

This Matlab program implements a neural network based pitch tracker in noise described in:

Han K. and Wang D.L. (2014): “Neural network based pitch tracking in very noisy speech,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, pp. 2158-2168 ([http://web.cse.ohio-state.edu/~dwang/pubs\\_year.html](http://web.cse.ohio-state.edu/~dwang/pubs_year.html)).

The program was originally written by Kun Han, and the posted version was subsequently cleaned by Yuzhou Liu.

---

Please execute DNNPitch.m or RNNPitch.m in Matlab to get pitch estimates. For example:

1. DNN based pitch tracking:  
sig = audioread('examples/noisy0dB.wav');  
F0 = DNNPitch(sig);
2. RNN based pitch tracking:  
sig = audioread('examples/noisy0dB.wav');  
F0 = RNNPitch(sig);

Note that, a 16 kHz time domain signal is required for both functions. To run the RNN code, you have to access a GPU and the Parallel toolbox in Matlab.

Following is a description of the main steps in DNNPitch.m and RNNPitch.m:

1. Given the input signal, extract spectral domain features in each frame.
2. Employ neural networks (DNN/RNN) to compute the posterior probability of the pitch state for each frequency bin.
3. Use Viterbi decoding to connect pitch states, and use a moving average to generate final pitch contours.

To train the neural network models from scratch, we provide the following codes in ./train folder:

gen\_pegfacea.m: Generate features for a noisy sentence

gen\_praatfea.m: Generate desired outputs for a clean sentence

trainNN: Given features and desired outputs for a corpus, train a DNN or RNN

Description of some other files and folders:

feature\_pecfac.m: Feature extraction code

cochleagram\_toolbox: Signal processing toolbox

DBNSuite: DNN toolbox

examples: Example waveform files, including a noisy signal and a corresponding clean signal

HMM: Viterbi decoding tool

Models: Trained DNN and RNN models

RNN: RNN toolbox

utilities: Utility files

voicebox\_icl: Voicebox toolbox