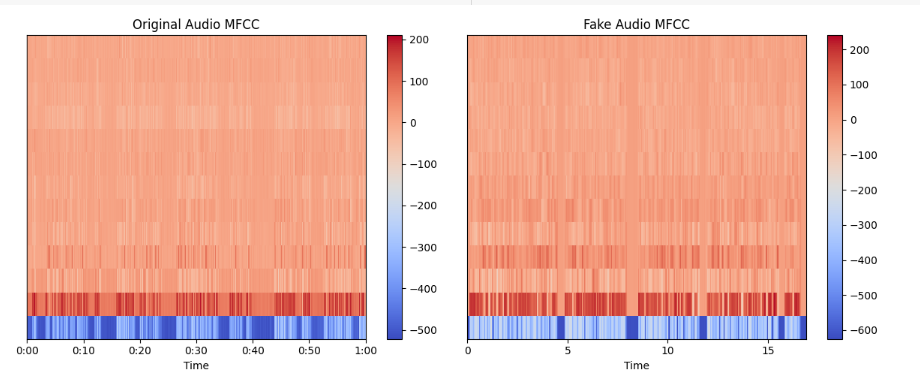
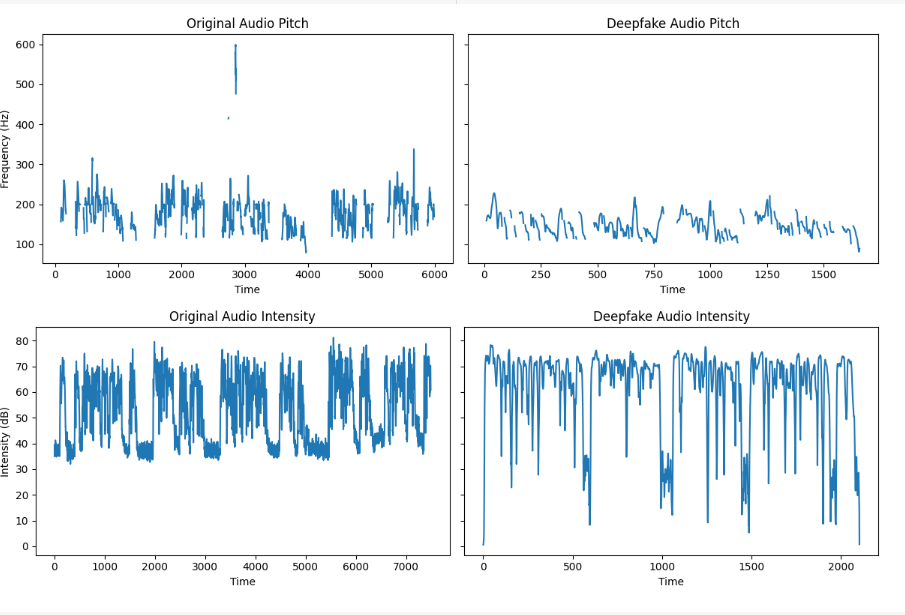
**ANALYSING THE DIFFERENT VOCAL FEATURES OF THE SAME SPEECH DATA**

**1.MFCC**



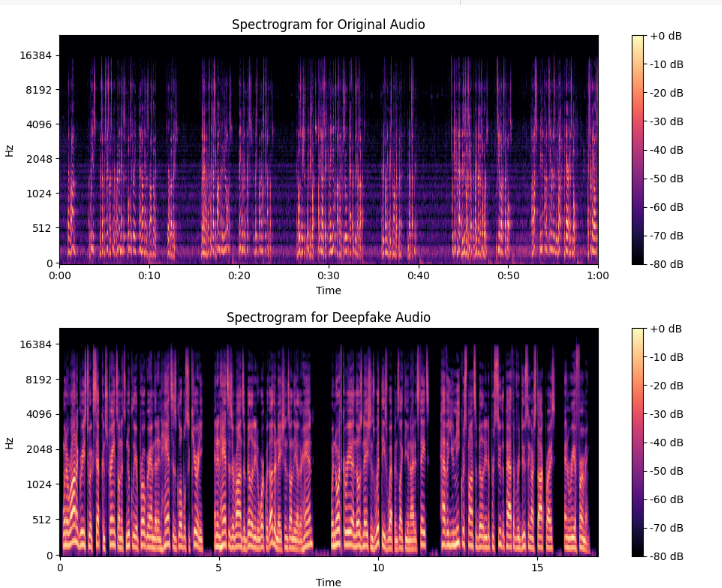
This suggests that the 'Original Audio' has a smaller variation in its spectral features. The 'Fake Audio' shows a broader range of values, indicating a wider variety of frequencies present or a difference in the processing technique.

**2.PROSODIC FEATURES:**



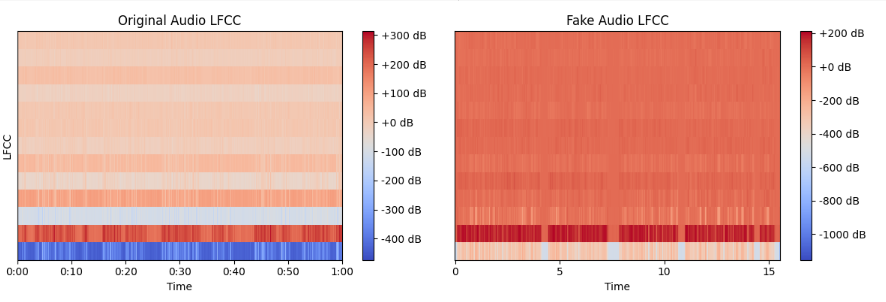
The deepfake audio shows a greater consistency in both pitch and intensity, lacking the natural variability seen in the original audio. This could suggest that the deepfake audio has been processed to have a more uniform and less dynamic character. The original audio displays a greater range of variability in both pitch and intensity, which is more characteristic of natural speech or audio.

**3.SPECTROGRAM**



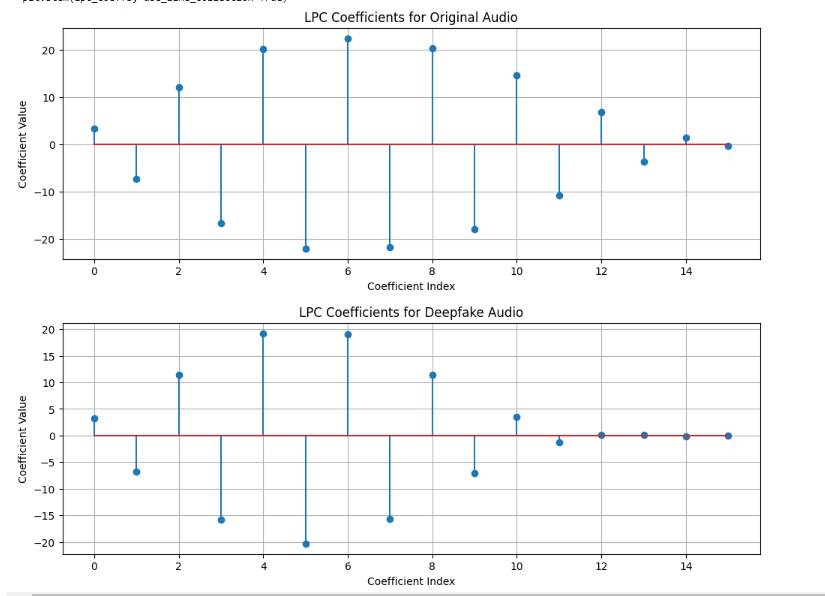
The spectrograms depict frequency over time with color intensity indicating power (decibels). The original audio has sharp, defined lines, suggesting clear and consistent sound patterns typical of genuine recordings. Deepfake audio displays more variability and less definition in frequency lines, likely from synthesis processes.

**4.LFCC**



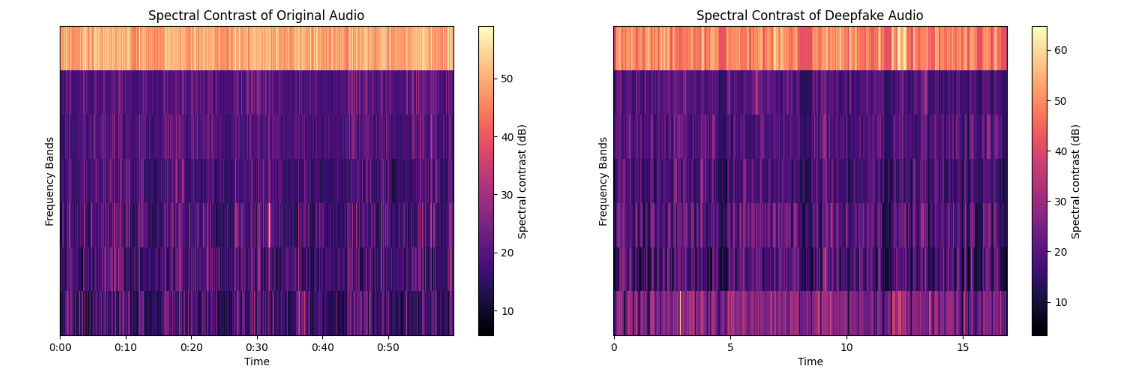
The original shows shorter duration and less variation in the cepstral coefficients, suggesting a cleaner and possibly less complex sound. The fake audio spans a longer timeframe with greater variance in LFCC values, indicating a potentially more complex or manipulated sound. Visually, the original's uniformity contrasts with the fake's broader and darker range of LFCCs, which could reflect alterations introduced in the fake audio synthesis process.

**5.LPC**



The LPC (Linear Predictive Coding) Coefficients for the original audio exhibit smaller variations around the centerline, indicating a more stable and possibly natural sound structure. The deepfake audio's LPC coefficients show much larger variations and greater extremities in values, suggesting a wider range of articulation which could be indicative of synthesized or modified audio.

**6.SPECTRAL CONTRAST:**

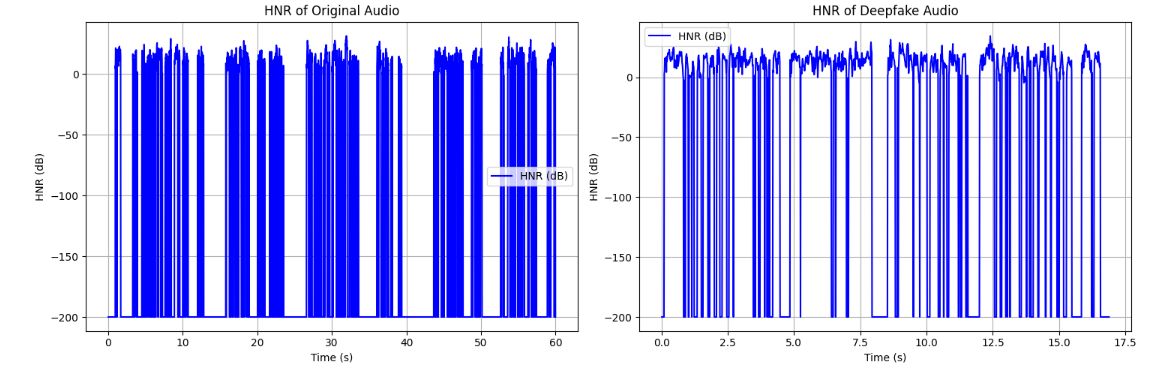


Spectral contrast measures the difference in decibel levels between peaks and valleys within a frequency range, highlighting texture and timbral aspects.

The original audio has a narrower time span with less pronounced contrasts, indicating a more consistent and possibly less complex sound. The spectral contrast appears smoother, which is common in unaltered audio signals.

In contrast, the deepfake audio extends over a longer period and displays more pronounced spectral contrasts, with clearer separations between frequency bands.

**7.HNR:**



In the "Original Audio," there's a consistent pattern of sharp peaks and valleys, suggesting a natural and periodic vocal structure with harmonic content. The peaks are regular and the HNR levels are quite stable, which is typical of a clean, unprocessed voice.

The "Deepfake Audio," while also showing peaks, has more variation in the HNR levels between these peaks. The peaks don’t reach as low as the original, indicating a higher noise floor or less distinct harmonic content, which could be a sign of artificial processing or synthesis.

CORRELATION RESULTS:

