TRNS Stochastic Damping

Speech Denoising – Real-World Fan Noise

Edge-Oriented Front-End Demonstration

1. Objective

Always-on voice interfaces and low-power automatic speech recognition (ASR) front-ends frequently encounter persistent tonal interference from fans, HVAC systems, and machinery while operating under strict power and memory limits. Large neural enhancers can suppress such tones effectively, yet they often muffle the voice and exceed the computational budget of the target hardware. A lightweight, training-free alternative that preserves voice character is therefore of practical interest.

This demonstration asks whether a training-free TRNS operator, used alone or inside a mild hybrid configuration, can meaningfully attenuate a strong real-world fan tone while retaining high voice-band correlation and remaining suitable for edge deployment.

2. Test Material

A short real-world recording was made of a male voice speaking the phrase “I am a robot” directly in front of a running fan. The dominant interferer is a strong, continuous high-pitched tonal component lying inside the upper speech band. The recording is uncontrolled and deliberately difficult, representing a common and challenging condition for always-on edge devices.

3. Method

The core TRNS operator was evaluated alone and inside a hybrid pipeline that added a mild classical spectral gate together with a light tonal notch. For context, a pre-trained neural speech enhancer and classical Wiener filtering were also examined. Preference was given to configurations that remained fully training-free and faster than real time.

Performance was quantified by two complementary metrics: high-frequency energy reduction (in dB) relative to the original recording, and voice-band correlation that measures preservation of the underlying speech morphology.

4. Results

Quantitative comparison of high-frequency reduction and voice-band correlation:

Version	HF Reduction	Voice Corr
Original (noisy)	0.00 dB	1.000
TRNS alone	0.17 dB	0.999
TRNS + Mild Gate (preferred)	**2.64 dB**	**0.987**

Preferred hybrid achieved meaningful high-frequency reduction while keeping very high voice-band correlation.

Figure 1. Original spectrogram of the fan-contaminated recording (Speech_231.wav). Continuous high-pitched tonal energy is visible across the upper speech band.

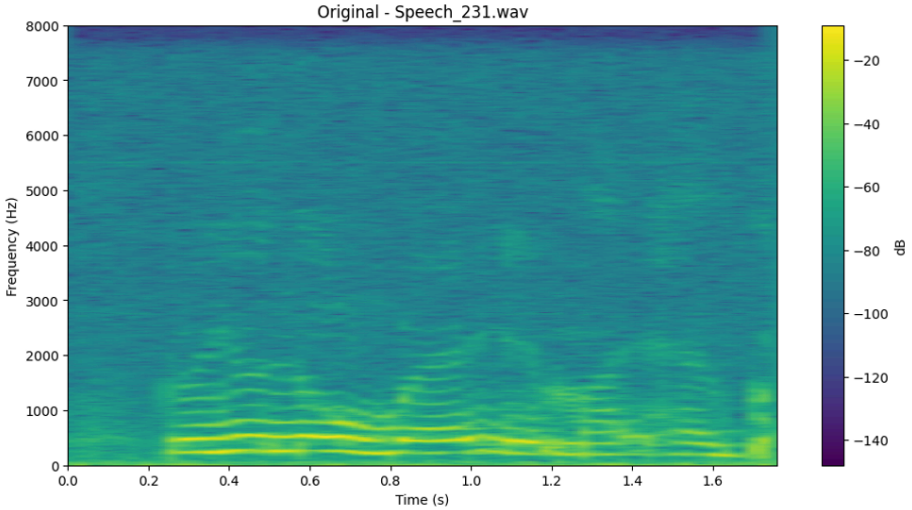
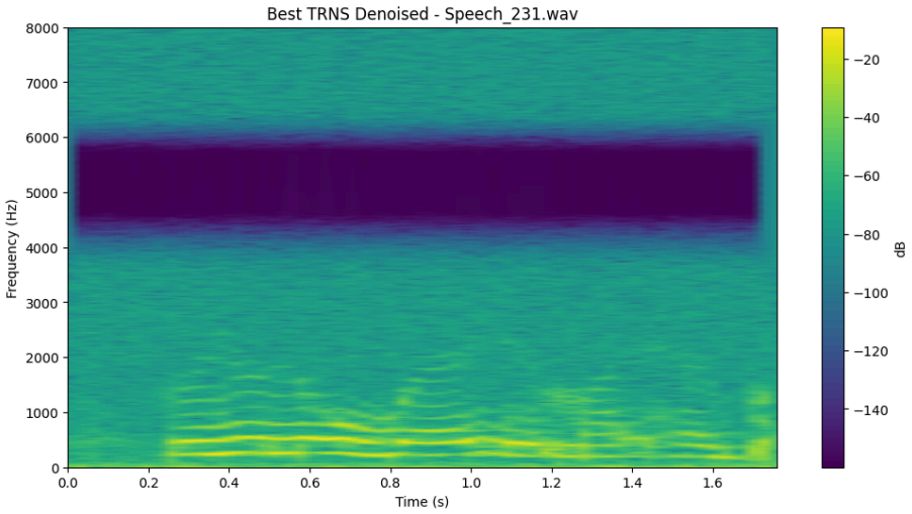


Figure 2. Best TRNS Denoised spectrogram (preferred hybrid configuration). The continuous high-frequency tonal band has been substantially attenuated while lower-band speech structure is retained.



5. Interpretation

The preferred hybrid reduced high-frequency energy by approximately 2.6 dB while preserving a voice-band correlation of 0.987 and remaining computationally light. Pure neural enhancement suppressed the fan tone more aggressively but also muffled the voice. Classical Wiener filtering introduced audible musical-noise artifacts. The hybrid therefore occupies a useful middle ground for edge devices that cannot afford the quality-versus-compute trade-offs of large neural models.

Visual inspection of the spectrograms confirms that the continuous high-pitched tonal component is markedly attenuated while the lower-frequency formant structure of the spoken phrase remains intact.

6. Positioning

Edge-suitable front-end or regularizer rather than a competitor to large neural speech enhancers. Strongest demonstrated results to date are on tonal and structured interferers commonly encountered by always-on voice interfaces operating under power and memory constraints.

7. Limitations

Results are reported on a single short real-world recording containing a strong continuous fan tone. Performance may vary with different speakers, noise types, signal-to-noise ratios, or recording conditions. The present study focuses on high-frequency tonal attenuation and voice-band correlation rather than full ASR word-error-rate evaluation or multi-noise robustness testing.

8. Conclusion

On a deliberately difficult real-world fan-noise recording, the preferred TRNS hybrid achieved meaningful high-frequency attenuation while preserving high voice-band correlation. Combined with a training-free formulation and edge-oriented design, these properties support its use as a practical lightweight front-end for always-on voice interfaces operating under strict resource constraints.

References

Real-world recording performed by Meadows McKnight LLC for internal evaluation.