

DeepFilterNet: A Low Complexity Speech Enhancement Framework based on Deep Filtering

Hendrik Schröter¹, Alberto N. Escalante-B.², Tobias Rosenkranz², Andreas Maier¹

¹Pattern Recognition Lab, Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany

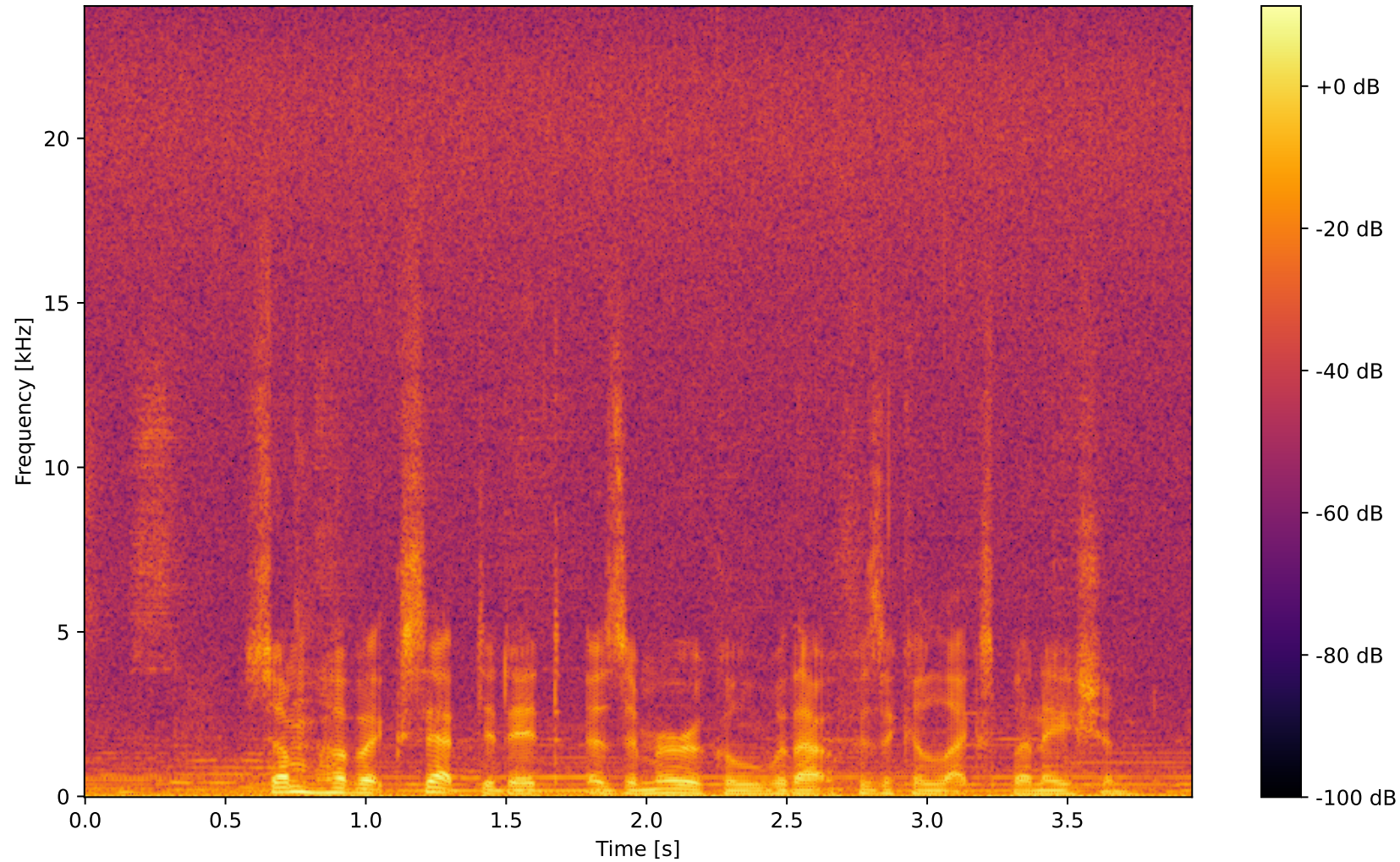
²WS Audiology, Research and Development, Erlangen, Germany

ICASSP 2022, Singapore

- Perceptual speech enhancement approach
- Full band: 48 kHz
- Real-Time Capabilities: Supported latencies from 5 ms to 40 ms
- Low Complexity: 1.8 M Parameters, 0.35 GMACS
- Open Source: github.com/Rikorose/DeepFilterNet
- Demo: <https://huggingface.co/spaces/hshr/DeepFilterNet>

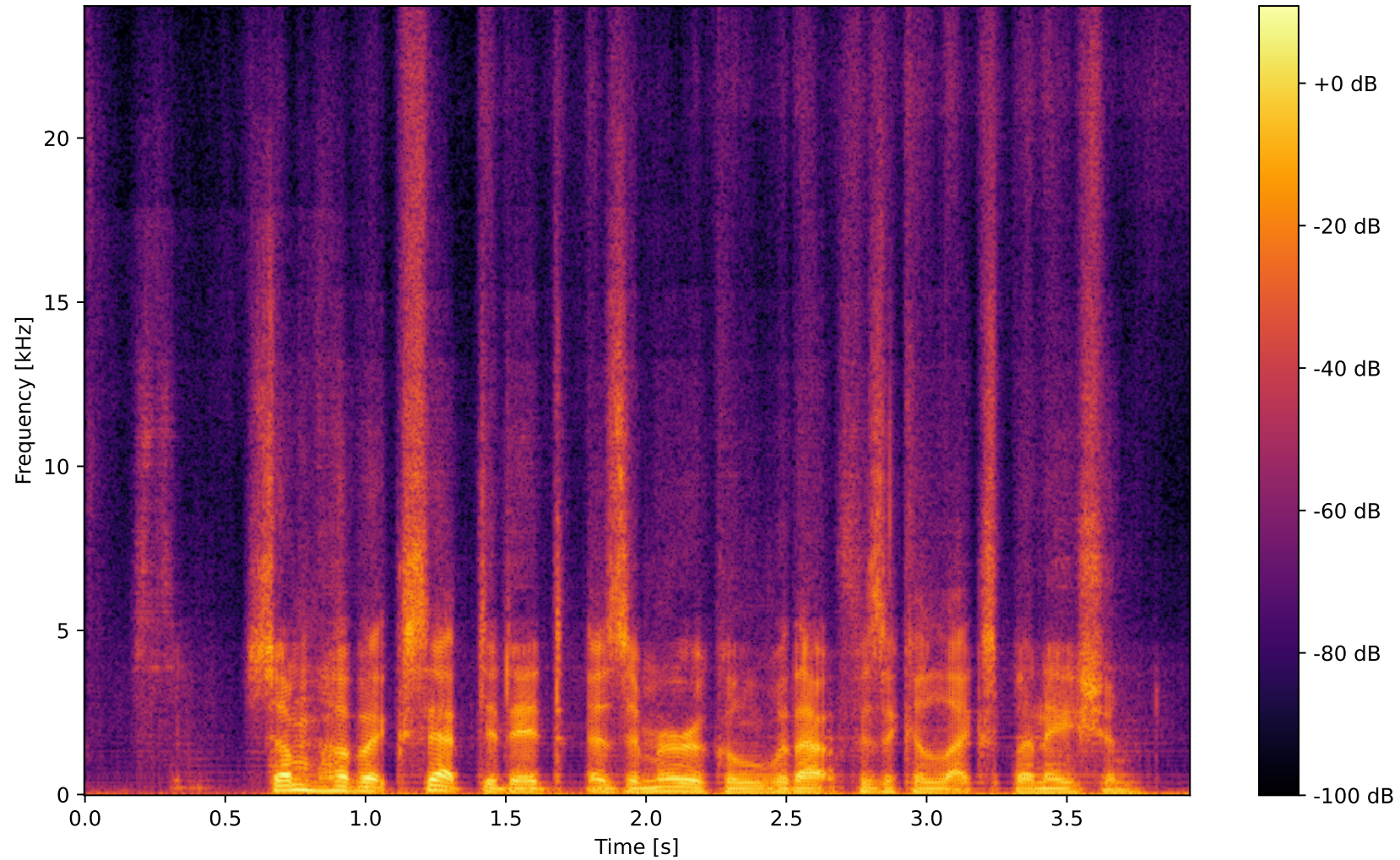
Perceptual Approach

Speech: Periodic + stochastic component



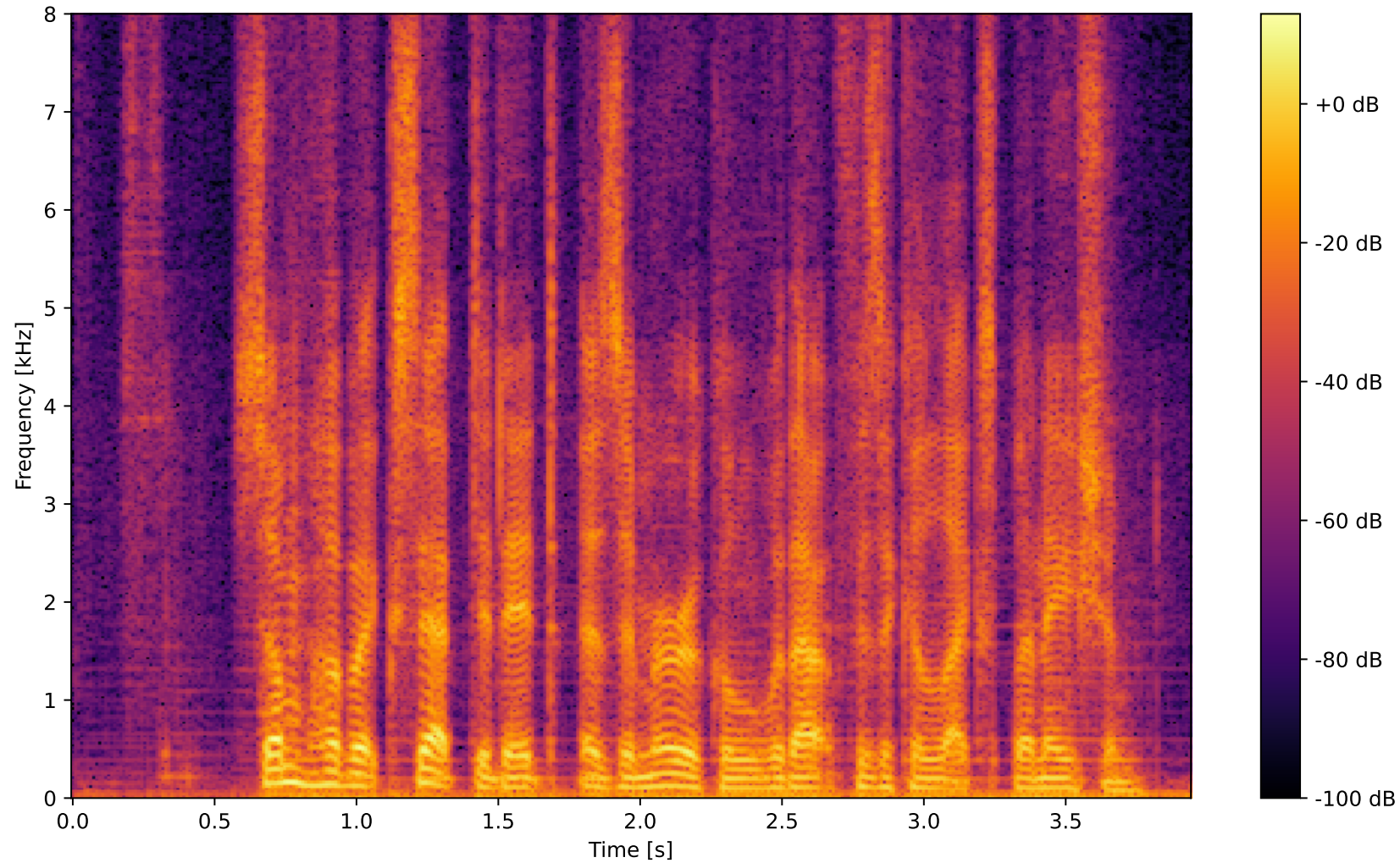
Perceptual Approach

Stage 1: Enhance speech envelope



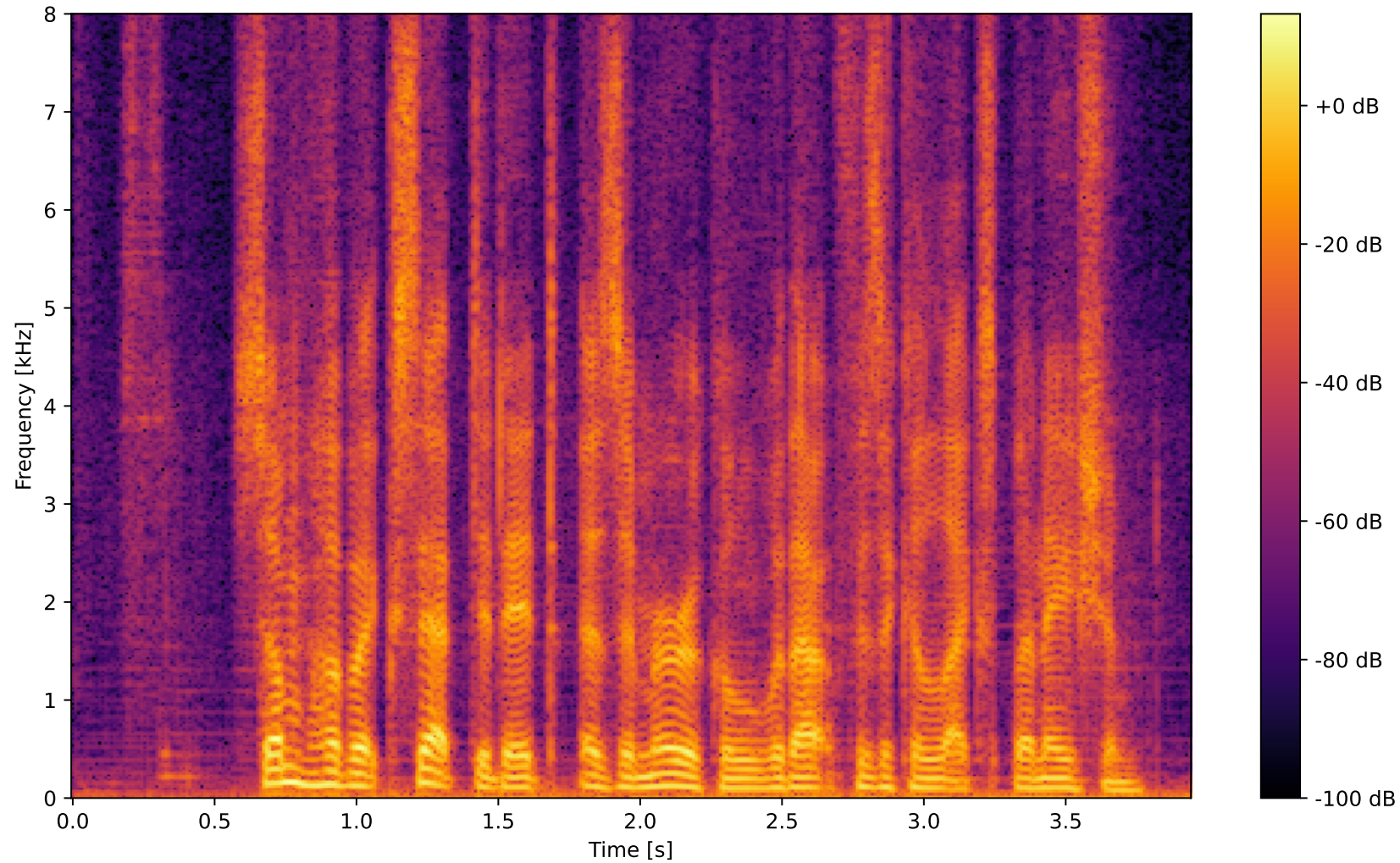
Perceptual Approach

Stage 1: Enhance speech envelope

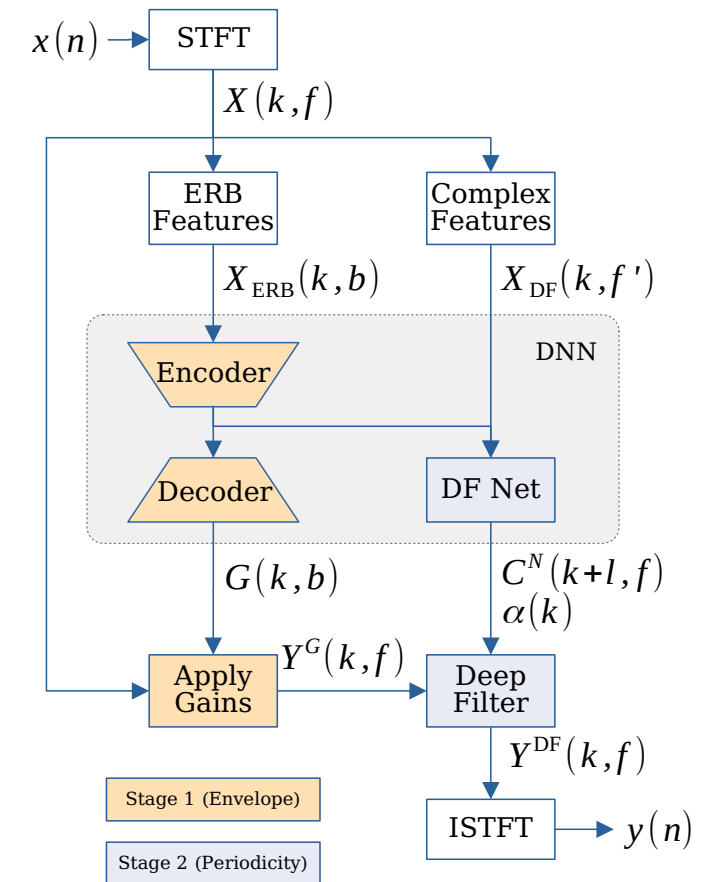


Perceptual Approach

Stage 2: Deep filtering to enhance periodic component



- Framework operates in frequency domain (FD)
- Support for various FFT and hop sizes depending on latency requirements
- Stage 1 enhances overall speech envelope
 - Operates logarithmically scaled magnitude features (ERB)
 - Only 32 ERB bins allow to construct a small network
 - U-Net like convolutional encoder/decoder structure
- Stage 2 enhances speech periodicity
 - Operates in complex FD up to a frequency F_{DF}
 - Deep Filtering enhances periodic speech components

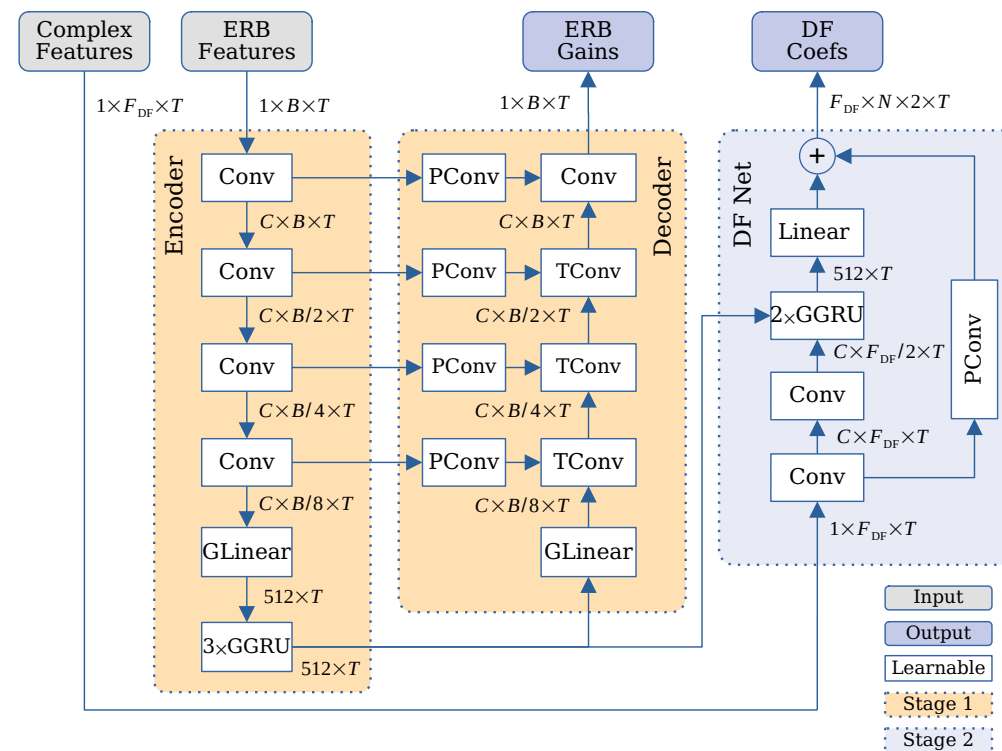


- Efficient & standard DNN layers (conv, bn, relu)
- Small DNN I/O
- DNN Sparseness via grouping [1]
- 1.8 M Parameters, 0.35 GMACS

Training Details:

- Trained on DNS 3 dataset [2]
- 32 ERB bands for stage 1 I/O
- Upper frequency for stage 2 DF: $f_{DF} = 5$ kHz
- Loss on compressed complex spectrum:

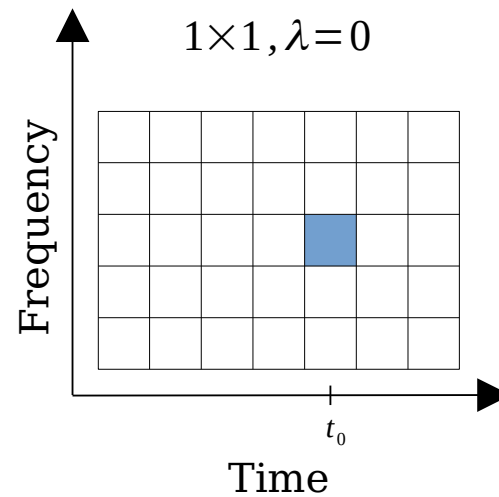
$$\mathcal{L}_{spec} = \sum_{k,f} || |Y|^c - |S|^c ||^2 + \sum_{k,f} || |Y|^c e^{j\varphi_Y} - |S|^c e^{j\varphi_S} ||^2$$



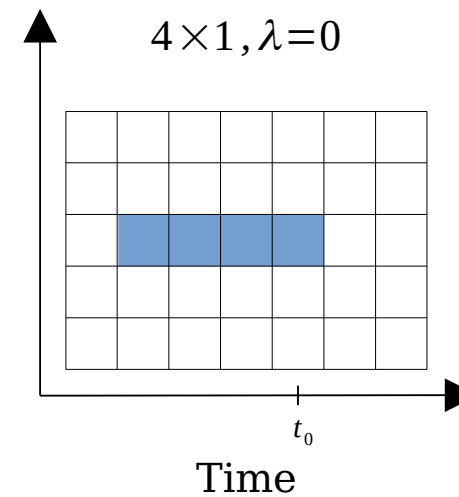
Deep Filtering

Originally proposed by [3], [4]

Complex Ratio Mask (CRM)



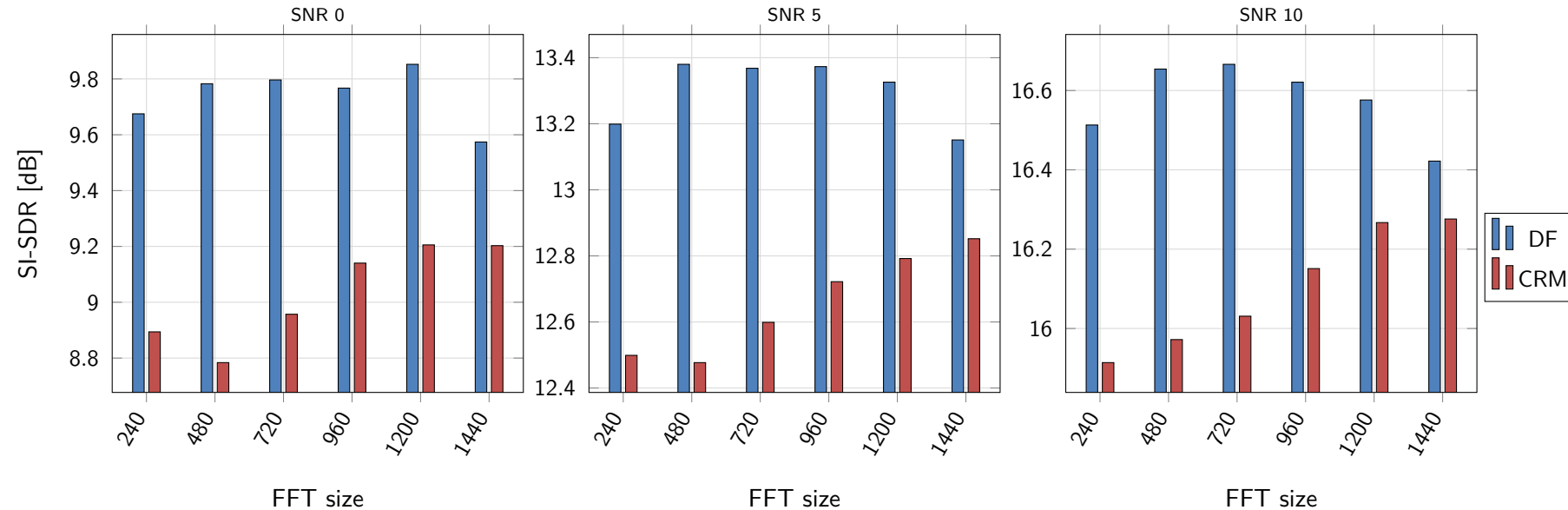
Deep Filtering (DF)



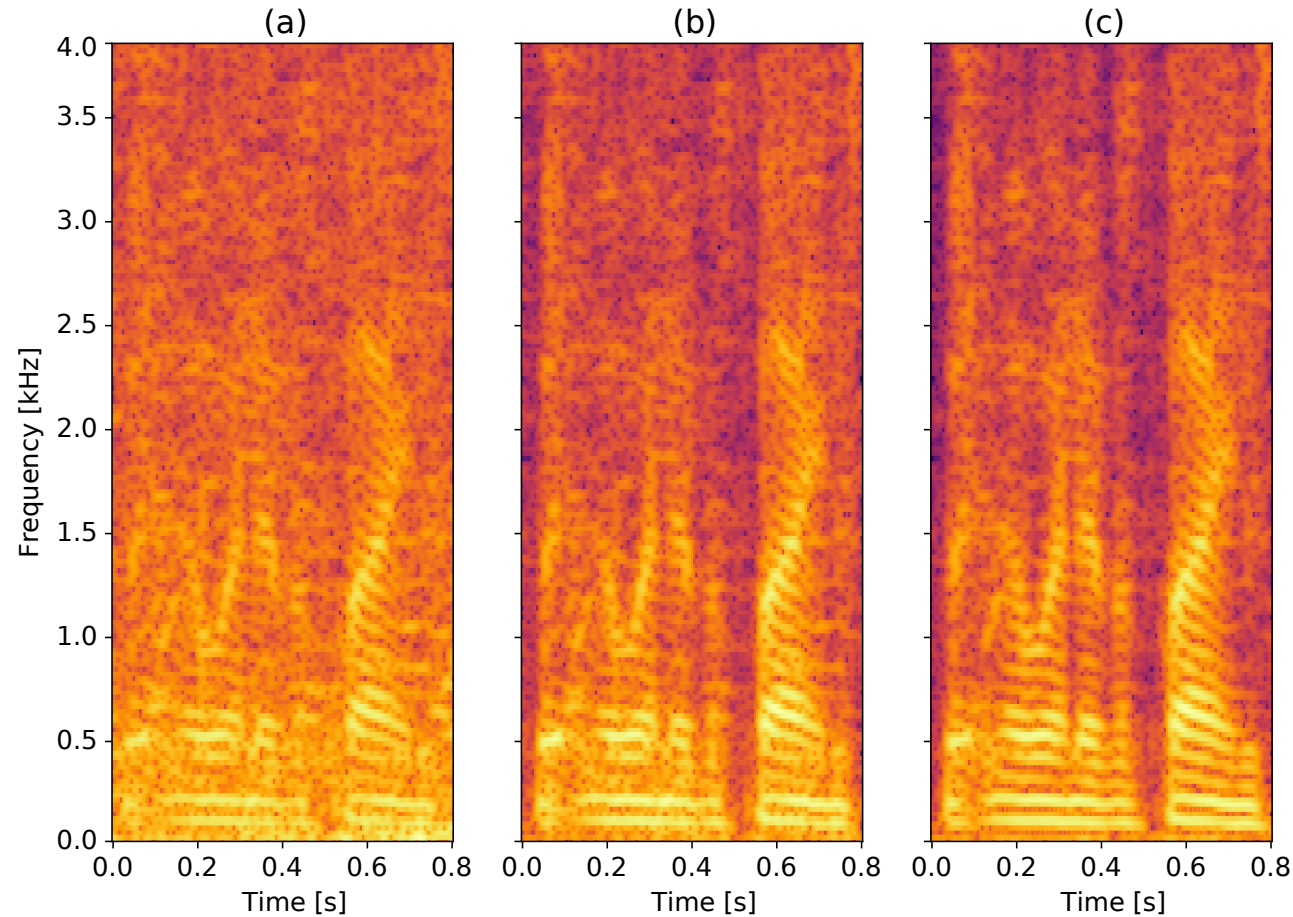
$$Y(k, f) = \sum_{i=0}^N C(k, i, f) \cdot X(k - i + l, f)$$

where C are the predicted FD deep filtering coefficient with filter order N .

Comparison between deep filtering and a traditional complex ratio mask (CRM) based on the internal test set.



Qualitative sample from the Voicebank/DEMAND test set. Noisy (a), CRM (b), DF (c).



Comparison with related work based on the Voicebank/DEMAND test set [5].

Model	Params [M]	MACS [G]	WB-PESQ [MOS]	SI-SDR [dB]
Noisy	-	-	1.97	8.41
PercepNet [6]	8.0	0.80	2.73	-
DCCRN [7]	3.7	14.36	2.68	-
DCCRN+ [8]	3.3	-	2.84	-
DeepFilterNet	1.8	0.35	2.81	16.63
w/o stage 2	0.9	0.25	2.57	13.81

We have code!

<https://github.com/Rikorose/DeepFilterNet>



Search or jump to... Pull requests Issues Marketplace Explore

Rikorose / DeepFilterNet (Public) Unpin Unwatch (4) Fork (17) Starred (72)

<> Code Issues (5) Pull requests (3) Discussions Actions Security Insights Settings

main 12 branches 9 tags Go to file Add file Code

File	Commit	Time
Rikorose build: Add Ito	5a515eb	23 minutes ago
.github	ci: Add stale bot	3 days ago
DeepFilterNet	df: Limit number of channels in hdf5 script	last month
assets	scripts: Fix dns4 noise processing; other misc	2 months ago
libDF	libdf: Compatibility with rubtato 0.11	28 days ago
pyDF-data	build: Add Ito	23 minutes ago
pyDF	build: Add Ito	23 minutes ago
scripts	scripts: Fix dns4 noise processing; other misc	2 months ago
.flake8	Add DeepFilterNet framework source code	6 months ago
.gitignore	gitignore: Ignore poetry.lock	4 months ago
Cargo.toml	build: Add Ito	23 minutes ago
LICENSE	build: Clarify license for crates.io	5 months ago
LICENSE-APACHE	Add license	6 months ago
LICENSE-MIT	Add license	6 months ago
README.md	Update citation	4 days ago
post-release.sh	build: Fix extra dependencies	4 months ago
pyproject.toml	build: Add top-level pyproject for better isort/black formatting	5 months ago
release.sh	build: Publish crates first before pushing release tag	4 months ago
rustfmt.toml	Add DeepFilterNet framework source code	6 months ago

302 commits

Noise suppresion using deep filtering

huggingface.co/spaces/hshr/deepfilternet

audio deep-learning speech pytorch speech-enhancement noise-suppression

Readme View license 72 stars 4 watching 17 forks

Releases (6)

v0.1.10 (Latest) on Dec 10, 2021

+ 5 releases

Contributors (3)

Rikorose Hendrik Schröter dependabot[bot] gleb-shshn Gleb

Languages

- Python 49.5%
- Rust 45.5%
- Shell 5.0%

README.md

DeepFilterNet

A Low Complexity Speech Enhancement Framework for Full-Band Audio (48kHz) based on Deep Filtering (Paper). Audio samples from the voice bank/DEMAND test set can be found [here](#).

New! Try out DeepFilterNet with your own voice at this [HuggingFace Space](#).



DeepFilterNet Samples

DeepFilterNet: A Low Complexity Speech Enhancement Framework for Full-Band Audio based on Deep Filtering

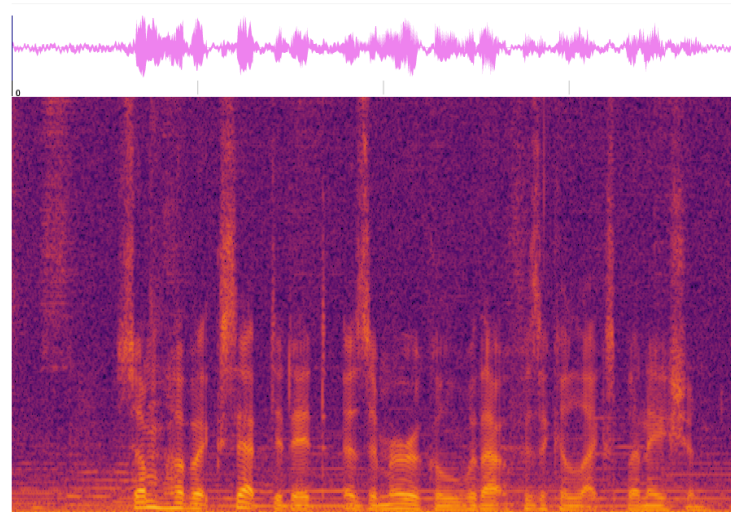
[Home](#) [About](#)

AUDIO SAMPLES

- [Sample 1](#)
- [Sample 2](#)
- [Sample 3](#)
- [Sample 4](#)
- [Sample 5](#)
- [Sample 6](#)
- [About](#)

Sample 1

Noisy



Play / Pause

[Noisy](#)

[DeepFilterNet \(stage 1 only\)](#)

[DeepFilterNet \(stage 1 + CRM\)](#)

[DeepFilterNet](#)

[Clean](#)

Demo

<https://huggingface.co/spaces/hshr/DeepFilterNet>



Hugging Face

Models Datasets Spaces Docs Solutions Pricing Log In Sign Up

Spaces: hshr/DeepFilterNet like 1 Running

App Files and versions

DeepFilterNet Demo

This demo denoises audio files using DeepFilterNet. Try it with your own voice!

Record your own voice (optional)

Alternative: Upload speech sample (optional)

Upload noise sample (optional)

snr

Clear Submit

Noisy 1.1s

Output 2

Enhanced

Output 4



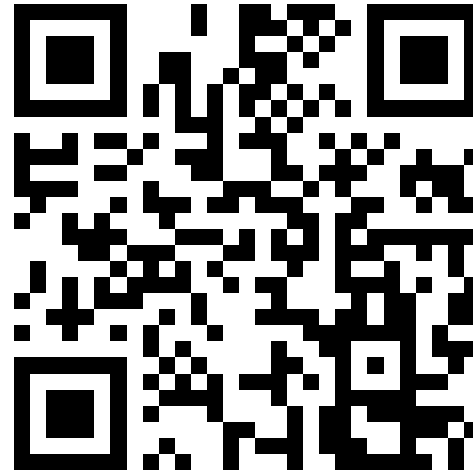
- [1] K. Tan and D. Wang, "Learning complex spectral mapping with gated convolutional recurrent networks for monaural speech enhancement," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 28, pp. 380–390, 2019.
- [2] C. K. A. Reddy, E. Beyrami, H. Dubey, *et al.*, *The INTERSPEECH 2020 Deep Noise Suppression Challenge: Datasets, Subjective Speech Quality and Testing Framework*, 2020. eprint: 2001.08662.
- [3] W. Mack and E. A. Habets, "Deep Filtering: Signal Extraction and Reconstruction Using Complex Time-Frequency Filters," *IEEE Signal Processing Letters*, vol. 27, pp. 61–65, 2020.
- [4] H. Schröter, T. Rosenkranz, A. Escalante Banuelos, M. Aubreville, and A. Maier, "CLCNet: Deep learning-based noise reduction for hearing aids using complex linear coding," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Barcelona), 2020. [Online]. Available: <https://rikorose.github.io/CLCNet-audio-samples.github.io/>.
- [5] C. Valentini-Botinhao, X. Wang, S. Takaki, and J. Yamagishi, "Investigating RNN-based speech enhancement methods for noise-robust Text-to-Speech," in *SSW*, 2016, pp. 146–152.
- [6] J.-M. Valin, U. Isik, N. Phansalkar, R. Giri, K. Helwani, and A. Krishnaswamy, "A Perceptually-Motivated Approach for Low-Complexity, Real-Time Enhancement of Fullband Speech," in *INTERSPEECH 2020*, Oct. 25–29, 2020.
- [7] Y. Hu, Y. Liu, S. Lv, *et al.*, "DCCRN: Deep complex convolution recurrent network for phase-aware speech enhancement," in *INTERSPEECH*, 2020.
- [8] S. Lv, Y. Hu, S. Zhang, and L. Xie, "DCCRN+: Channel-wise Subband DCCRN with SNR Estimation for Speech Enhancement," in *INTERSPEECH*, 2021.
- [9] K. Tan and D. Wang, "Complex spectral mapping with a convolutional recurrent network for monaural speech enhancement," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2019, pp. 6865–6869.

Thank You!

Contact



Code



Samples

