

Введение

Целью задания было создание нейронки для классификации изображений с использованием данного нам набора данных. Там было 100000 фотографий и их классификация. Тестовая выборка состояла из такого же кол-ва фоток, но они были размера 40 на 40. Нужно было сделать правильную классификацию этого набора фоток. Конечно я не смог получить датасферу и не понял, как ей пользоваться. НО для чекпоинта ничего более моего личного компьютера не понадобилось. После миллиона часов загрузки фоток, я просто использовал простые сверточные слои с пулингом и полносвязными слоями для достижения минимального качества, в итоге было где-то 17% и я довольный пошел навстречу новому году).

После НГ

В новом году оказалось, что простое тыканье в конволюционные сети - гиблое дело. Нужно было взять что-то готовое, чтобы потом его доработать. Изучив документацию torchvision, я решил воспользоваться архитектурой ResNet-34, скопировав код из исходников и адаптировав его под задачу. Я убрал из модели начальный слой с ядром 7x7 и слой MaxPool для лучшей работы с изображениями 40x40. Провел обучение модели на 10 эпохах. Вместо оптимизатора Adam, начал использовать SGD, вроде как он эффективнее. Ну и конечно кросс энтропию не забыл в качестве критерия. В результате 29% точности и как будто минимум уже выполнен. Добавил эпох до 10 штук и планка в 30% побита.(31%). Но теперь стало понятно, что признаки начинали переобучаться. Добавляя еще эпох я получал худшее качество. Кстати писал я все в катле.

Борьба с переобучением

Теперь уже без напряжения быть самым тупым на направлении я приступил использовать аугментацию, как способ разнообразить данные, так как видно было, что модель подгонялась уже и качество не то что не росло, а падало. Я взял базовые методы типа RandomHorizontalFlip, RandomRotation. Думал, что это не сильно улучшит модельку, но прям буст был неплохой. По итогу 39.98%. И я решил на этом пока приотостановиться. Сейчас добавлю ещё 1 эпоху, чтобы поднять качество. Но скорее всего модель начала уже переобучаться и до дд я не успею глянуть что было на 9 эпохе. Randomcrop ухудшил мою модельку, поэтому убрал его, мб неправильную площадь взял, потому что лектор сказал, что это хорошая вещь, но я не стал сильно стараться докапываться до истины. Можно было использовать Mixup и Cutmix, но качества уже хватало, так что я решил особо не париться на этот счет, думаю мне % 5-10 ещё дало бы.

Небольшой вывод

По ходу работы 1000 раз пожалел, что не использовал графики. Они бы точно упростили работу и улучшили понимание. Но опыт прикольный и полезный, жду 2ое бдз, чуть больше постараюсь)

Подытоживая:

- Простая архитектура: ~17%
- ResNet-34: ~31%
- ResNet-34 + простая аугментация: ~39.98%

Выводы

- 1 Простая архитектура оказалась недостаточно мощной для данной задачи, что продемонстрировало важность использования проверенных архитектур, таких как ResNet.
- 2 Добавление методов аугментации данных, существенно улучшило точность модели.
- 3 Метод RandomCrop работал плохо, так как часто "обрезал" ключевые области изображений.
- 4 Финальная модель достигла точности ~39.98%, что соответствует требованиям задачи