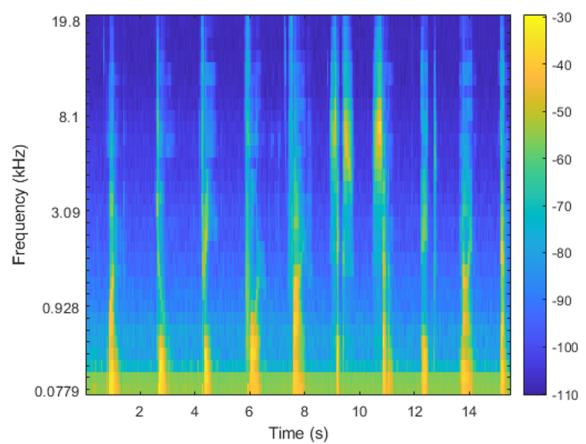
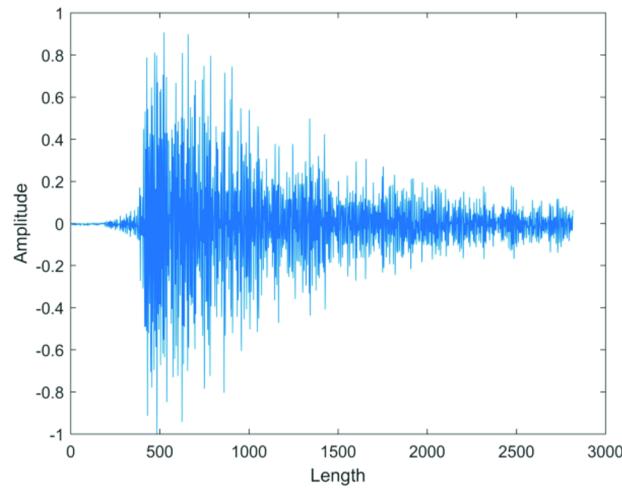


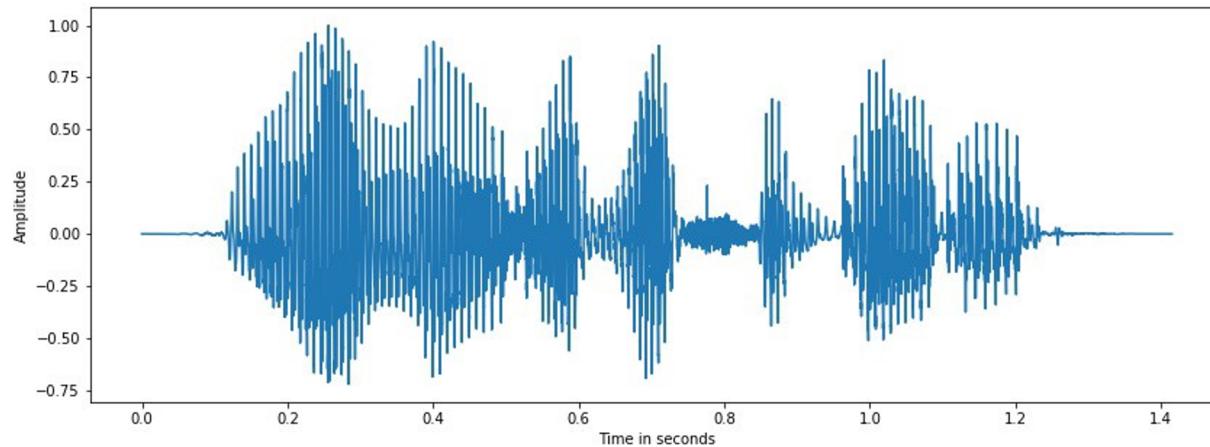
Seminar 10: VERY basics of sound processing



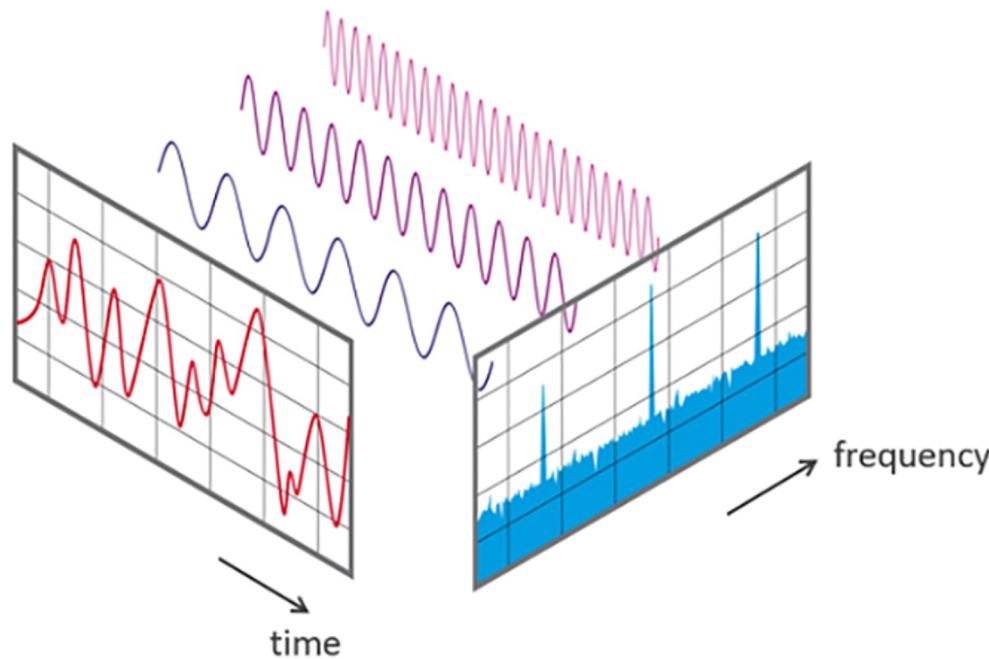
Actually, AI does
not exist. It's just
ML.



What is sound and how it look like as a signal How to process it? Any Ideas?

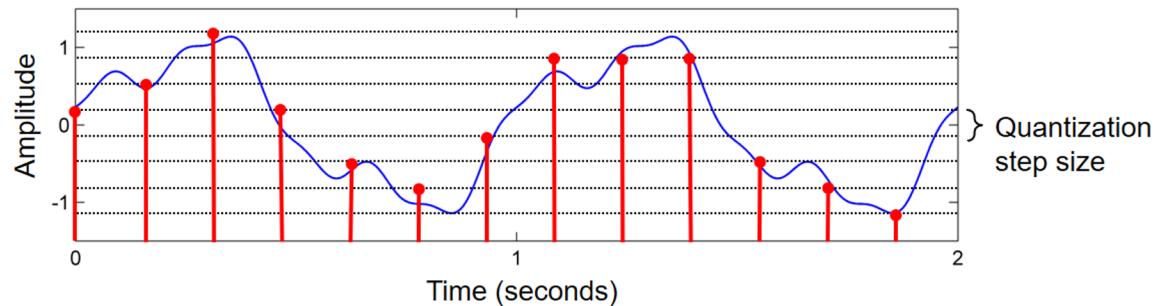
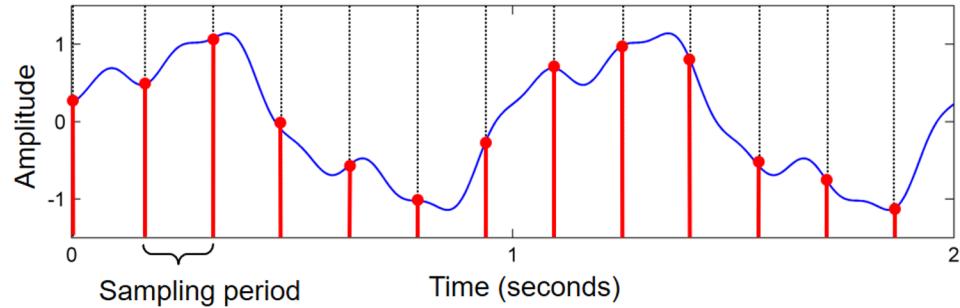


Fourier transform in a nutshell



Discretization

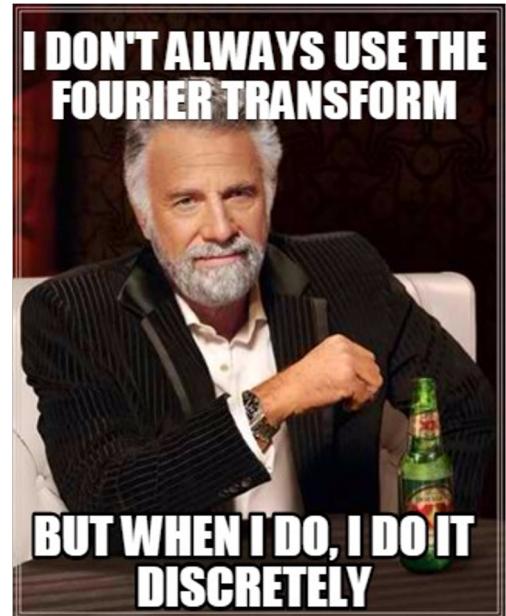
Figure 2.13 from [Müller, FMP, Springer 2015]



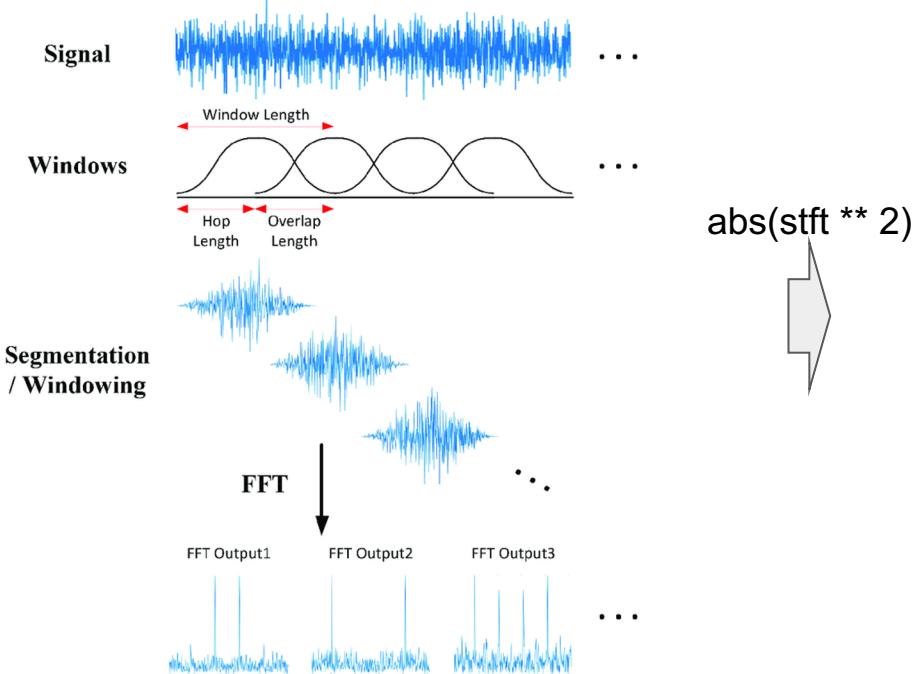
Discrete fourier transform

Time Duration		
Finite	Infinite	
Discrete FT (DFT) $X(k) = \sum_{n=0}^{N-1} x(n)e^{-j\omega_k n}$ $k = 0, 1, \dots, N - 1$	Discrete Time FT (DTFT) $X(\omega) = \sum_{n=-\infty}^{+\infty} x(n)e^{-j\omega n}$ $\omega \in (-\pi, +\pi)$	discr. time n
Fourier Series (FS) $X(k) = \int_0^P x(t)e^{-j\omega_k t} dt$ $k = -\infty, \dots, +\infty$	Fourier Transform (FT) $X(\omega) = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t} dt$ $\omega \in (-\infty, +\infty)$	cont. time t
discrete freq. k	continuous freq. ω	

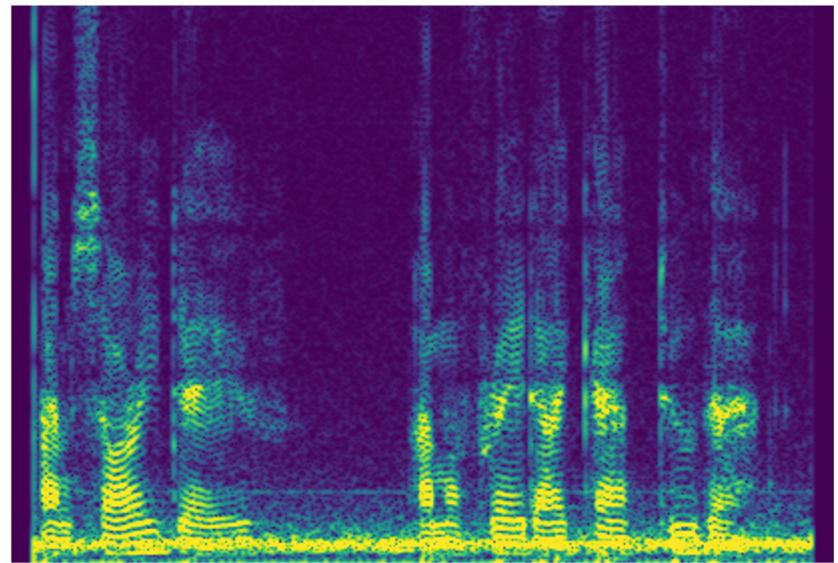
99% of DL programmers:



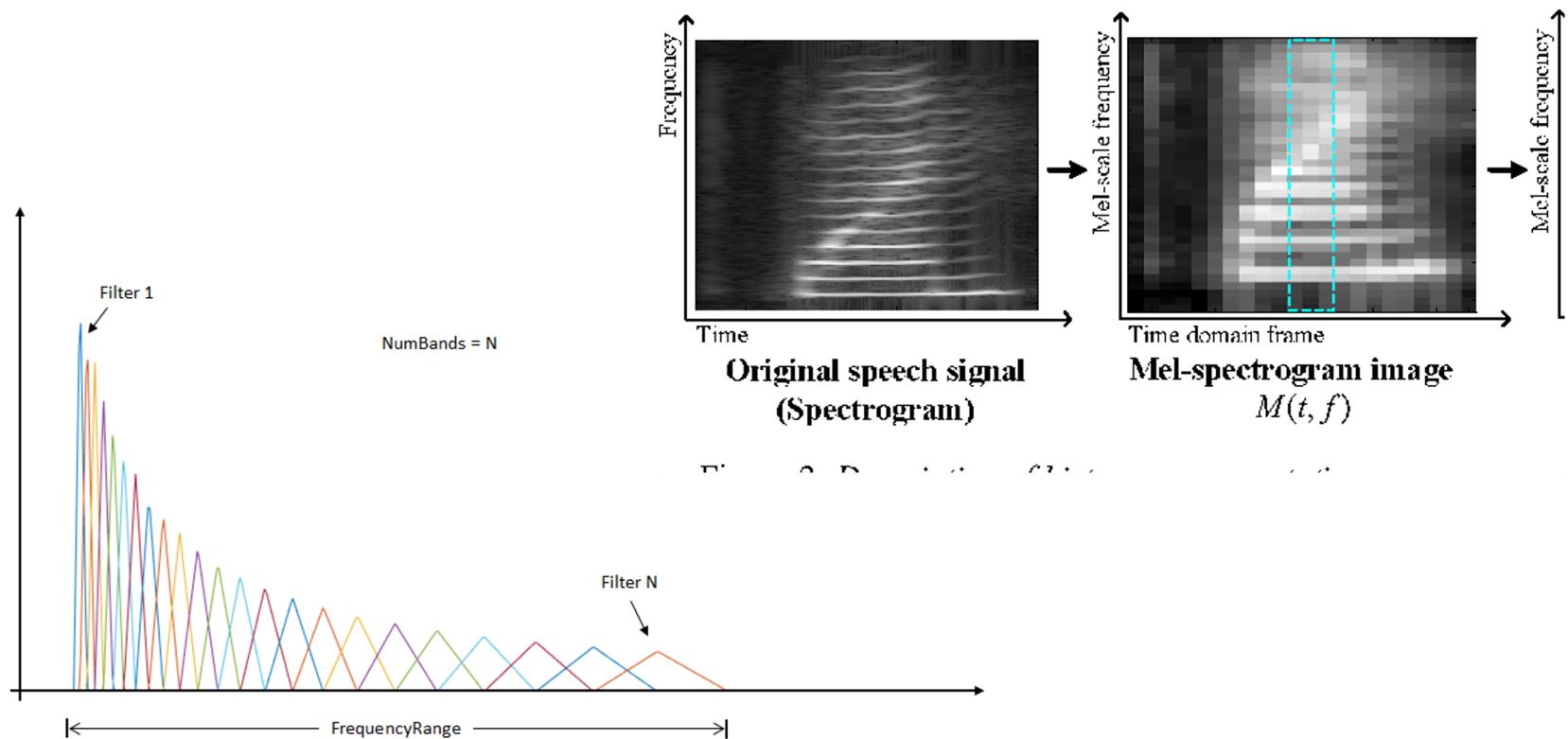
Short Time Fourier Transform in a nutshell



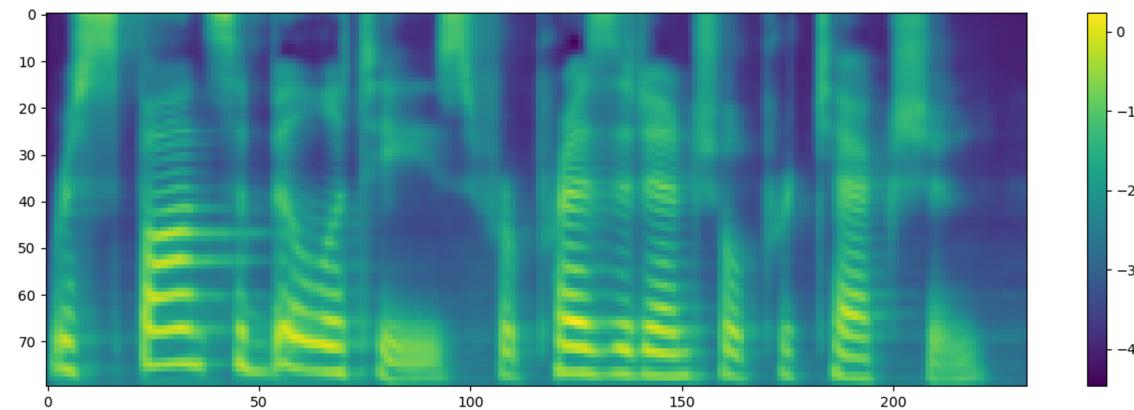
Spectrogram



Mel spectrogram

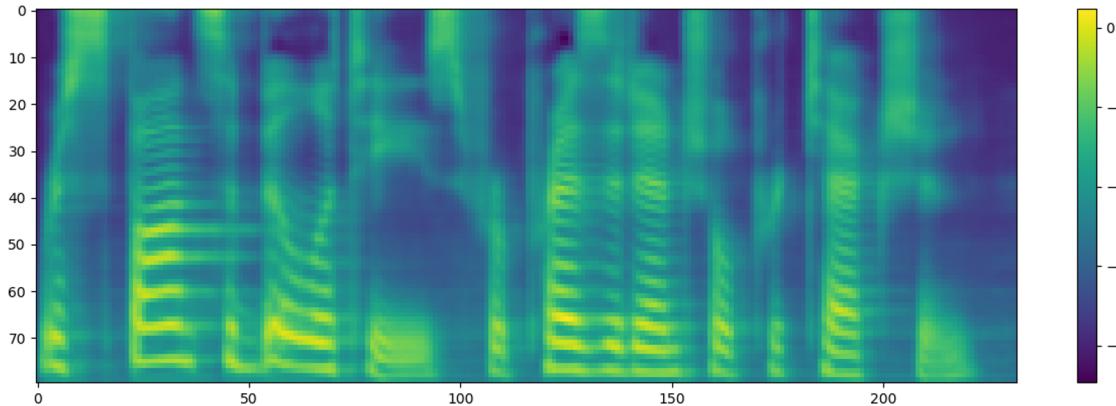


What do you see?

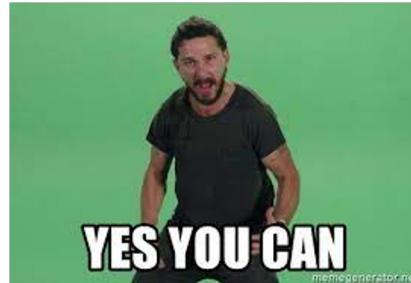


CNN for mel spectrograms

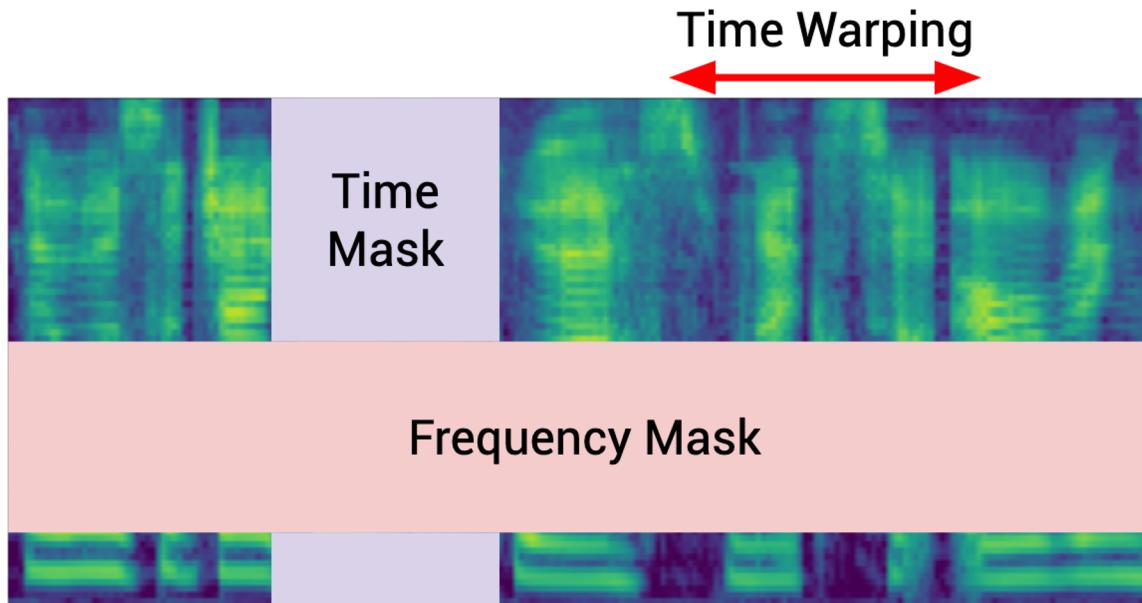
What do you see?



So we can use conv nets?



Augmentations



Other tasks and applications

