Removing Noise from Speech with Deep Learning

Andor Kiss
Balázs Glávits
Márk Konrád

Introduction

Our task was to reduce noise from speech using deep learning.

The goal was to preserve sound quality as much as we can, while reducing the noise.

Motivation

Cool noise reducing hardware.

Figure 1: Sennheiser GSP-500

But this is hardware, and we are computer scientists, not electrical engineers.

Motivation

Noise cancelling software.

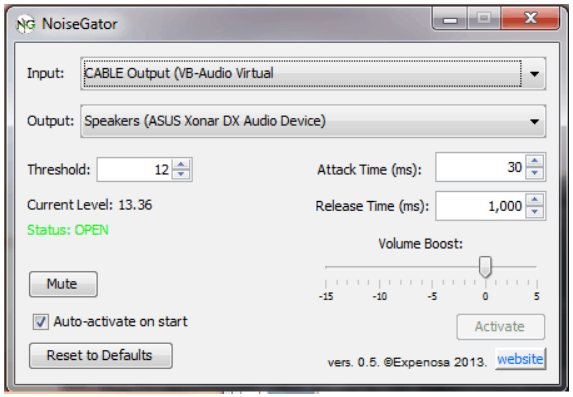

Figure 2: NoiseGator Software

If sound is above the treshold, it goes through.
Else it is cancelled.

Not flexible enough.
Deep learning could do a better job.

Data pipeline

Training phase.

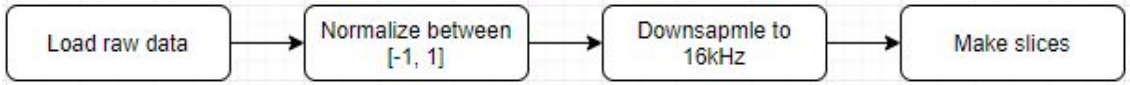

Figure 3: Training preprocessing

We do this on the noisy and clean data as well.
Input: Noisy slices
Output: Clean slices

Data augmentation: Overlapping slices

Full data pipeline

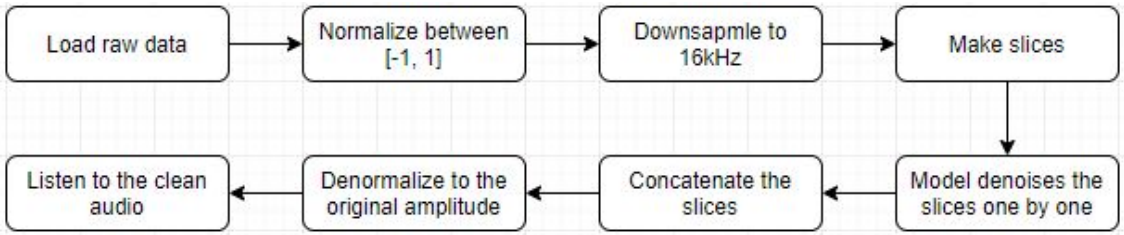

Figure 4: Inference preprocessing

Model is a black box now, it will be elaborated later.

Original WaveNet

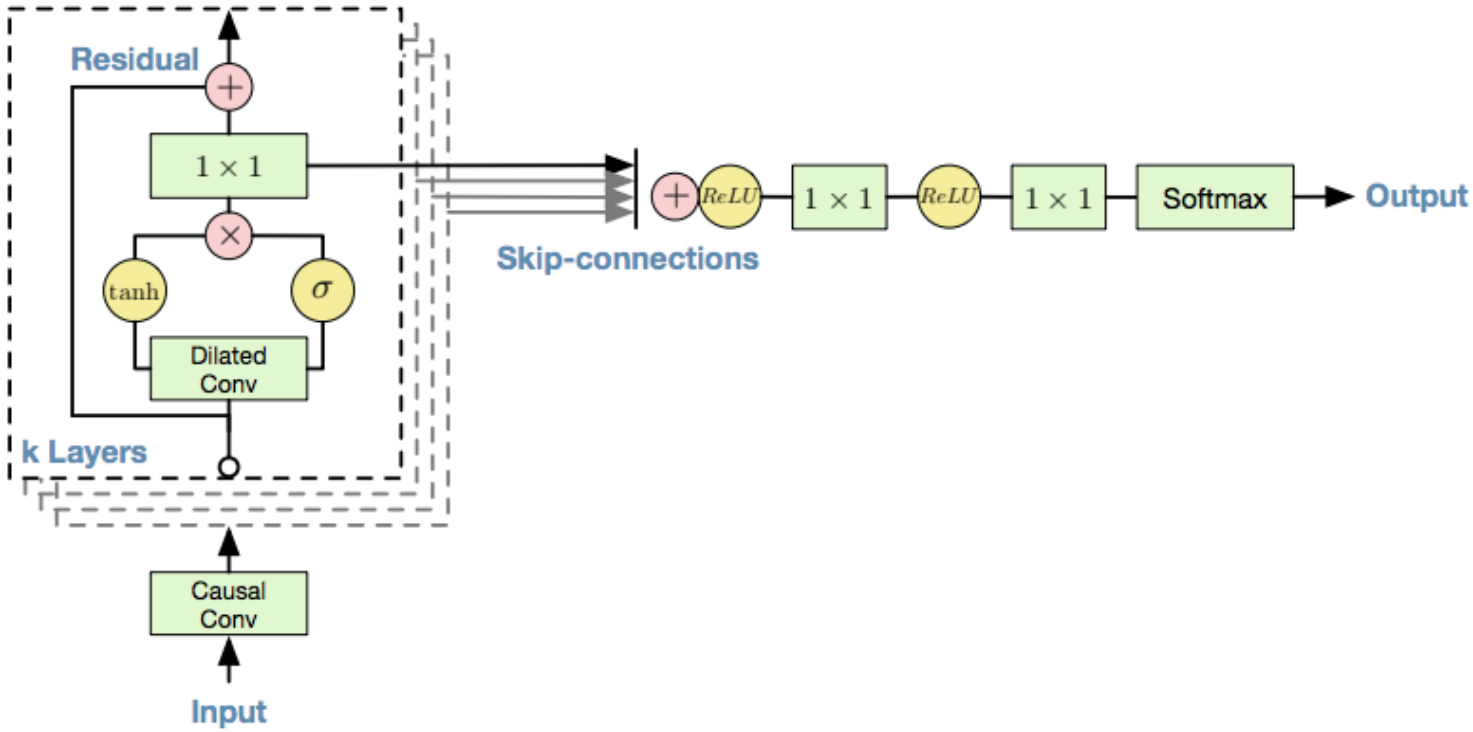

Figure 5: WaveNet

Causal convolutions, mu-law transform and softmax distribution.

Modified wavenet

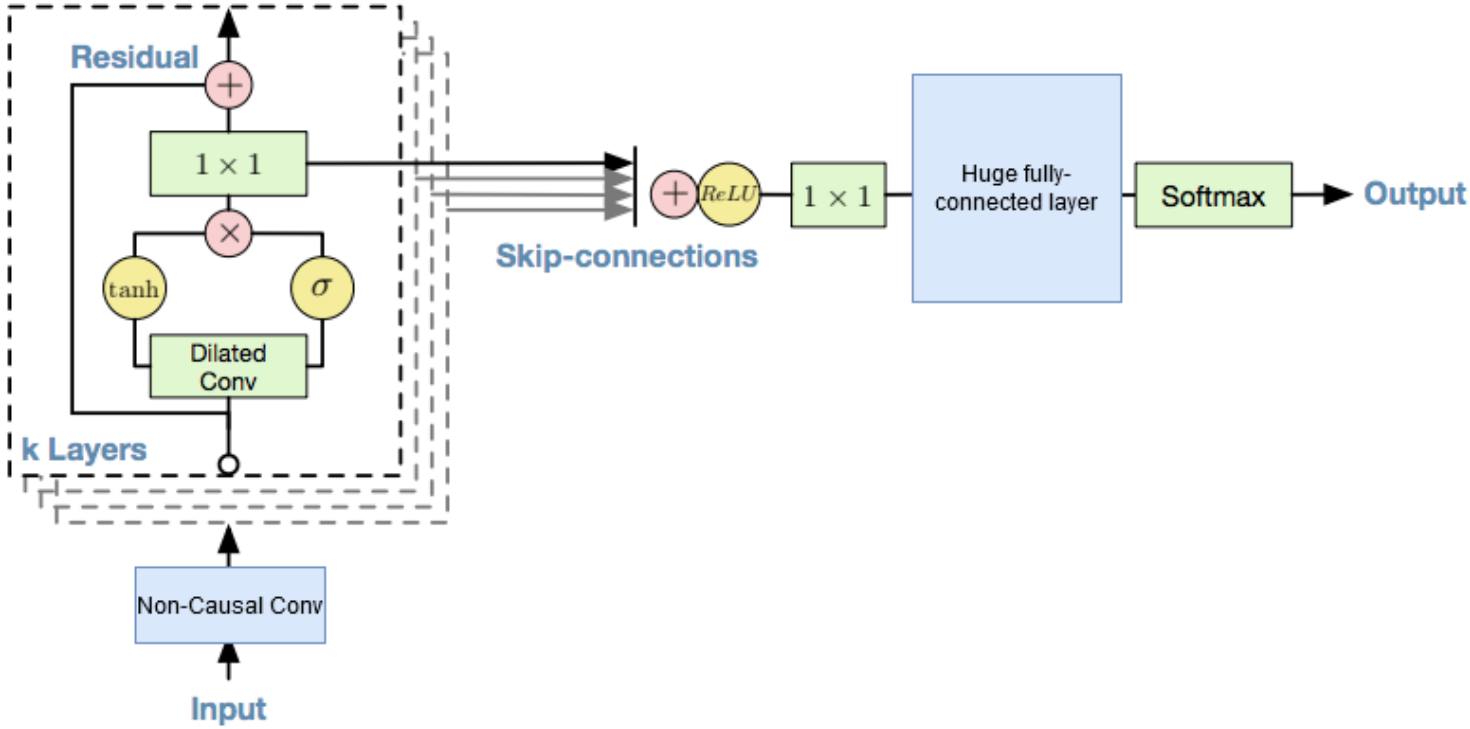

Figure 6: Modified WaveNet

Non-causal convolutions, and dense output layer.

Regression with dense layer

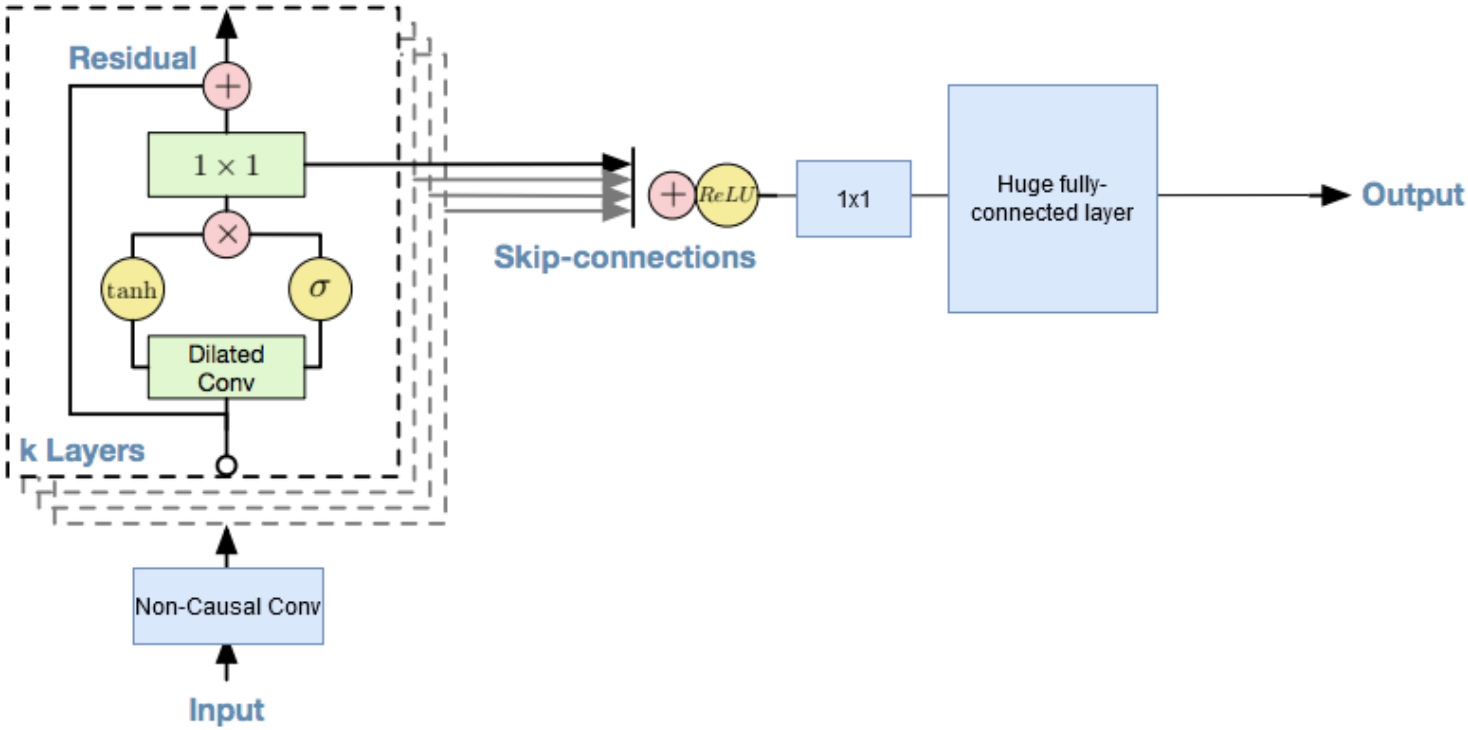

Figure 7: Regression with dense layer

WaveNet with non-causal convolutions, regression, and flatten + dense output layers

WaveNet based autoencoder

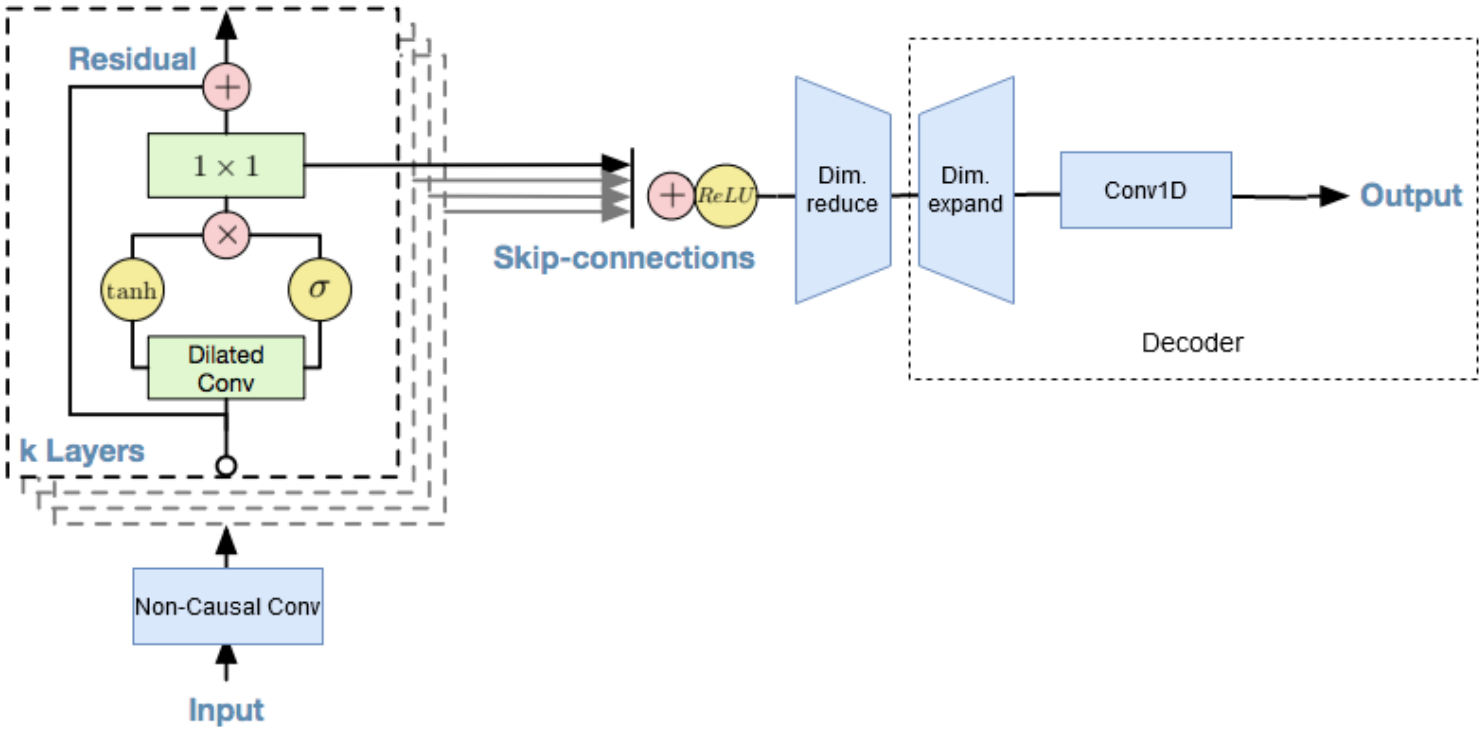

Figure 8: Wavenet based autoencoder

Regression with convolutional layers

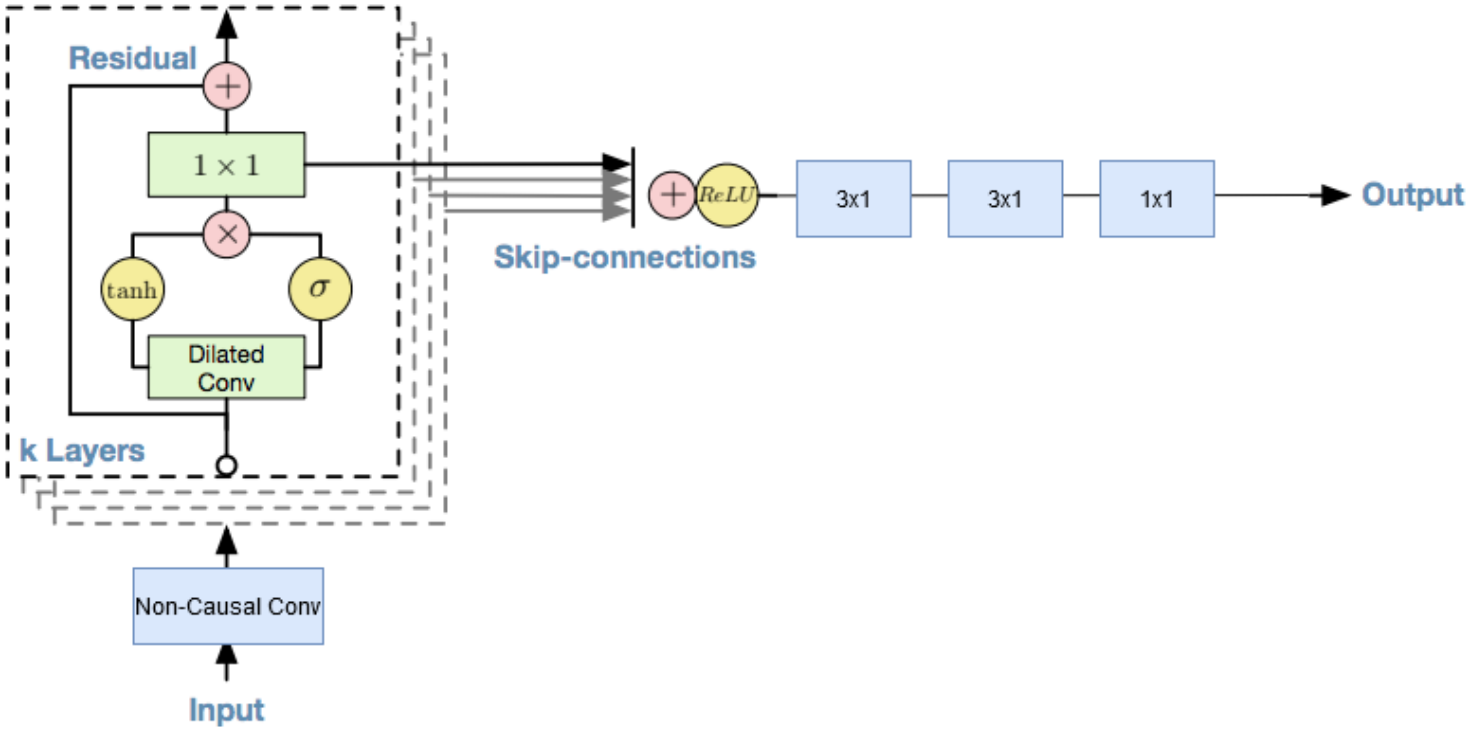

Figure 9: Wavenet based autoencoder

WaveNet with non-causal convolutions, regression, and extra one dimensional convolutional layers on the output.

Training

- Google Cloud Platform
- Clean & Noisy slice generator
- MAE loss
- SGD optimizer
- ReduceLROnPlateau

Demo

Thank you for your attention

Sources:

- Aaron van den Oord, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, AlexGraves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. “WaveNet: A GenerativeModel for Raw Audio”. In: (2016) arXiv:1609.03499
- Dario Rethage, Jordi Pons, and Xavier Serra. “A Wavenet for Speech Denoising”. In: (2018) arXiv:1706.07162