

Выбор планировщика обучения нейронных сетей для задачи верификации дикторов

Двойникова А.А., Садикова А.А.

(Университет ИТМО)

В работе рассматривается применение различных планировщиков обучения (англ. scheduler) нейронных сетей для задачи текстонезависимой верификации дикторов. Для экспериментальных исследований использовалась база данных VoxCeleb1, которая содержит в себе различные речевые высказывания 1211 дикторов. Система верификации строилась следующим образом: из аудиоданных извлекались mfcc признаки, которые подавались на вход нейронной сети ResNetSE34L. Также происходила онлайн аугментация данных с помощью двух различных баз данных MUSAN и Simulted room impulse response database. MUSAN представляет набор аудиозаписей различной музыки, речи и шума, а Simulted room impulse response импульсные отклики в различных помещениях, с помощью которой можно добавлять в исходные речевые данные эффект реверберации. Сравнение эталонных признаков с целевыми происходит по принципу косинусной близости.

Экспериментальные исследования проводились с различными планировщиками обучения, представленными в библиотеке PyTorch языка программирования Python, а именно StepLR, OneCycleLR, CycleLR, CosineAnnealingWarmRestarts, CosineAnnealingLR, ExponentialLR, MultiStepLR, MultiplicativeLR, LambdaLR, а также рассматривались 2 новых планировщика, которые не встроены в библиотеку PyTorch, с названиями Cyclic Cosine Annealing и Incremental Cosine Annealing. Суть алгоритма Cyclic Cosine Annealing заключается в том, что весь процесс обучения делится на M циклов, каждый из которых начинается с большой скорости обучения, а заканчивается маленькой. Алгоритм Incremental Cosine Annealing является модификацией Cyclic Cosine Annealing. Суть алгоритма Incremental Cosine Annealing заключается в следующем, для каждого цикла увеличивается максимальное значение скорости обучения на c , где $c = 0.001$, причем параметр c в 10 раз меньше начальной скорости обучения, равному 0.01.

В статье представлены сравнительные результаты точности верификации, выраженные метрикой Equal Error Rate (EER). EER показывает значение пересечения ошибок первого (англ. False Acceptance Rate, FAR) и второго рода (англ. False Rejection Rate, FRR). Как показывают экспериментальные исследования наименьшая EER = 3,2857 достигается при использовании планировщика обучения StepLR и оптимизатора adam.

Садикова А.А. (автор)