



image sources:

Video:
<https://ai.meta.com/blog/v-jepa-yann-lecun-ai-model-video-joint-embedding-predictive-architecture/>
 (2024 May 12th)

Mel-spectrogram:
https://docs.nvidia.com/deeplearning/dali/user-guide/docs/examples/audio_processing/spectrogram.html
 (2024 May 12th)