

- LSTM Model (Long Short-Term Memory Model)
 - It can capture long term dependencies
 - Widely used
 - 3 GATES:
 - Input → defines the shape of the input sequence
 - Forget → decides how much info to keep
 - Output → what should become the hidden state
- GRU (Gated Recurrent Unit Model)
 - ONLY 2 GATES
 - Update → decides how much previous information to keep and how much new information to add
 - Reset → decides how much of the past information to forget
 - Compared to LSTM it has reduced complexity
 - Generally faster and more efficient than LSTM
- CRN (Convolution Recurrent Network Model)
 - Combines two layers to benefit from both local feature extraction and temporal sequence modeling
 - TWO LAYERS
 - Convolution → process and extract local features (effective in identifying patterns)
 - Recurrent → capture temporal dependencies and helps maintain continuity using RNN layers like LSTM and GRU
 - Can work near real time
- UNET
 - Called UNet because it has a “U shape”
 - Encoder → feature extractor that works by progressively reducing the spatial dimensions while increasing feature channels
 - Decoder → reconstructs feature map back to original, while enhancing
 - Treats the audio signal as a 2D image