















speech quality (PESQ) – a new method for speech quality assessment of telephone networks and codecs,” in *Proc. ICASSP*, 2001.

Cees H. Taal, Richard C. Hendriks, Richard Heusdens, Jesper Jensen, "An algorithm for intelligibility prediction of time–frequency weighted noisy speech,” *IEEE Trans on Audio, Speech, and Language Processing*, 2011.

Christophe Veaux, Junichi Yamagishi and Simon King, "The voice bank corpus: Design, collection and data analysis of a large regional accent speech database," in *Proc. O-COCOSDA/CASLRE*, 2013

Dean, David and Sridharan, Sridha and Vogt, Robert and Mason, Michael. "The QUT-NOISE-TIMIT corpus for the evaluation of voice activity detection algorithms," in *Proc. Interspeech*, 2010.

Fu-An Chao, Jieh-weih Hung and Berlin Chen, "Cross-Domain Single-Channel Speech Enhancement Model with BI-Projection Fusion Module for Noise-Robust ASR," in *Proc. ICME*, 2021.

Fu-En Wang et al., “BiFuse: Monocular 360° depth estimation via bi-projection fusion,” in *Proc. CVPR*, 2020

Ian Goodfellow, Yoshua Bengio and Aaron Courville, “Deep learning,” *MIT Press*, 2016

Jingjing Chen, Qirong Mao, Dong Liu, "Dual-Path Transformer Network: Direct Context-Aware Modeling for End-to-End Monaural Speech Separation," in *Proc. Interspeech*, 2020

Joachim Thiemann, Nobutaka Ito, and Emmanuel Vincent, "Demand: a collection of multi-channel recordings of acoustic noise in diverse environments," in *Proc. ICA*, 2013

Kahneman, D., Sibony, O., and Sunstein, C. R. "Noise: a flaw in human judgment," *First edition, New York, Little, Brown Spark*, 2021

Philipos C. Loizou, "Speech Enhancement: Theory and Practice," *CRC Press*, 2013

Stephane Mallat, “A Wavelet Tour of Signal Processing,” *2nd ed., San Diego, CA: Academic*, 1999

Yi Luo and Nima Mesgarani, "Conv-TasNet: Surpassing Ideal Time–Frequency Magnitude Masking for Speech Separation," *IEEE/ACM Trans on Audio, Speech, and Language Processing*, 2019.

Yusuf Isik, Jonathan Le Roux, Zhuo Chen, Shinji Watanabe, and John R. Hershey, "Single-channel multi-speaker separation using deep clustering,” in *Proc. Interspeech*, 2016