MEFT
Boosting the Future

Contribution ID: **75**   Type: **Workshop 2025/2026**

Noise isolation and voice isolation for crowds and festivals

Thursday 29 January 2026 18:15 (15 minutes)

Monaural speech enhancement is the task of improving how clear and understandable speech sounds when it's recorded in noisy, reverberant environments using only a single microphone. Traditional signal-processing methods depend on stochastic assumptions about how speech and noise behave, which makes them struggle in real-world situations where noise changes over time or the signal-to-noise ratio is low. More recently, deep learning methods have outperformed these approaches by learning rich spectral representations directly from data, performing particularly well when nonstationary noise is present.

This report focuses on frequency-domain speech enhancement methods, with a particular emphasis on deep learning–based techniques. It covers the main building blocks of these systems, including feature extraction, network architectures, training targets, and loss functions. Special attention is given to complex-valued models, especially convolutional recurrent networks, which currently achieve state-of-the-art results. This report sets the basis for a project exploring the state-of-the-art speech enhancement deep learning techniques and their potential applications.

Field of Research/Work

Beyond Physics

Author: FELICIANO, André (Instituto Superior Técnico - University of Lisbon)