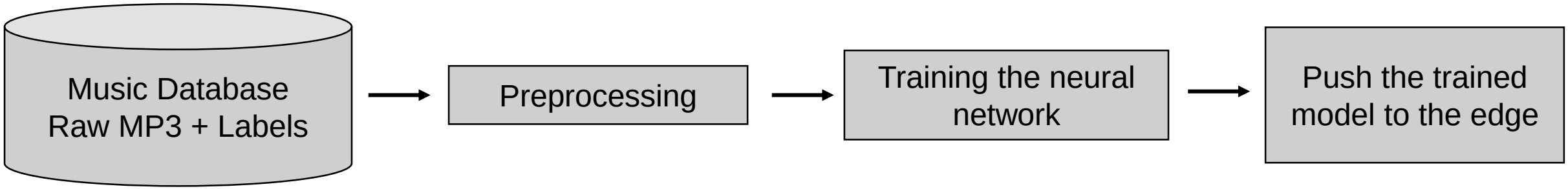
The pipeline



- Host the MP3 files and annotations in cloud storage.
- Dataset used: MagnaTagATune

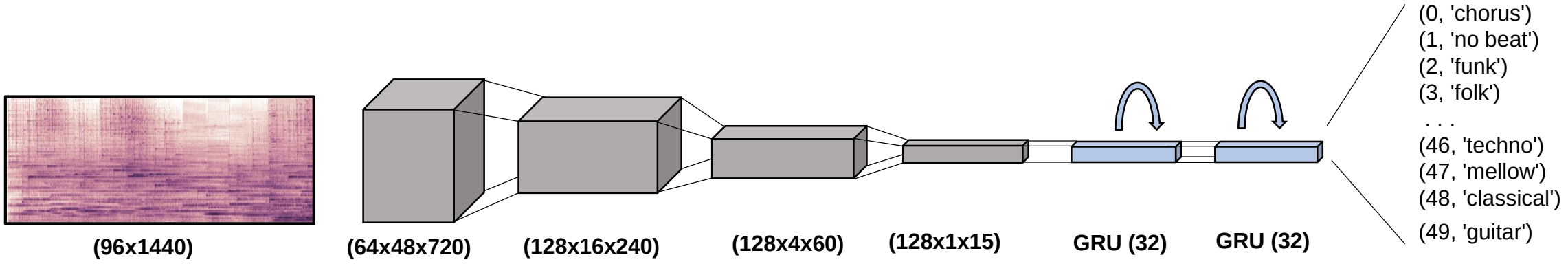
- Convert raw MP3 files to Mel-Spectograms.

- Teach the neural network to classify music based on its genre.

- Prediction phase on mobile or edge devices.

UTSA OPENCLOUD INSTITUTE

Deep Neural Net Architecture

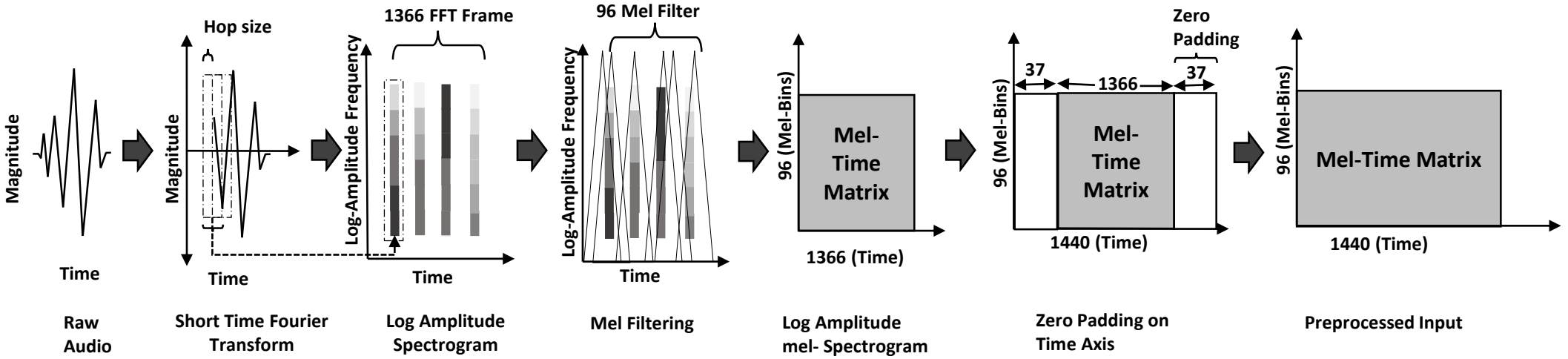


Log Amplitude Mel-Spectrogram

Convolutional Neural Network (CNN) Stacks

Recurrent Neural Network (RNN) Stacks

Pre-Processing



MP3 Input Music
Mel-Power Spectrum
mel power spectrogram
CRNN Model
Existence Percentage of Genres in a Music
Guitar 0.81
Classical 0.62
Mellow 0.55
Techno 0.46
String 0.42
Vocal 0.38
Electronic 0.34
Drum 0.32
Instrumnetal 0.29
Rock 0.25
Fast 0.22
Male 0.21
Beat 0.18
Female 0.18
Piano 0.15
Ambient 0.13
Violin 0.13
Synth 0.11
Indian 0.11
Singer 0.09
Opera 0.09
Harpsichord 0.06
Loud 0.06
Quite 0.06
Flute 0.06
Pop 0.06
Sitar 0.06
Solo 0.06
Choir 0.06
New Age 0.06
Dance 0.05
Strange 0.05
Harp 0.03
Metal 0.03
Cello 0.01
Jazz 0.00
Country 0.00
Eastern 0.00
Decomposed Music by Genres
String
Mellow
Guitar
Classical
Techno