

# Removing Noise from Speech with Deep Learning

## First Milestone

### 1 Task summary

Our goal is to reduce (or in the best case, entirely remove) the noise from audio recordings containing noisy speech. The idea is to use the WaveNet[1] architecture to generate clean audio from the noisy one.

### 2 Data acquisition and exploration

For our training and testing data, we used a dataset called "Noisy speech database for training speech enhancement algorithms and TTS models" by the University of Edinburgh. It consists of  $\sim 23000$  clean-noisy pairs from 56 different speakers. The samples are stored in separate .wav files of varying length.

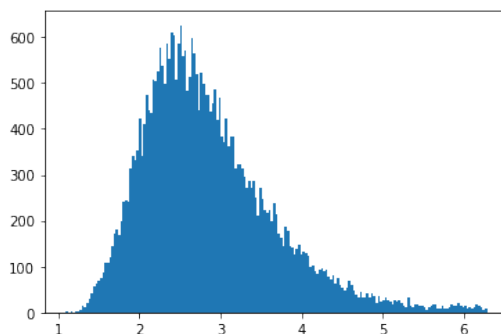


Figure 1: Duration distribution histogram of a subset of speeches

## 3 Data preprocessing

### 3.1 Data selection by duration

We intend to zero pad the data in order to make the samples the same length. To reduce the number of zeroes, we only use the  $n$  samples that are closest to each other in duration. This way, we decrease the number of training samples, but hopefully, the efficiency of our network will be much better.

### 3.2 Zero padding

Our network architecture only supports samples of the same length. To fulfill this constraint, we add zeroes to the end of each recording. The number of necessary zeroes is calculated from the difference between the longest and the current sample.

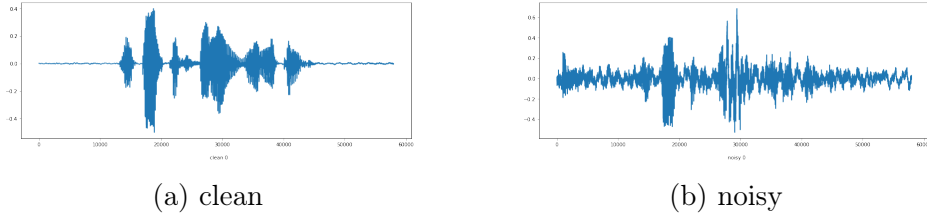


Figure 2: An example clean-noisy pair

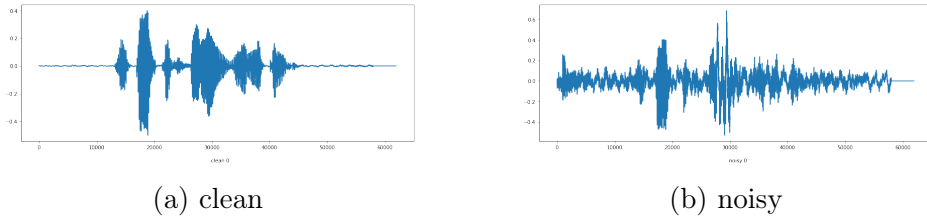
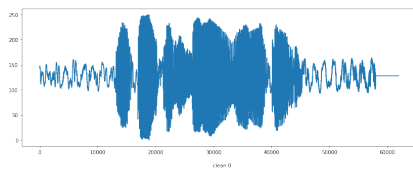


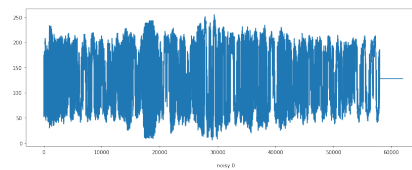
Figure 3: An example clean-noisy pair after zero padding

### 3.3 $\mu$ -law transformation

To make future computations faster we, reduce the samples to 8 bit with  $\mu$ -law transformation.



(a) clean



(b) noisy

Figure 4: An example clean-noisy pair after padding and  $\mu$ -law transformation

## References

- [1] Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. *WaveNet: A Generative Model for Raw Audio*. 2016. arXiv: 1609.03499 [cs.SD].