Remove reverb from sound

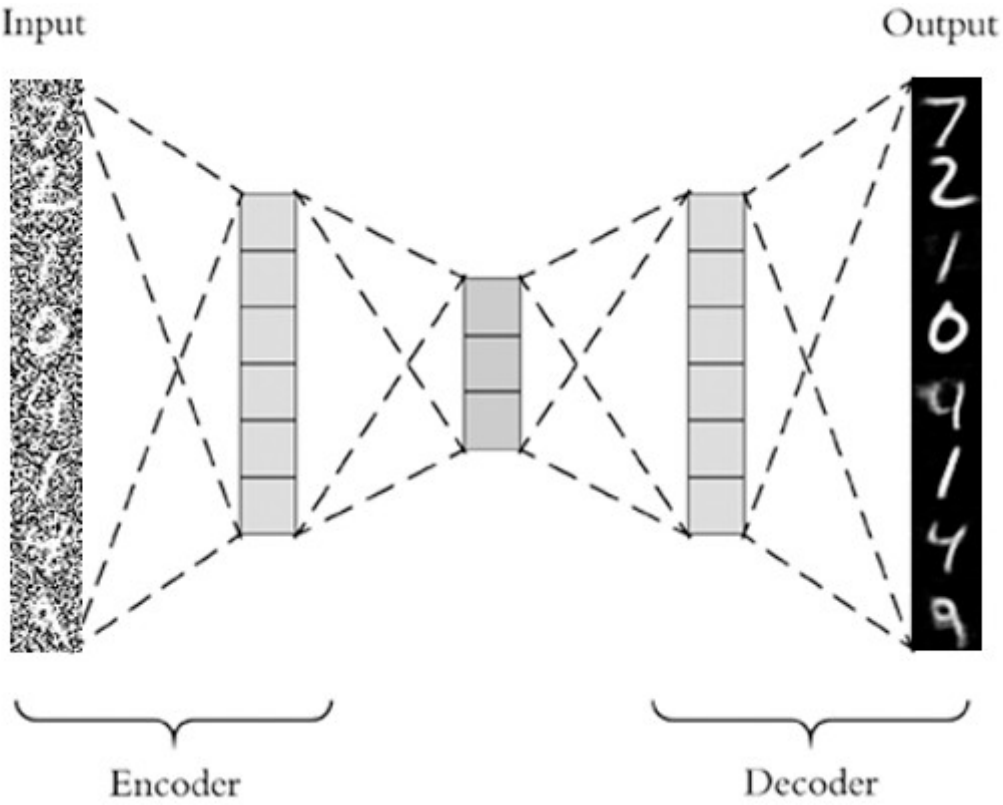
Possible strategies

- Deconvolve signal with a linear combination of known IR responses (Reverb Classifier)
- Predict IR with NN and deconvolve
- Predict dereverberated signal directly with AE

Autoencoder (AE)

- WaveNet
- Spectrograms
- Autoregressive

- Combinations

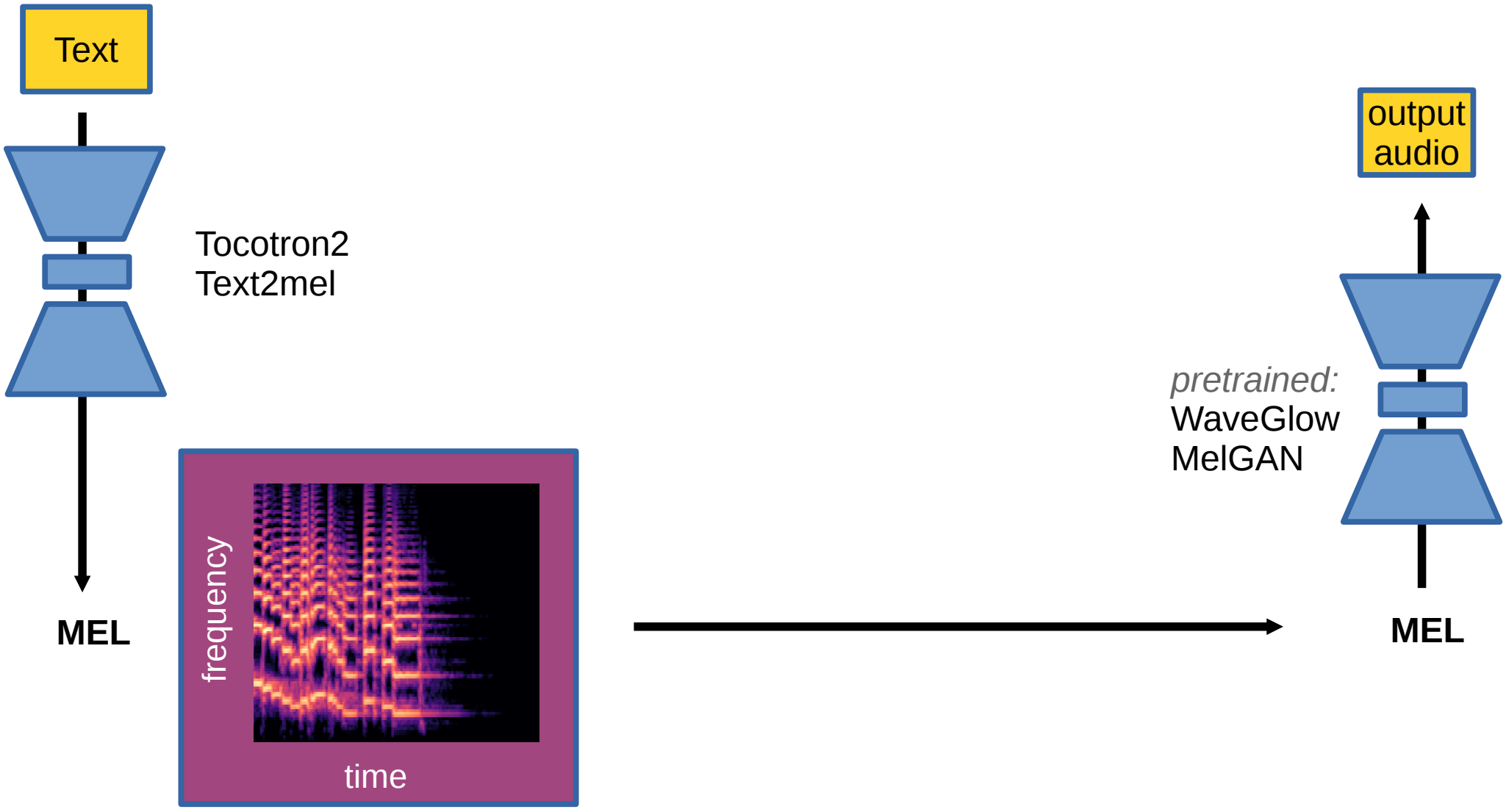


source: pyimagesearch.com/wp-content/uploads/2020/02/keras_denoising_autoencoder_header.png

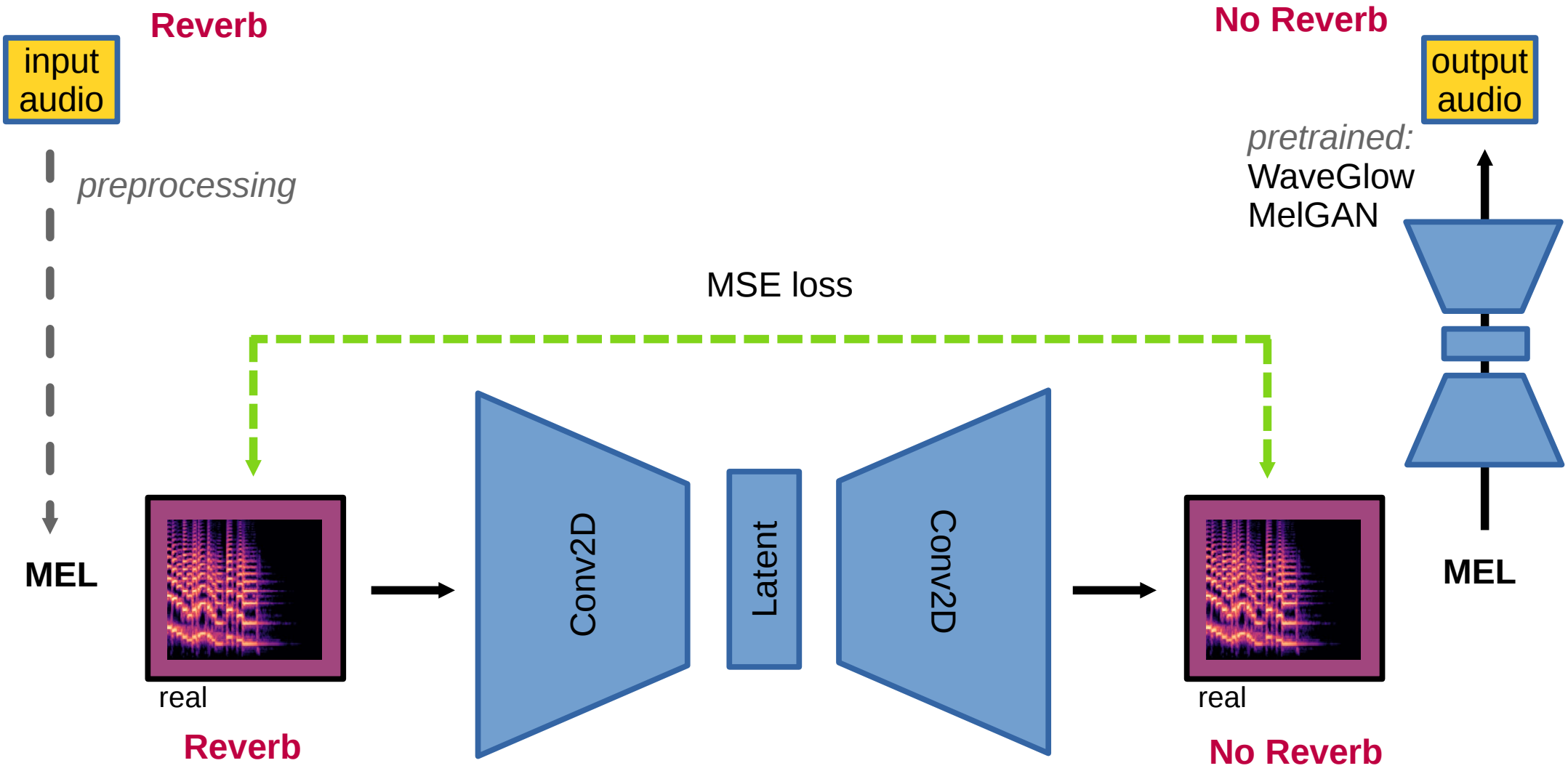
Possible strategies

- Work with MEL spectrograms
  - AE and WaveGlow (to slow for RT execution)
  - AE and WaveGAN (rather slow but possible)
- Work with audio signal
  - LSTM (bad quality)
  - Use WaveNet AE (slow convergence, no results, bad for reverb)
  - Train WaveGAN to omit reverb in translation (transfer learning)
- Work with STFT and complex numbers
  - Unknown territory!

SOTA Methods (Text-to-Speech)



SOTA Methods *Slow!*



MelGAN - transfer learning

Reverb

input audio

target

feature loss

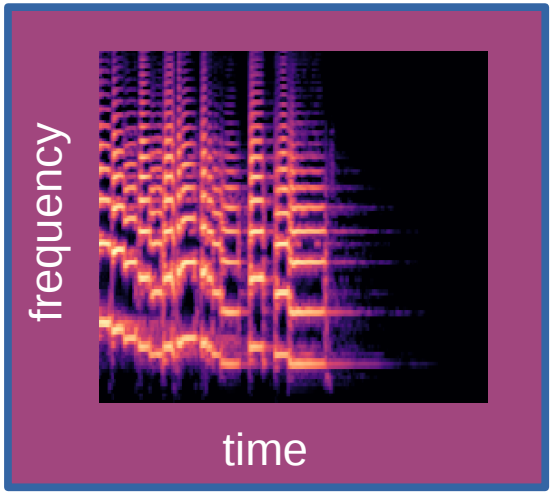
600 epoch

output audio

No Reverb

MelGAN
pretrained on speech

MEL

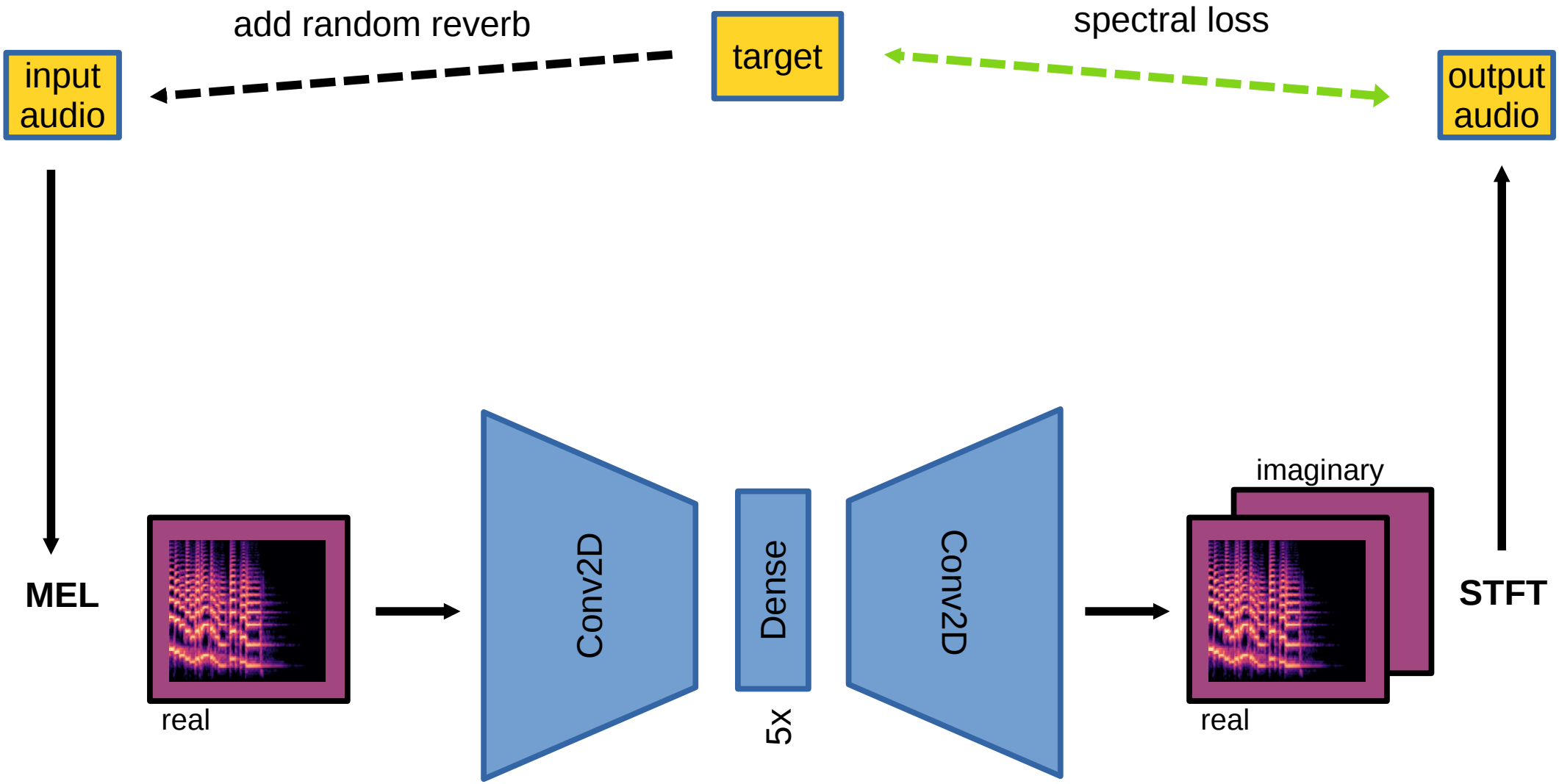


MEL

Unseen speaker

MelGAN trained on speech

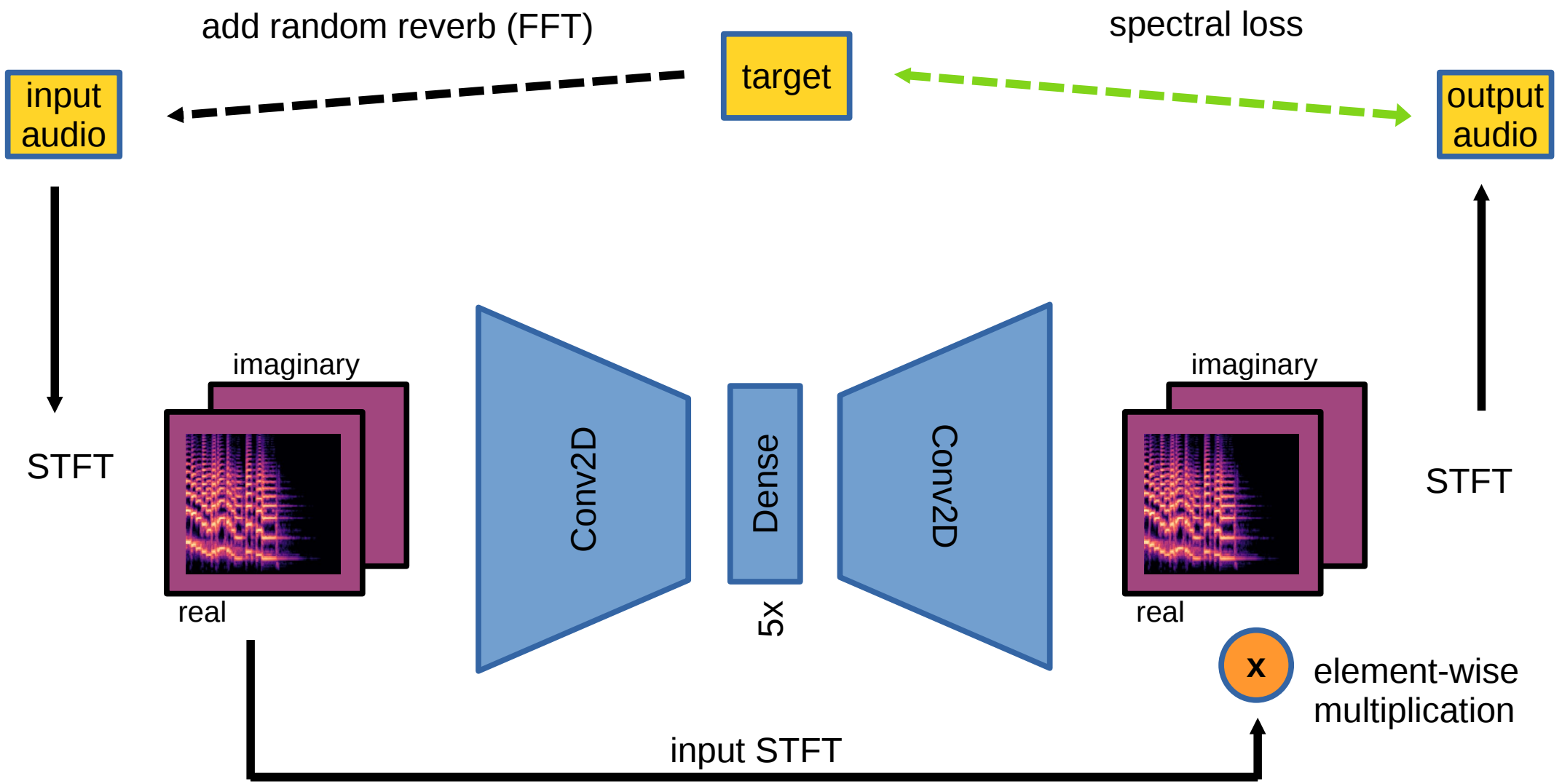
MEL-to-STFT (AE)



Unseen speaker

Mel-to-STFT test

STFT (AE)



real data | test | add room reverb | or a big reverb