

Отчёт по заданию 1

Галимов Даниил

31 марта 2025 г.

1 Введение

В данном отчёте представлены результаты экспериментов по реализации слоя LogMelFilterBanks и обучению простой CNN модели для бинарной классификации команд «yes» и «no» из датасета SpeechCommands. Эксперименты проводились с изменением числа Mel-фильтров (`n_mels`) и использованием групповой свёртки.

2 Эксперименты с числом Mel-фильтров

Были протестированы модели с `n_mels` равными 20, 40 и 80. Ниже приведены итоговые метрики для каждой конфигурации:

<code>n_mels</code>	Test Accuracy	Параметры	FLOPs (КМас)	Средняя TrainLoss	Средняя ValAcc
20	96.84%	5746	187.2	0.12	96.9%
40	96.72%	6706	283.2	0.13	96.9%
80	98.54%	8626	475.2	0.14	97.8%

Таблица 1: Сравнение моделей при различных значениях `n_mels`.

На рисунках 1 и 2 показана динамика Train Loss и Validation Accuracy по эпохам для различных значений `n_mels`.

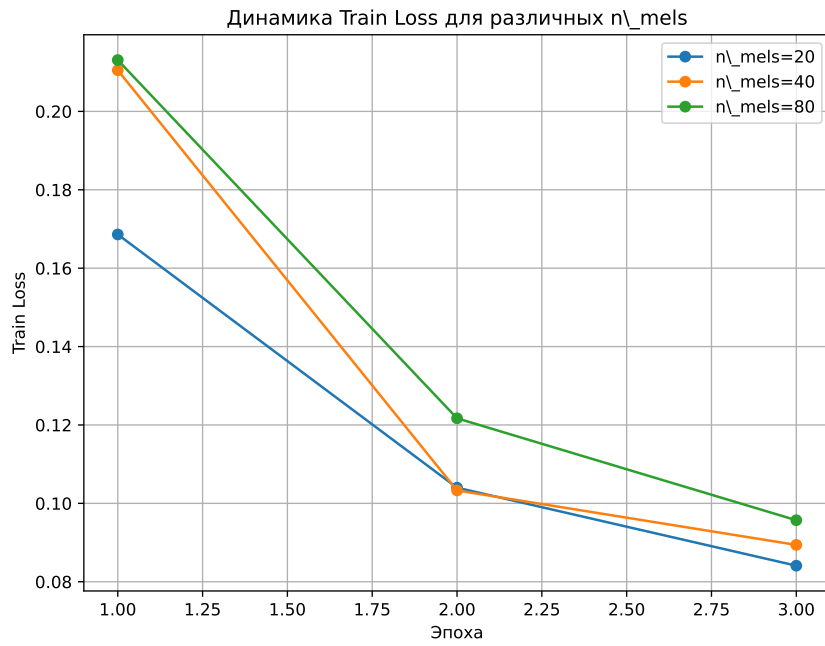


Рис. 1: Динамика Train Loss для различных значений n_mels .

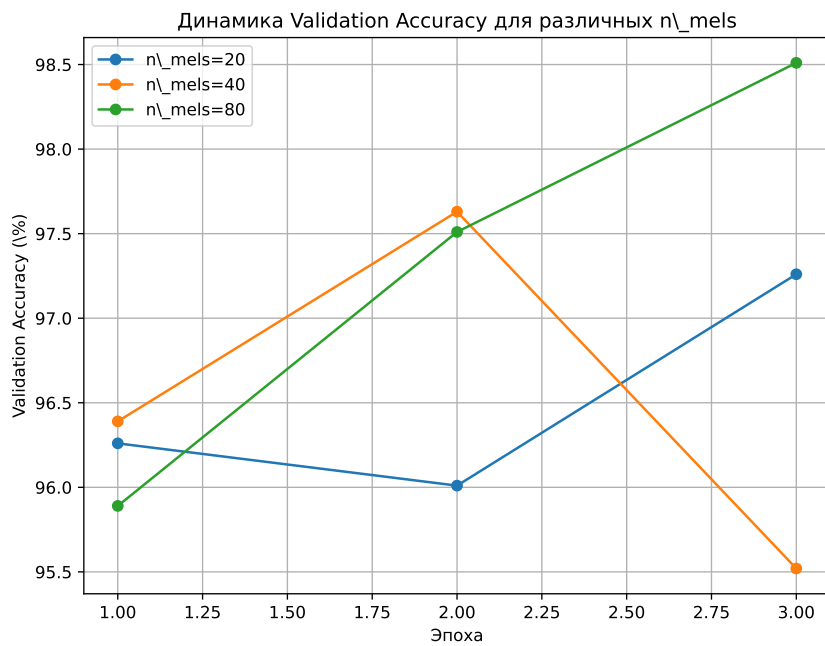


Рис. 2: Динамика Validation Accuracy для различных значений n_mels .

3 Эксперименты с групповой свёрткой

Эксперименты проводились с параметром `groups` равным 2, 4, 8 и 16. Итоговые результаты приведены в таблице:

Группы	Test Accuracy	Параметры	FLOPs (КМас)
2	97.09%	5938	244.8
4	98.67%	4594	129.6
8	96.12%	3922	72.0
16	95.15%	3586	43.2

Таблица 2: Метрики при различных значениях групповой свёртки.

На рисунке 3 показаны зависимости числа параметров, FLOPs и Test Accuracy от числа групп.

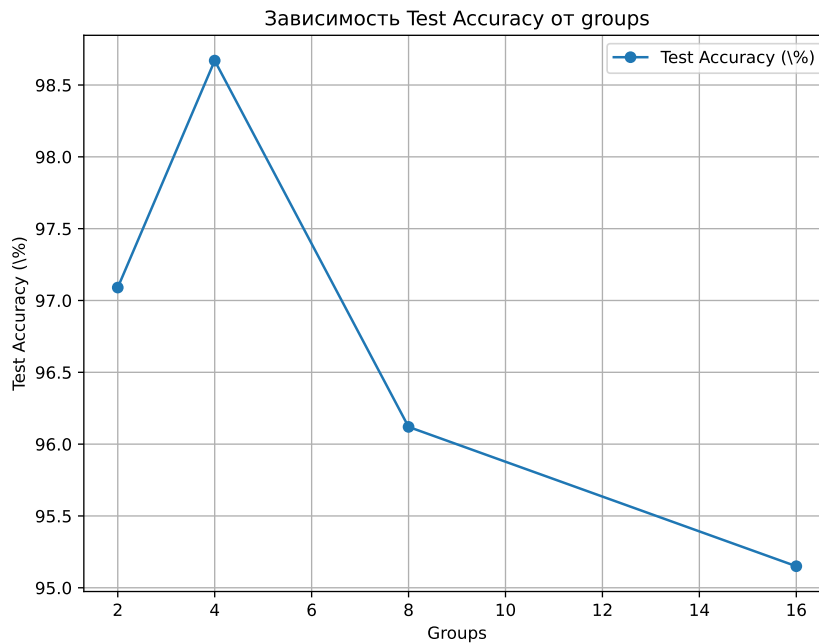


Рис. 3: Зависимость метрик от числа групп.

4 Анализ результатов

По результатам экспериментов можно сделать следующие выводы:

- Увеличение числа Mel-фильтров приводит к росту числа параметров и вычислительной нагрузки (FLOPs). При `n_mels=80` наблюдается улучшение тестовой точности, однако это сопровождается увеличением вычислительных затрат.
- Применение групповой свёртки позволяет существенно сократить число параметров модели. Оптимальные результаты достигнуты при `groups=4`, тогда как при дальнейшем увеличении числа групп качество модели начинает снижаться.

5 Заключение

Модель с `n_mels=80` и `groups=4` демонстрирует оптимальное соотношение между точностью и вычислительными затратами.