

Отчет по четвертому домашнему заданию по курсу “Глубинное обучение в обработке звука”

Выполнил: Михаил Малафеев

Я выжил.

В ходе выполнения четвертого домашнего задания была реализована модель HiFi-GAN (ну типа). Собственно все было взято из статьи (или из тех, на которые ссылались). Кроме расписания `learning rate`'а, его взял уже по привычке `CosineAnnealingLR`, чтобы иметь возможность оставить надолго. Конфиг генератора взят самый экономный (3) по очевидным причинам причинам) Из экспериментов и сложностей могу отметить, что модель достаточно неплохо учится и для `L1Loss`'а по самой аудиодорожке (.wav), вместо мелспек. Это интересное наблюдение было получено в ходе исправления кривых рук и обращения к репозиторию статьи, поскольку в ходе прохождения этого курса я разучился воспринимать что-либо человеческое и неправильно воспринял раздел **Mel-Spectrogram Loss**. Возможно, такое обучение можно исследовать или поискать статьи, которые делают функцию потерь по аудио, но кажется это просто медленнее (но это неточно). Логи можно посмотреть тут: [1](#), [2](#), [3](#). (как всегда несколько, потому что единственное стабильное в датасфере - это ее падения). В README.md описан процесс воспроизведения результатов обучения.

Не знаю, что сказать еще в рамках этого отчета, поэтому скину мем.

