**Detecting Spoofing Attacks Using VGG and SincNet: BUT-Omilia Submission to ASVspoof 2019 Challenge**

1. **Key Technical Innovations:**

* Fusion of two VGG Networks for Physical Access (PA) attacks.
* One network trained on single-channel features.
* Another trained on two-channel features.
* Fusion of VGG and SincNet architectures forLogical Access (LA) attacks.
* SincNet uses parameterized sinc function**s** to model the filters in the first convolutional layer, enhancing performance on raw audio.
* Separate systems were designed specifically for PA and LA, acknowledging their different characteristics.

1. **Reported Performance metrics:**

* Physical Access (PA): Achieved 86% relative improvement over the official baseline.
* Logical Access (LA):Good performance on known spoofing attacks. Challenges in generalization to unseen attacks not available during training.

1. **Promising aspects of this approach:**

* It combines VGG networks (spectrogram/CQT analysis) with SincNet (raw waveform processing) to detect both synthetic speech (TTS/VC) and replay attacks
* It is trained on 1-second audio segments, enabling near real-time detection
* End-to-end classification without external backends reduces latency
* Mean/variance normalization and energy-based voice activity detection adapt to varying conversation dynamics.

1. **Potential Limitations or challenges:**

* Achieved 0.2% EER on seen attacks but performance degrades when on unseen attacks.
* Fusion architecture increases resource requirements vs single models.
* SincNet 251-sample filters need GPU acceleration for real-time use.