Exploiting Voice Cloning in Adversarial Simulation

Mark Foudy

Boston Hacker

@0xM4rk7homas

DEFCON 32

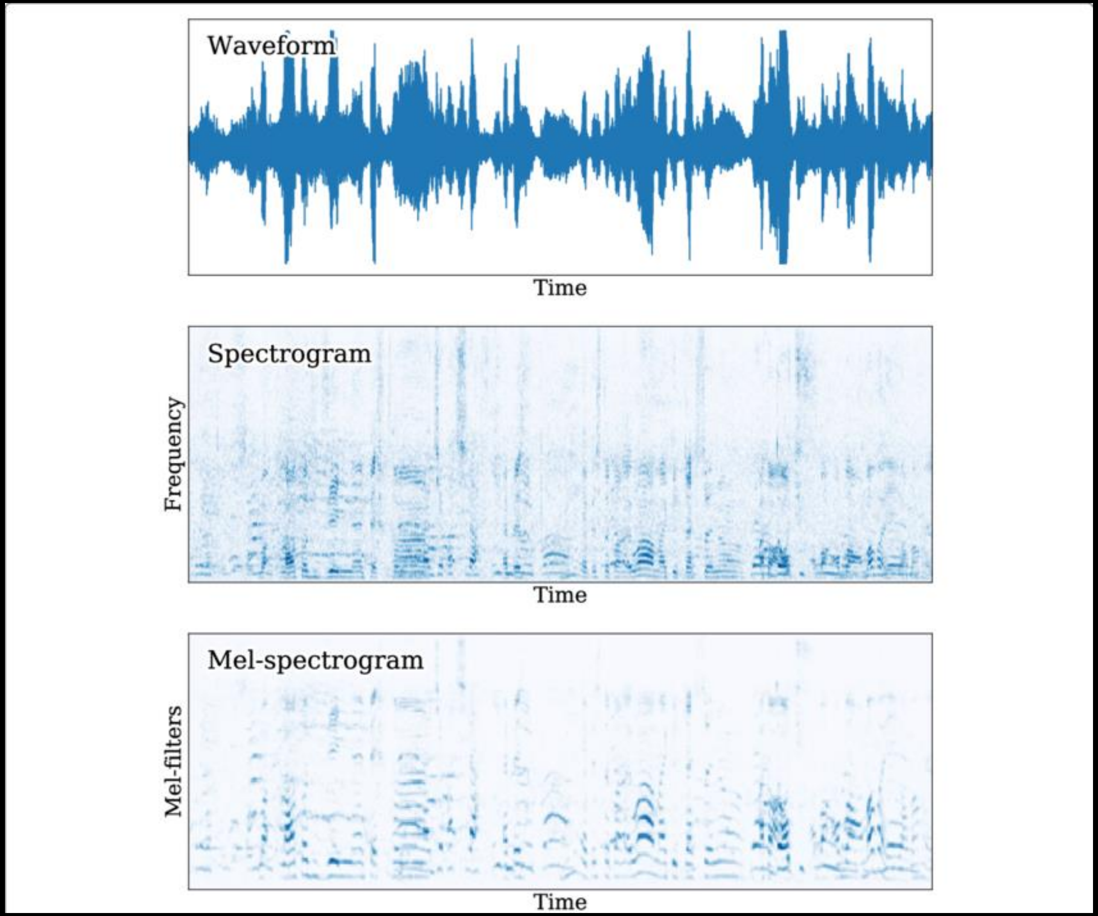
Adversary Village

https://github.com/MarkFoudy/ACOUSTIC-Standards-for-Modifying-Spoof-Speech

Who am I?

- Father, Husband, Hacker
- AI Offensive Security Researcher
- Founder of Neurodiverse Hackers

Focus Areas



Voice Verification Services (VVS)

Anti-Spoofing Verification Services (ASVS)

Mel Spectrogram -A visual representation of sound frequencies, scaled to the mel scale to reflect how humans perceive pitch.

Voice Verification Services (VVS)

Companies Providing VVS

FORTRESS IDENTITY

secure biometric identification

VERINT

NUANCE

Financial Sector Utilizing these Services

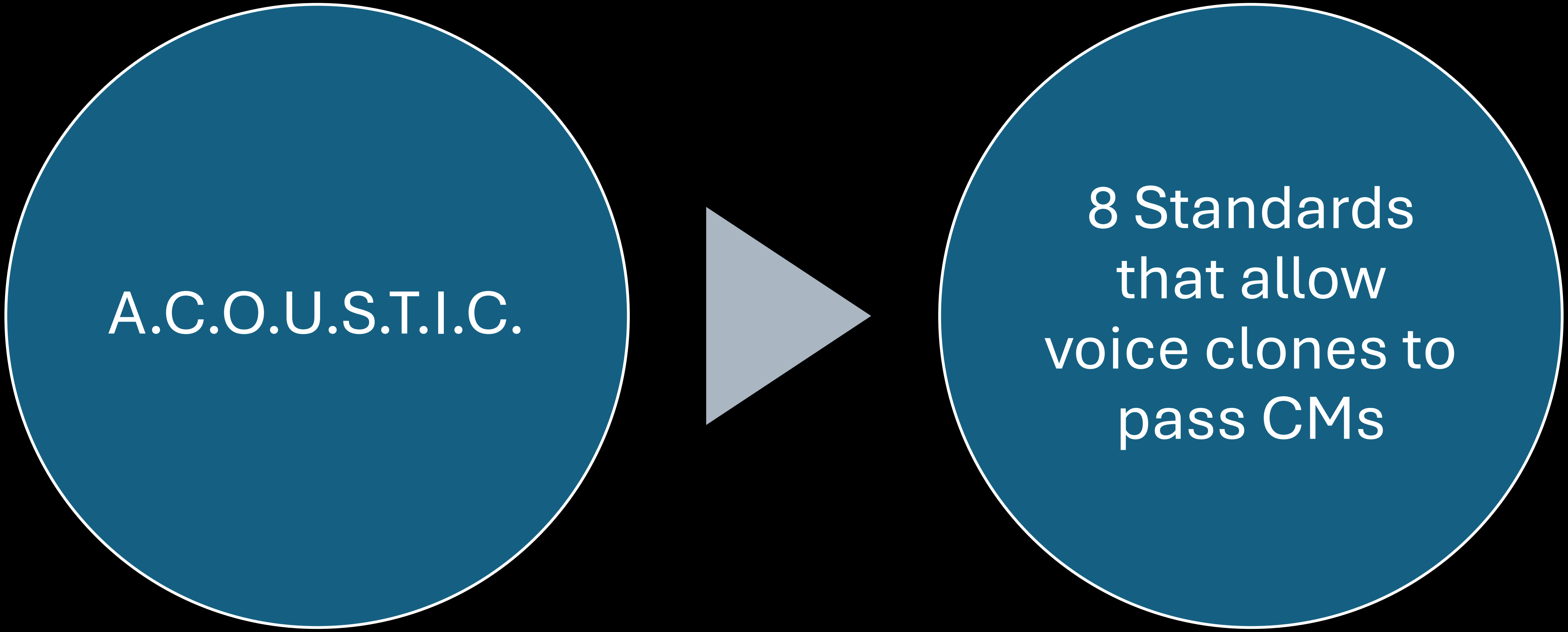
Fidelity

INVESTMENTS

WELLS FARGO

citibank

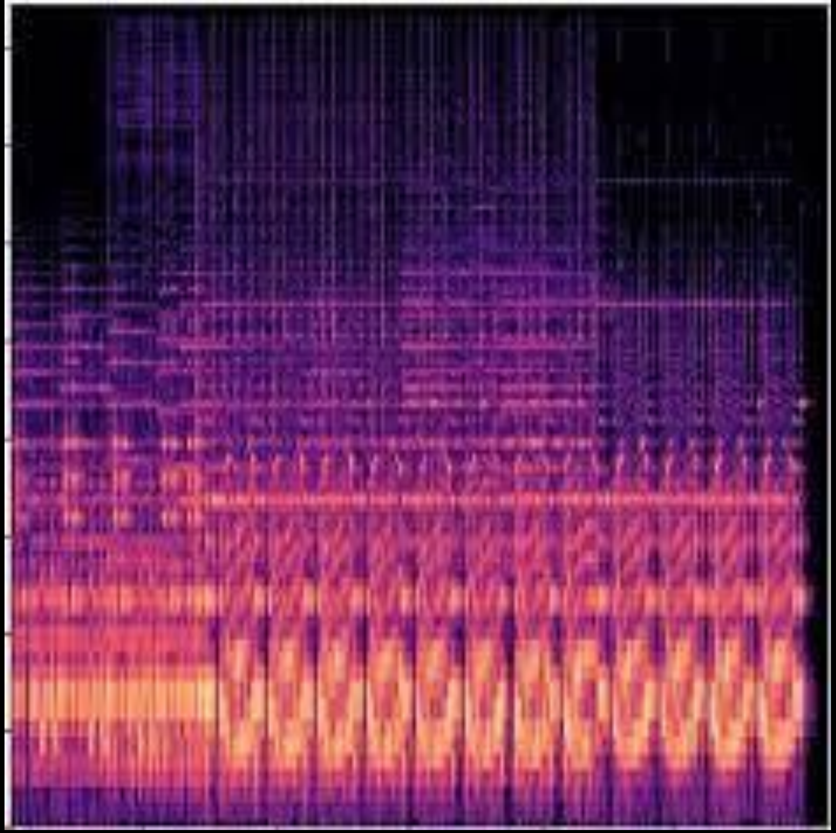
Methodology



A.C.O.U.S.T.I.C.

8 Standards that allow voice clones to pass CMs

A.C.O.U.S.T.I.C – Standards for Modifying Spoofed Speech



A – Adjust Silence Intervals

C – Center Spectrum Boosting

O – Optimize Echo Simulation

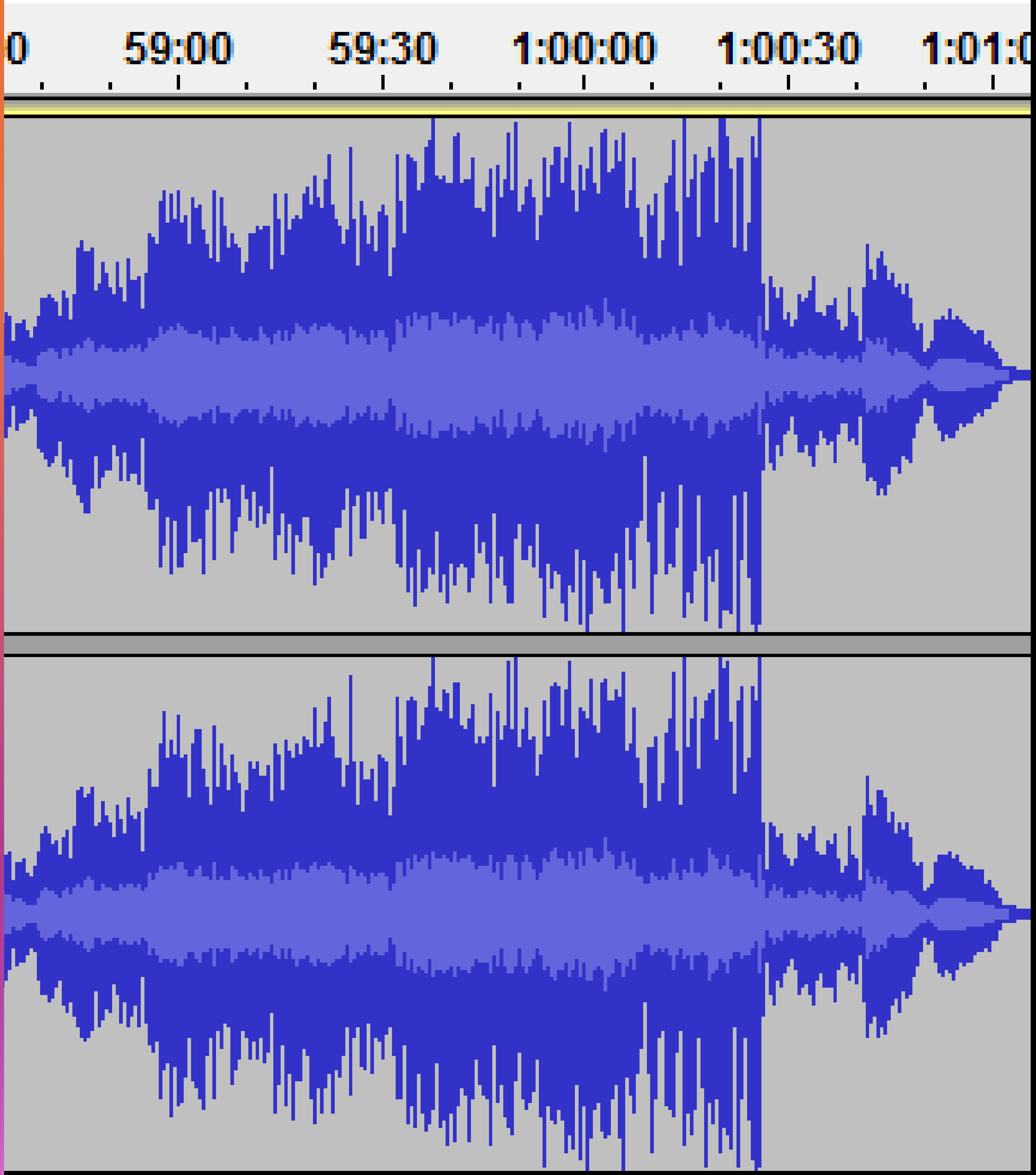
U – Upgrade Frequency Pre-emphasis

S – Spectral Noise Reduction

T – Tune Adversarial Speaker Regularization

I – Integrate Additive Noise

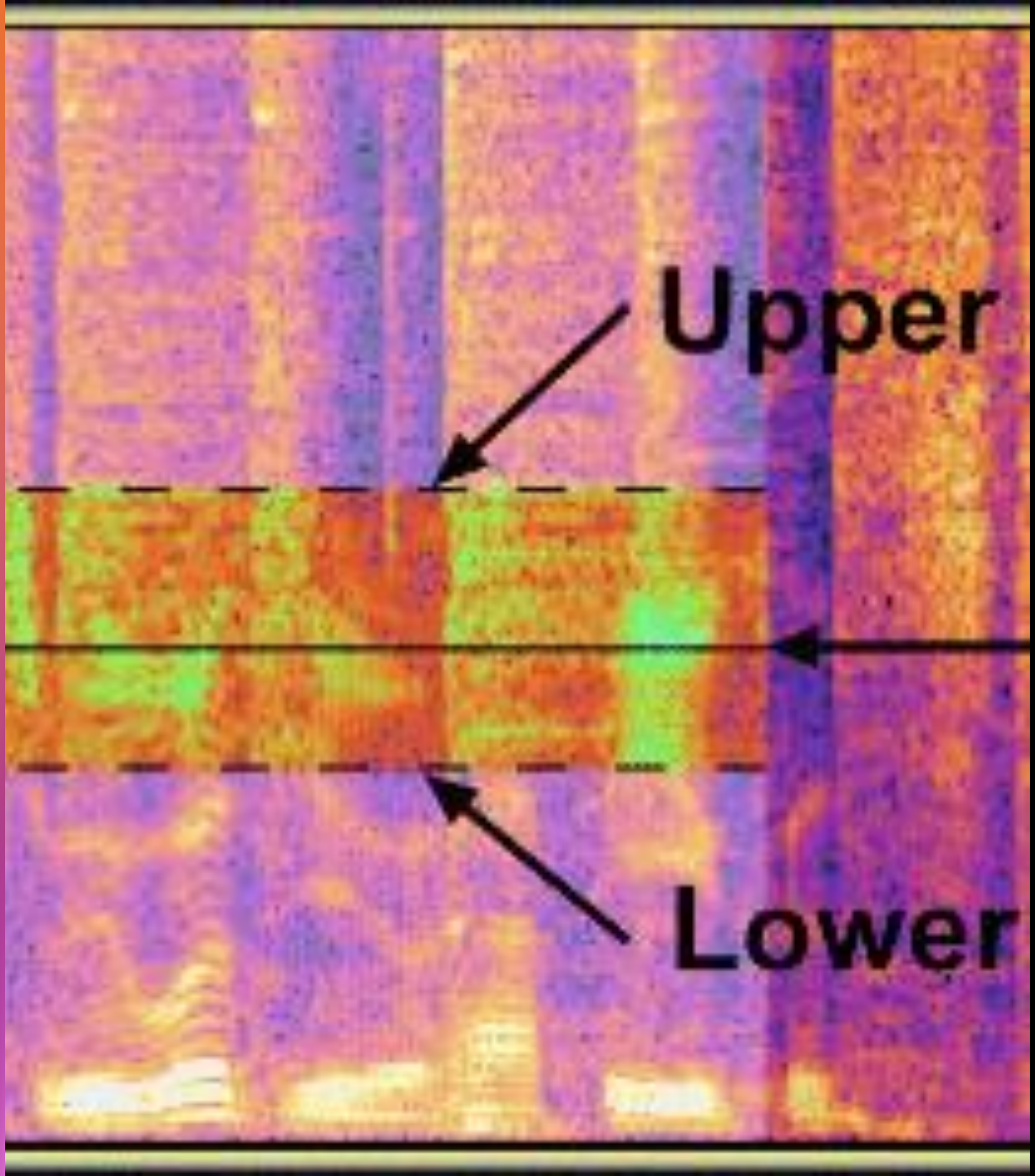
C – Create Model-Agnostic Approach



A.C.O.U.S.T.I.C. - A

A = Adjust Silence Intervals

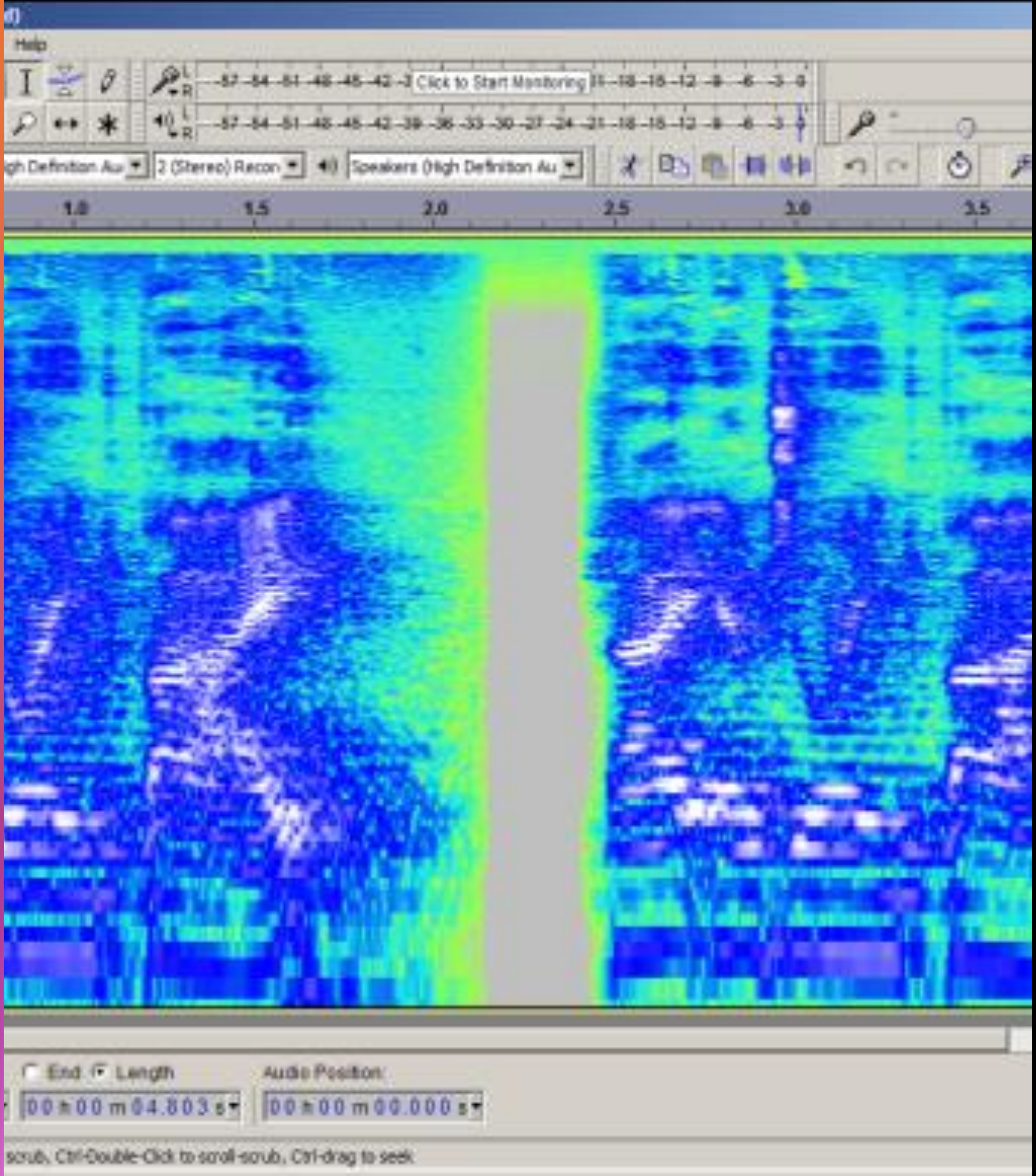
- Replacement of Leading & Trailing Silence
- Elimination of Inter-word Redundant Silence



A.C.O.U.S.T.I.C. - C

C = Center Spectrum Boosting

- Importance of Frequency Manipulation
- Calibration that creates a balance between amplification and authenticity

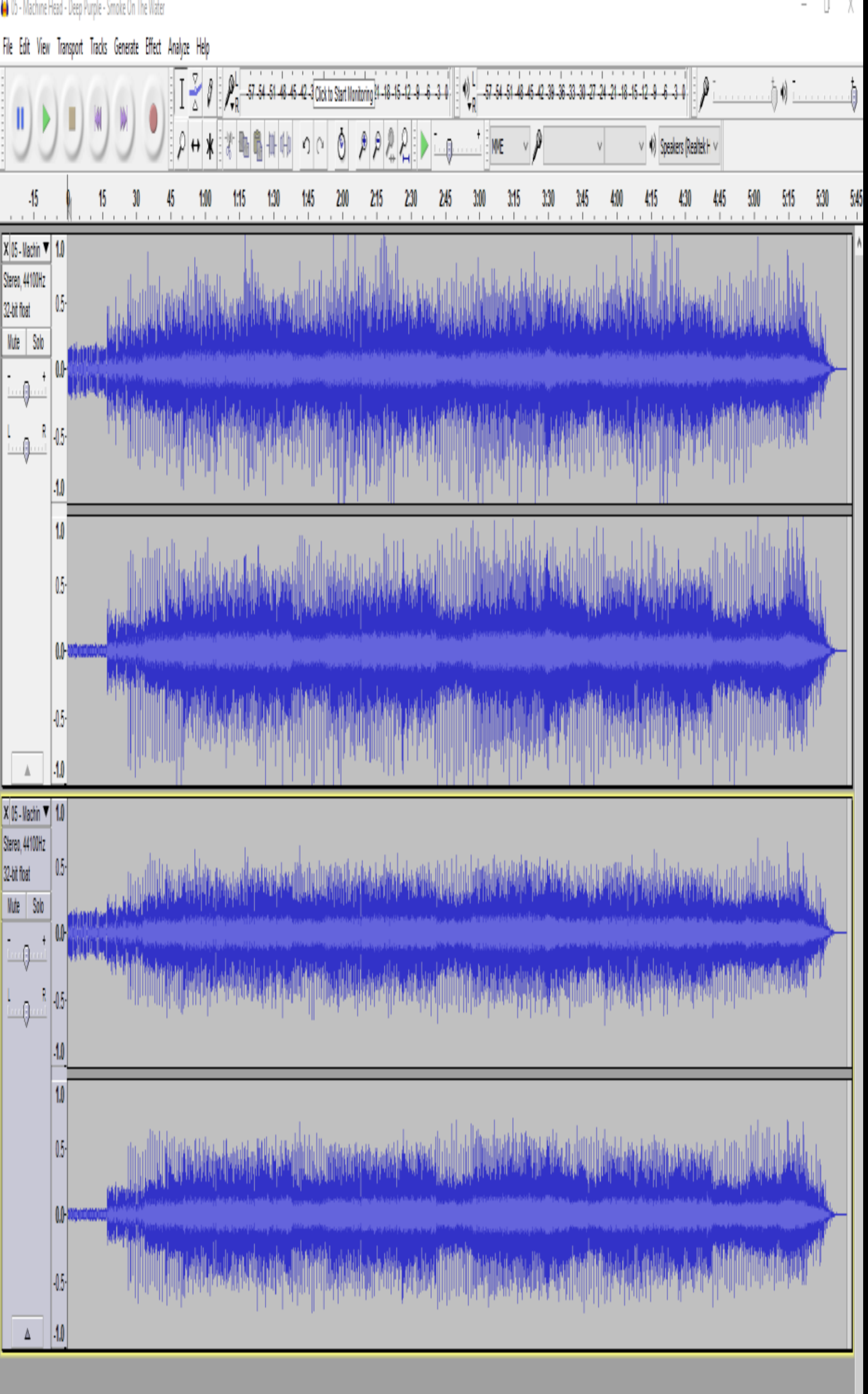


scrub, Ctrl-Double-Click to scroll-scrub, Ctrl-drag to seek

A.C.O.U.S.T.I.C. - O

O =Optimize Echo Simulation

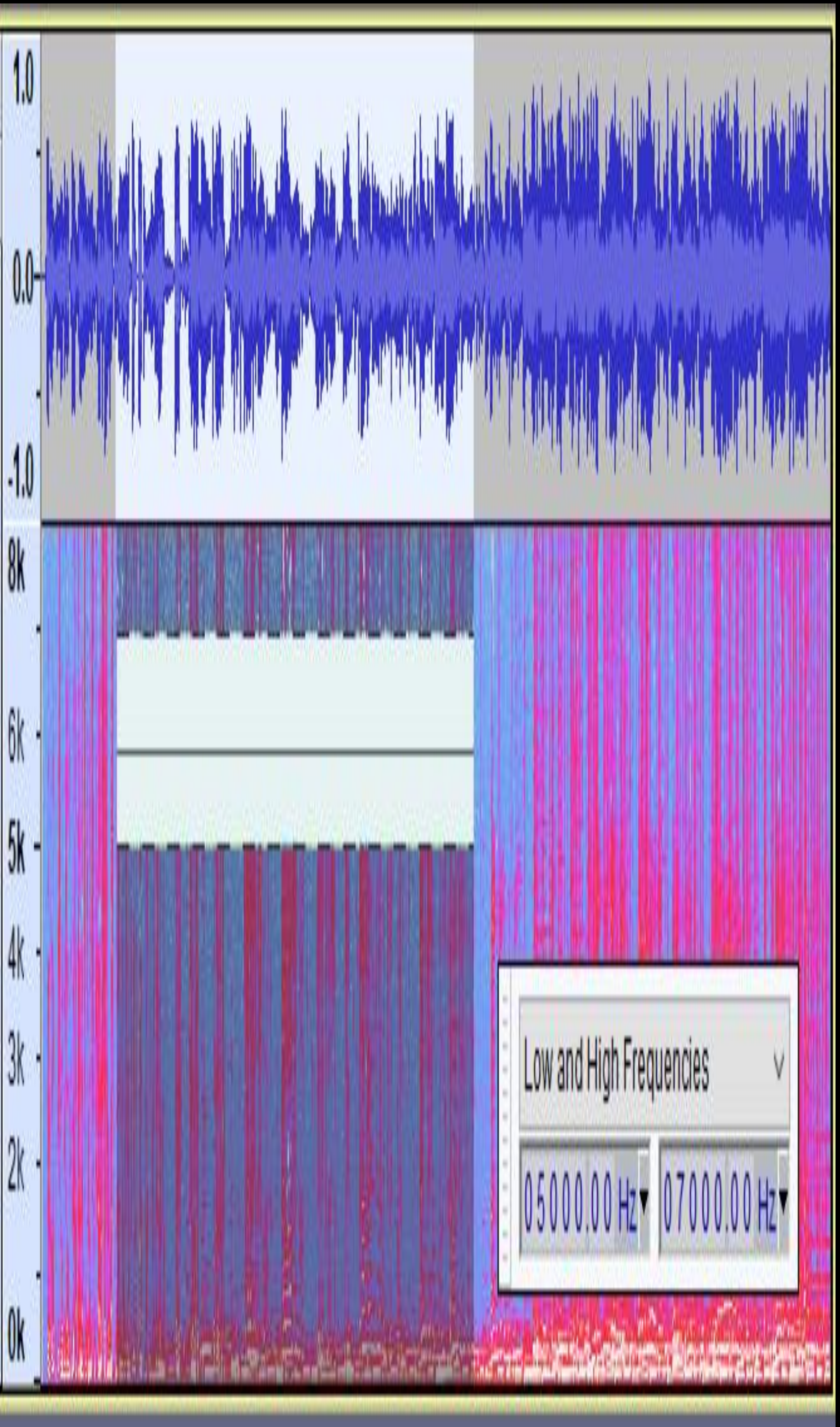
- Role of Echo in Speech Authenticity
- Simulating Echoes in Synthetic Speech


05 - Machine Head - Deep Purple - Smoke On The Water
File Edit View Transport Tracks Generate Effect Analyze Help
Click to Start Monitoring
Speakers (Realtek
05 - Machin
Stereo, 44100Hz
32-bit float
Mute Solo
1.0
0.5
0.0
-0.5
-1.0

A.C.O.U.S.T.I.C. - U

U = Upgrade Frequency Pre-emphasis

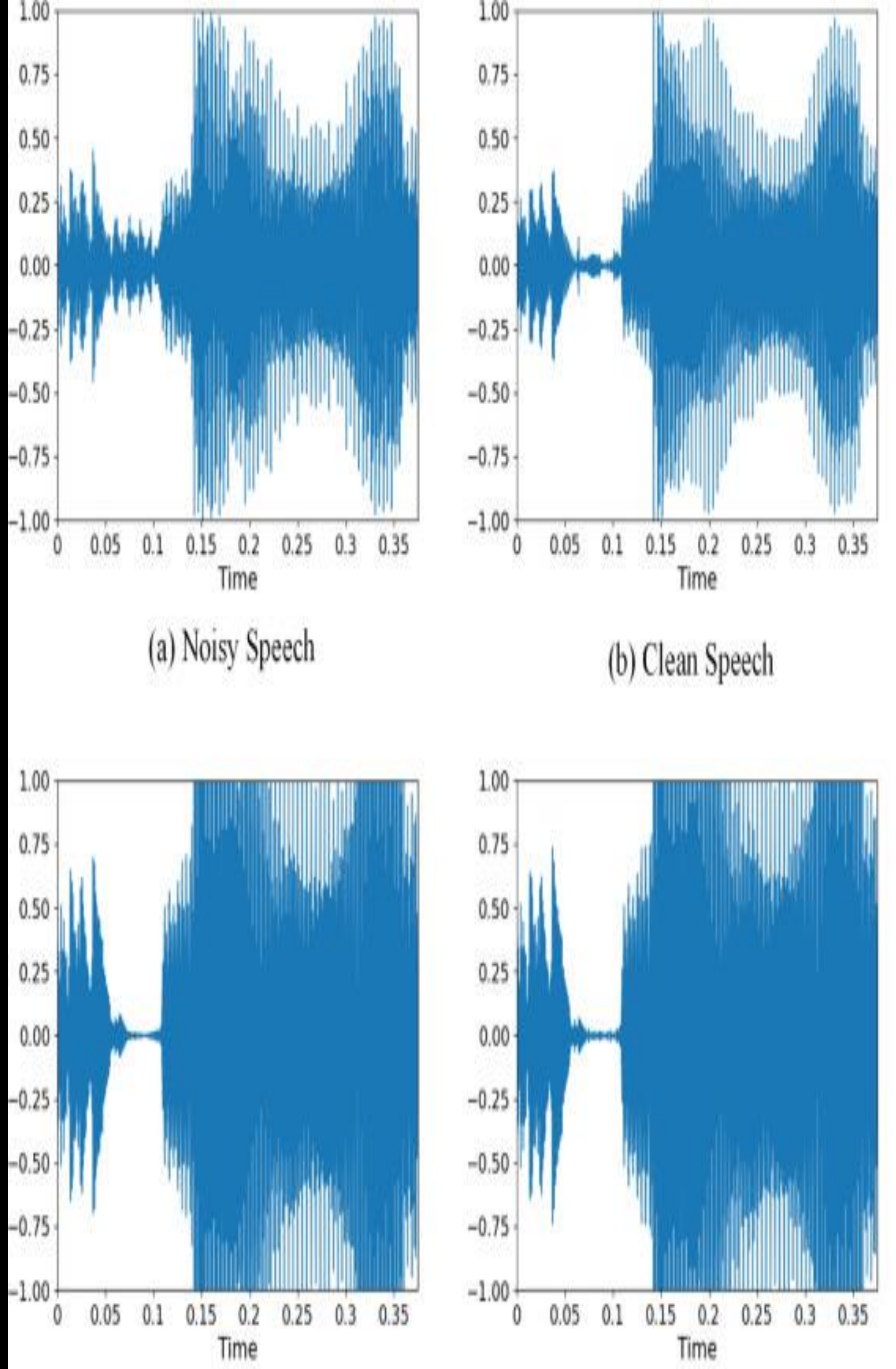
- The Role of Frequency Pre-emphasis
- Enhancing Clarity and Intelligibility
- Suppressing Unwanted Frequencies
- Maintaining Spectral Balance



A.C.O.U.S.T.I.C. - S

S = Spectral Noise Reduction

- Role of Spectral Noise Reduction
- Spectral Gating
- Dynamic Noise Filtering
- Additive Noise Incorporation



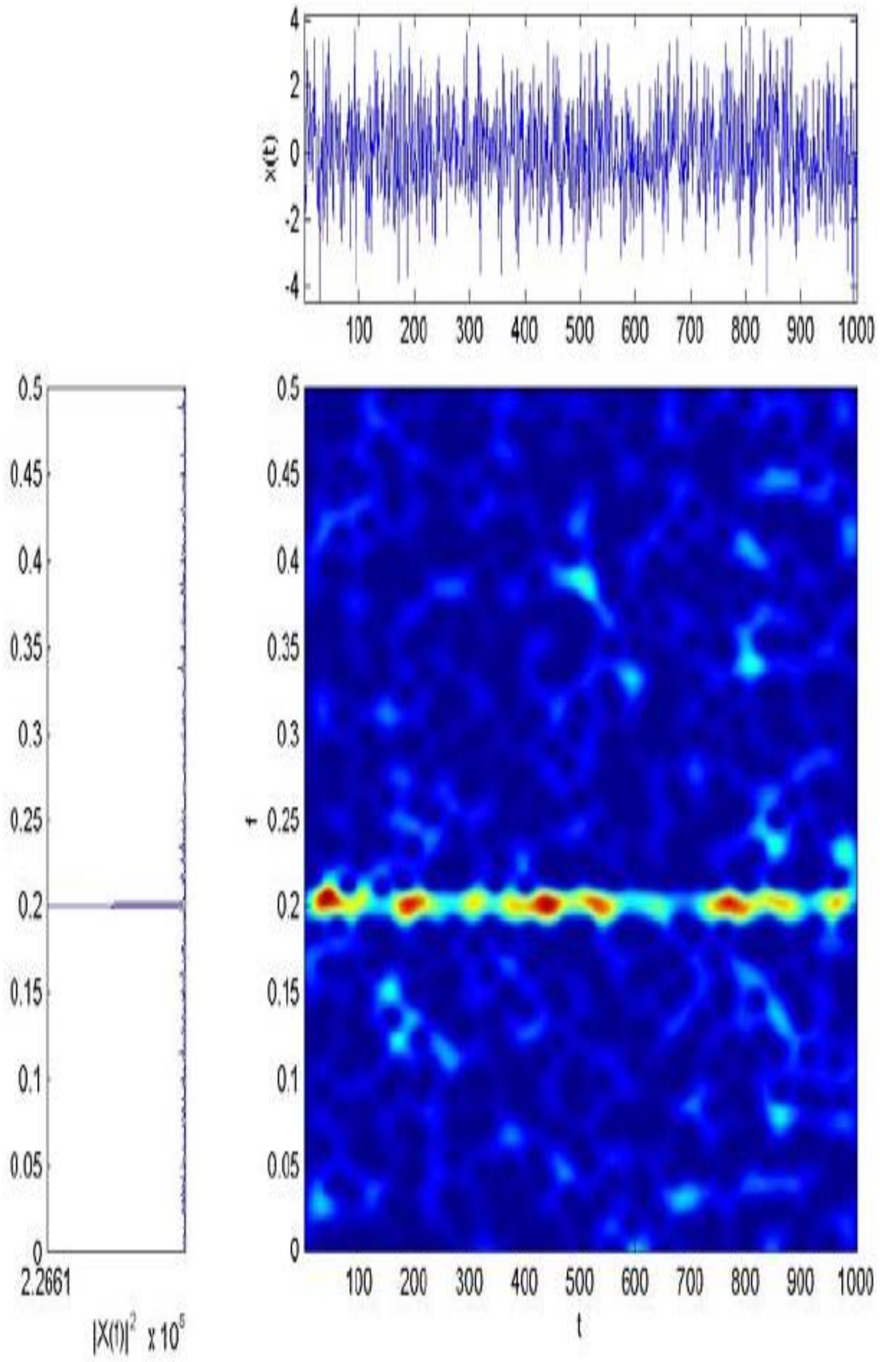
(a) Noisy Speech

(b) Clean Speech

A.C.O.U.S.T.I.C. - T

T = Tune Adversarial Speaker Regularization

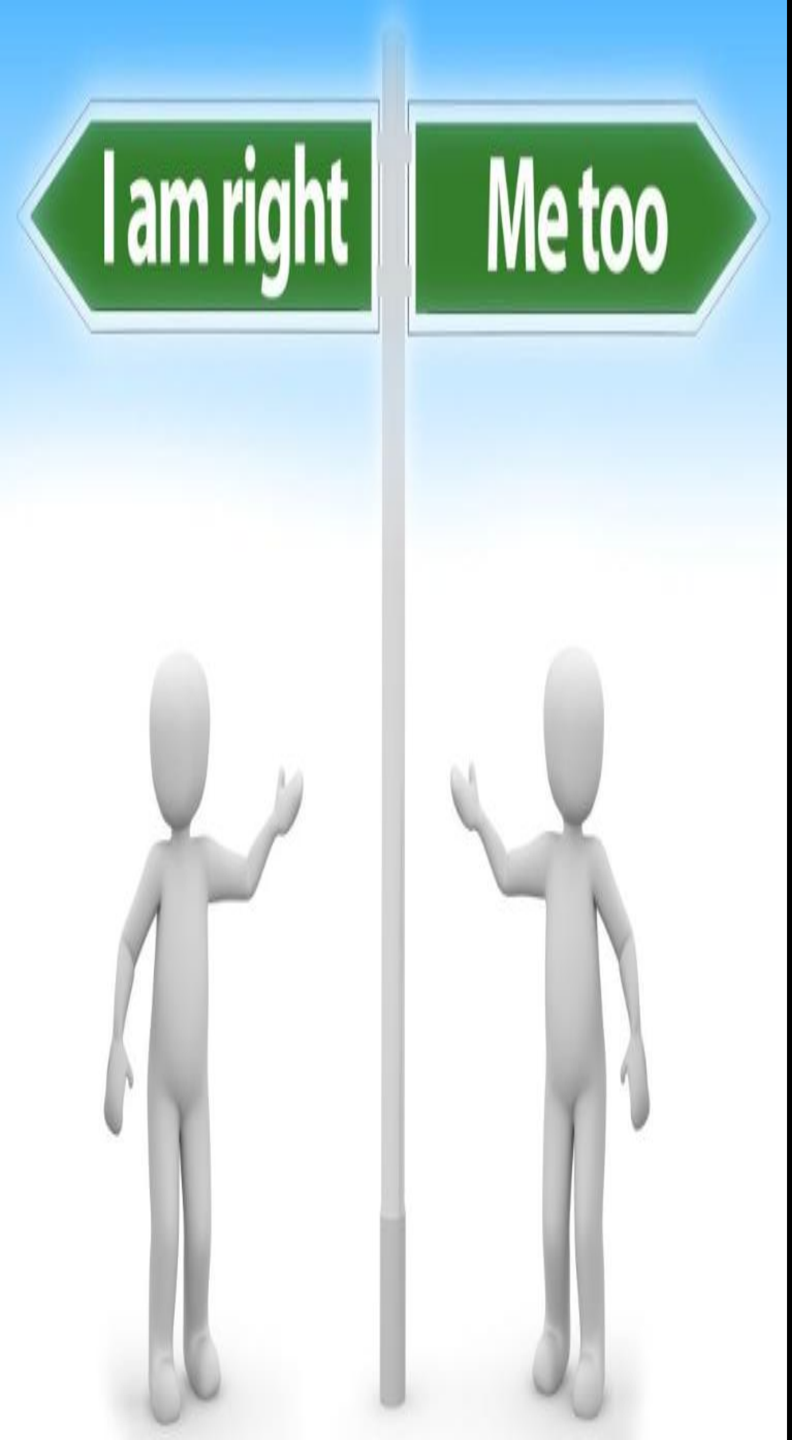
- Role of Adversarial Speaker Regularization
- Adversarial Training
- Speaker Voiceprint Engraving
- Regularization Techniques



A.C.O.U.S.T.I.C. - I

I = Integrate Additive Noise

- Role of Additive Noise
- Noise Source Selection
- Dynamic Noise Integration
- Temporal & Spatial Consistency



A.C.O.U.S.T.I.C. - C

C =Create Model-Agnostic Approach

- Role of a Model-Agnostic Approach
- Generalizable Techniques
- Flexible Frameworks
- Continuous Learning

Conclusion

- Enhanced Security Against Fraud
- Improving Detection Algorithms
- Preserving User Trust
- Adaptation to Evolving Threats

Images

https://manual.audacityteam.org/man/spectral_selection.html

https://forum.audacityteam.org/t/echo-reduction-and-optimizing/43511

https://manual.audacityteam.org/man/spectral_selection.html

https://forum.audacityteam.org/t/removing-silent-spaces-solved/37702

https://audiophilestyle.com/forums/topic/38914-pre-emphasis/

https://www.sciencedirect.com/science/article/abs/pii/S0885230824000019

https://www.researchgate.net/figure/Spectrogram-of-a-sinusoid-with-additive-white-Gaussian-noise_fig3_255728914

https://blogs.sas.com/content/subconsciousmusings/2020/05/07/model-agnostic-interpretability/

Above all, thank you to the Immaculata, Charlotte, Michael, Therese, and Gianna for always rooting me on.