

Качество и понятность кода

Оценка: 3 балла из 3.

Метрика качества

В лидерборде 0.93528, доп. баллы за этот пункт начинаются с 96.5%. У вас размер изображений всего 224x224, с одной стороны это экономит вычислительные ресурсы, но с другой – ограничивает точность.

Работа с обучающими данными

Сделано очень много аугментаций, но я не вижу результатов (визуализации) этих аугментаций, потому что в ноутбуке вообще нет выводов ячеек. Но в целом кажется что все с аугментациями хорошо.

Есть однако большая проблема: валидация (`test_datagen`) делается также с аугментациями. Так не должно быть. Представьте, что одну модель вы обучили на слабых аугментациях, другую на сильных. Если валидация делается с аугментациями, то модель с сильными аугментациями покажет более низкую точность на валидации. Если же делать валидацию нормально, без аугментаций, то все может быть наоборот. Поскольку валидация влияет на выбор моделей и гиперпараметров, важно делать ее без аугментаций.

Оценка: 2 балла из 3.

Работа с архитектурой модели

Используется всего одна модель, при этом в нее добавляется скрытый полносвязный слой и BatchNorm, но такое решение никак не обосновывается. Выбор модели сам по себе неплохой, Xception хорошая модель, но согласно критериям оценивания выбор нужно обосновывать – либо сравнением моделей на тестах, либо научными статьями или статьями из интернета.

Оценка: 0 баллов из 3.

Работа с процессом обучения

Используется `lr_find()` и `plot_lr()`, что хорошо, хотя результаты непонятны – поскольку почему-то пропали выводы ячеек. Судя по коду, сам процесс обучения построен неверно: сначала вы обучаете всю модель целиком, а затем замораживаете половину слоев. Так никогда не делают: наоборот сначала замораживают часть слоев, а потом обучают целиком.

Оценка: 1 балл из 3.

Работа с процессом инференса

По критериям оценивания нужно использовать ТТА и ансамблирование, в коде этого не нашел.

Оценка: 0 баллов из 3.

Создание работающего прототипа

Этого пункта у вас не нашел.

Сумма: 3+2+1=6 баллов.