphin-Poulat, C. Plapous, and R. Nicol, "GCNN for classification of domestic activities", *Tech. Rep. in DCASE2018 Challenge*, Orange Labs, July 2018.
- [22] J. F. Gemmeke, D. P. W. Ellis, D. Freedman, A. Jansen, W. Lawrence, R. C. Moore, M. Plakal, and M. Ritter, "Audio set: An ontology and human-labeled dataset for audio events", in *Proc. ICASSP*, 2017.
- [23] L. JiaKai, "Mean teacher convolution system for DCASE 2018 Task 4", *Tech. Rep. in DCASE2018 Challenge*, PFU Shanghai, July 2018.
- [25] W. Lim, S. Suh, and Y. Jeong, "Weakly labeled semi-supervised sound event detection using CRNN with inception module", *Tech. Rep. in DCASE2018 Challenge*, Electronics and Telecommunications Research Institute, Korea, July 2018.
- [26] H. Dinkel, and K. Yu, "Duration robust weakly supervised sound event detection", in *Proc. ICASSP*, 2020, arXiv:1904.03841v3.
- [27] A. Mesaros, T. Heittola and T. Virtanen, "Metrics for polyphonic sound event detection", in *Appl. Sci*, 6, 162, 2016.