Step	Original Script	Updated Script	Reason for Update / Improvement
Load Audio	librosa.load at 22.05 kHz	Same	No change, consistent resampling.
Trim Silence	librosa.effects.trim(y, top_db=20)	librosa.effects.split + concatenate, top_db=40	More aggressive silence trimming to remove low-level noise.
Amplitude Normalization	Normalize to [-1,1] via y / np.max(np.abs(y))	LUFS normalization using pyloudnorm to target -23 LUFS	LUFS ensures perceived loudness is consistent across all files; amplitude clipping avoided via np.clip.
Mel-Spectrogram	librosa.feature.melspectrogram with 80 mel bands	Same parameters	No change, standard Mel-spectrogram.
Log Scaling	librosa.power_to_db	Same	Perceptual scaling; preserves audio dynamics.
Standardization	(mel - mean) / std	Same, but applied consistently with added clipping check	DNN/CNN stability; ensures inputs are zero-mean, unit variance.
3-Channel Stacking	Not done	Added; stacked resized spectrogram into 3 channels	CNNs pretrained on RGB images (like ResNet) often require 3-

Step	Original Script	Updated Script	Reason for Update / Improvement
			channel input; improves transfer learning.
Image Size / PNG Generation	4×4 inch figure, axis off, saved 1-channel PNG	299×299 resized image with axes for visualization, 3-channel stack	Higher resolution helps CNNs but increases storage. Axis added for easier visualization.
Save Audio	sf.write normalized audio	Same, but after LUFS normalization and clipping prevention	Prevents extreme values/clipping.
Dataset Split	Train/Val/Test via train_test_split 60/20/20	Same	No change.
Logging / Errors	None	Added try-except blocks, tqdm progress, optional CSV/logs	Makes preprocessing robust and trackable.