

«Deep Neural Networks in Speech Enhancement»

Байтеков Никита

ВМК, ММП

21 июня 2019 г.

Введение

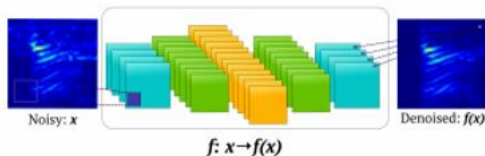
- ▶ Настоящие аудиоданные обычно повреждены посторонним шумом
- ▶ Шум можно убрать с помощью глубоких нейронных сетей (DNN)
- ▶ Применение: автоматическое распознавание речи или говорящего, мобильные коммуникации

Виды архитектур DNN

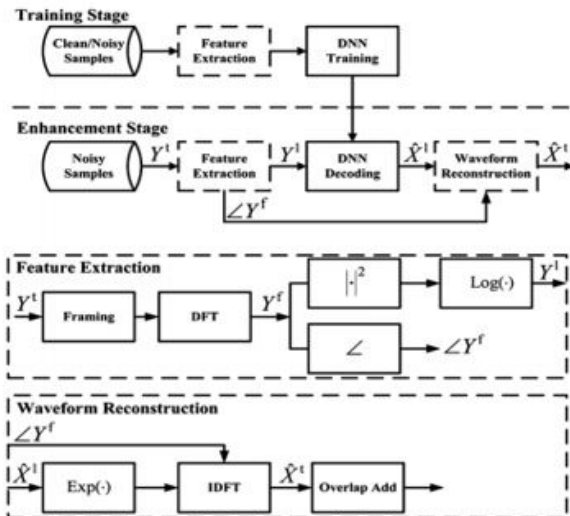
- ▶ Автоэнкодер
- ▶ Рекуррентные нейронные сети (RNN)
- ▶ Ограниченные машины Больцмана (RBM)

Используемый подход

- ▶ Автоэнкодер на лог-спектре сигнала в качестве признаков
- ▶ Спектр из STFT шириной в 256 сэмплов с 50% наложением
- ▶ Обучаем с учётом 16 последних фреймов
- ▶ Метрика качества — MSE
- ▶ Стандартный Adam

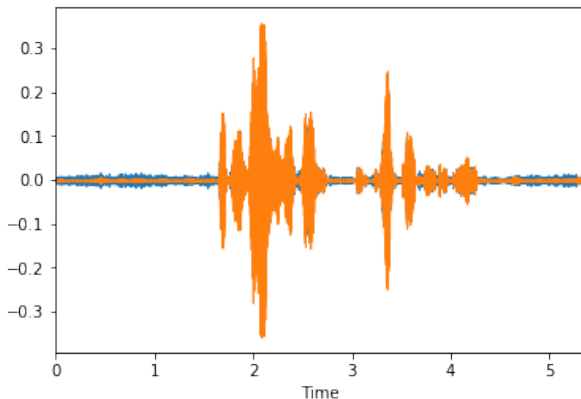


Общая схема



Эксперименты

- ▶ Датасеты Babble/Clean из стандартного датасета TCD-TIMIT
- ▶ Шум успешно фильтруется, итоговый сигнал близок к ground truth



Источники

- ▶ «Speech Enhancement In Multiple-Noise Conditions using Deep Neural Networks» by Kumar, Florencio
- ▶ «Experiments on Deep Learning for Speech Denoising» by Liu, Smaragdis, Kim
- ▶ «Speech Denoising Deep Neural Network» by Chaitanya