**Aatithaya Paliwal**

**Part 1: Research & Selection**

1. **Reviewed GitHub Repository**
   * Examined the curated collection of papers and resources on audio deepfake detection from [this repository](https://github.com/media-sec-lab/Audio-Deepfake-Detection).
2. **Identified 3 Promising Models/Forgery Detection Approaches**

**(1) Waveform-Based CNN Approach**

* + **Key Innovation:** Uses a convolutional neural network (CNN) directly on raw audio waveforms for feature extraction.
  + **Performance Metrics:** Achieved over 90% accuracy in some benchmark datasets.
  + **Why This Approach?**
    - Does not require complex handcrafted features.
    - Potential for real-time detection due to CNN’s efficiency.
  + **Limitations:**
    - Performance may degrade on unseen deepfake techniques.
    - Computational cost can be high for long audio clips.

**(2) Spectrogram-Based Transformer Model**

* + **Key Innovation:** Converts audio into spectrograms and applies transformer-based architectures (e.g., ViT or Audio Spectrogram Transformer).
  + **Performance Metrics:** Reports indicate an AUC score above 0.95 on ASVspoof datasets.
  + **Why This Approach?**
    - Leverages attention mechanisms to capture subtle differences in frequency patterns.
    - Strong generalization across different deepfake techniques.
  + **Limitations:**
    - Transformer models can be resource-intensive.
    - Requires significant pretraining for optimal results.

**(3) LPC & MFCC Feature-Based Hybrid Model (RNN + CNN)**

* + **Key Innovation:** Combines traditional speech features (Linear Predictive Coding, MFCC) with deep learning (CNN for feature extraction + RNN for temporal dependencies).
  + **Performance Metrics:** Accuracy of ~88% on VoxCeleb-based datasets.
  + **Why This Approach?**
    - Uses well-established speech processing techniques, ensuring interpretability.
    - More lightweight than transformers, making it viable for deployment.
  + **Limitations:**
    - May require domain-specific tuning.
    - Performance is dependent on dataset quality.

1. **Final Thoughts**
   * These three approaches provide a mix of interpretability, efficiency, and accuracy.
   * Further testing is needed to determine the best fit for real-world scenarios.