

Соревнование: классификация пиццы

Екатерина Самородова

Май 2023

Содержание

любимая деда мороза



миксгрин



ветчина грибы



палочки



пепперони



мясная



пепперони



цыпленок флорентина



цыпленок грин



Исходные данные и
анализ



Аугментация и
модели

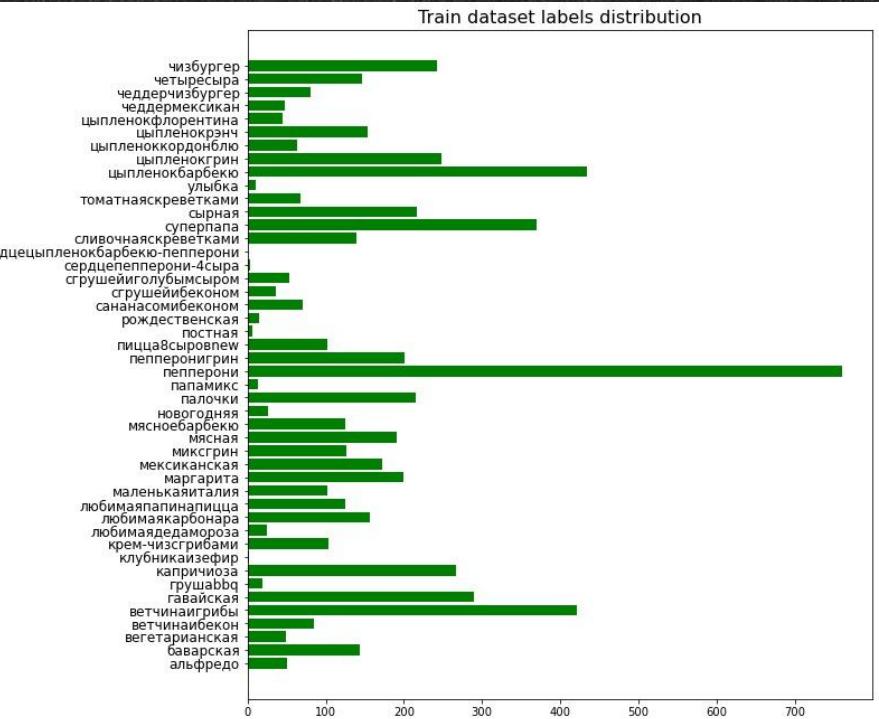


Результаты

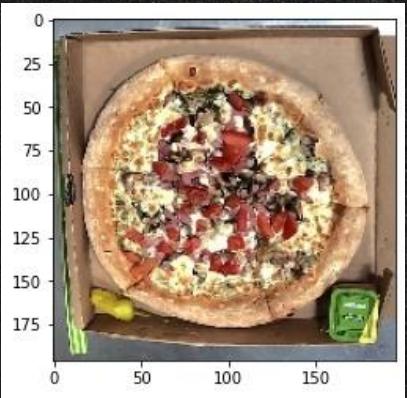


Исходные данные

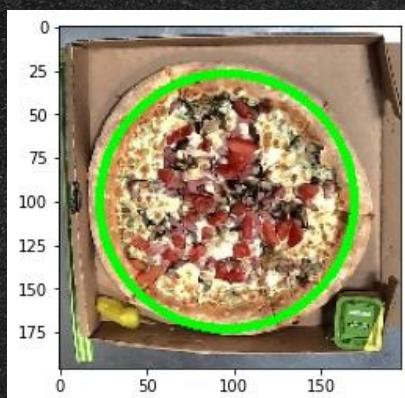
- + достаточно много данных
- + хорошо разделяемые классы
- + “блочный” формат ответа
- ошибки в данных (2 вида)
- дисбаланс классов
- плохо разделяемые классы



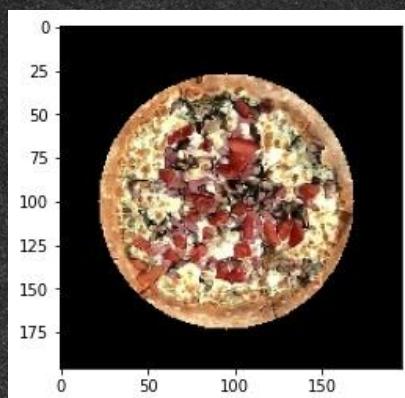
Аугментация: первый этап



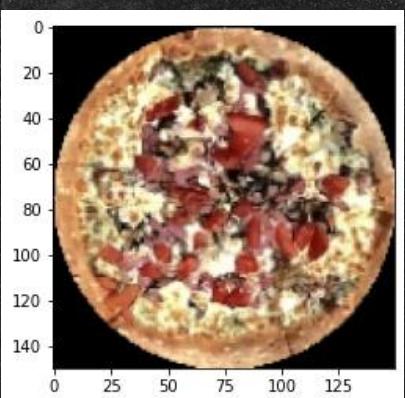
Исходное (0001.jpg)
изображение пиццы



Выделяем главный круг
при помощи cv2



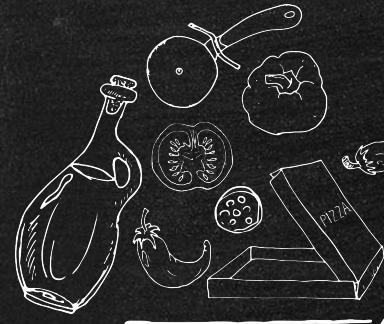
Вырезаем лишний фон
вокруг начинки



Обрезаем края и
перемасштабируем

Идея: для определения вида пиццы начинка важнее, чем корочка, коробка, перец и соус на фотографии.

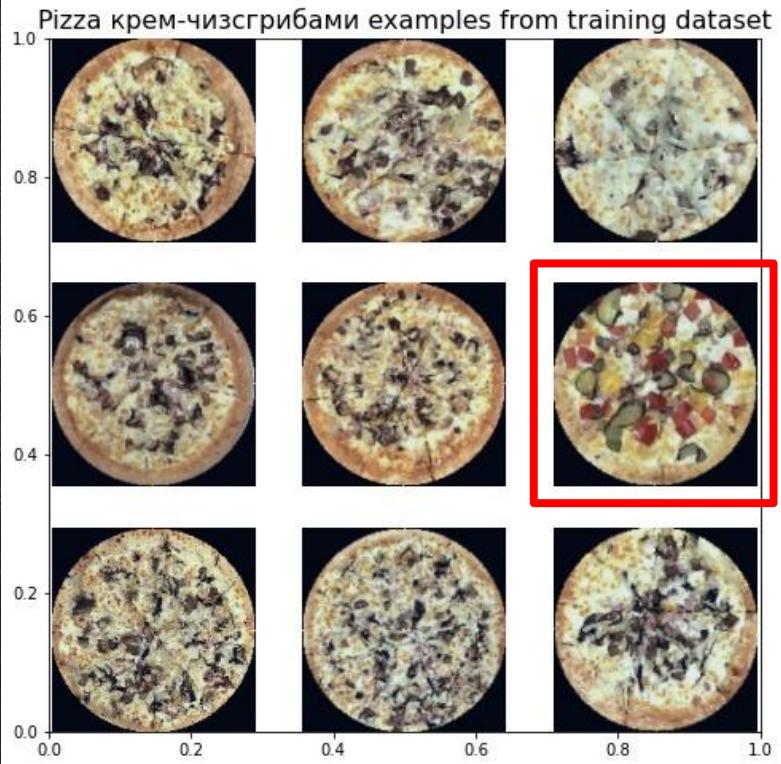
Результат: уменьшение шума и размера картинки.



Анализ исходных данных

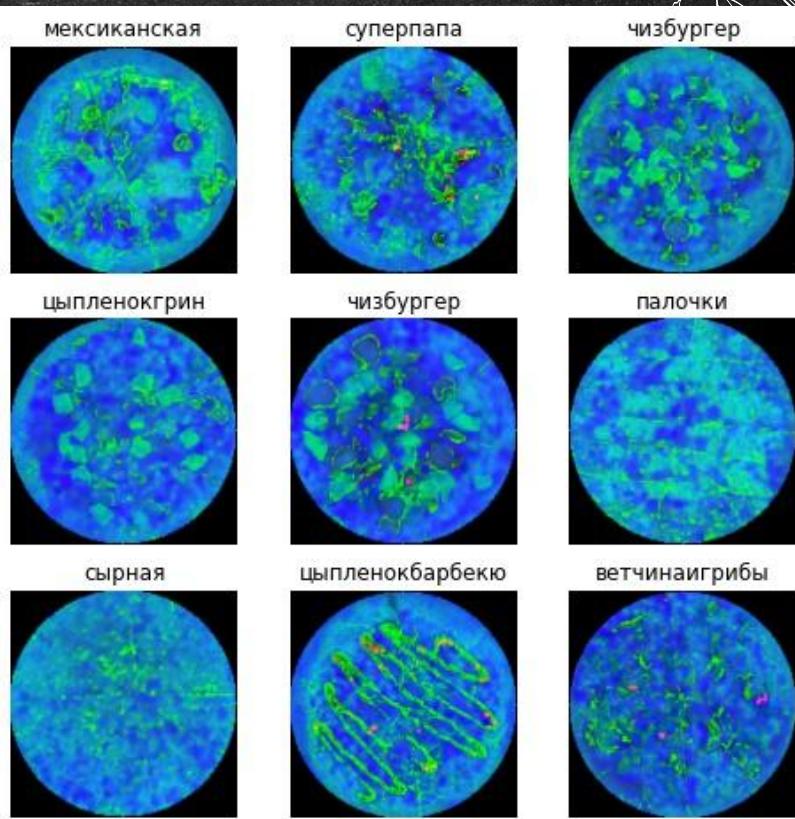
Проблема: наборы обучающих и тестовых данных иногда содержат ошибки, смазанные картинки и т.д.

Решение: можно исправить метки с ошибками или просто удалить из обучающего набора неподходящие (некорректные) изображения.



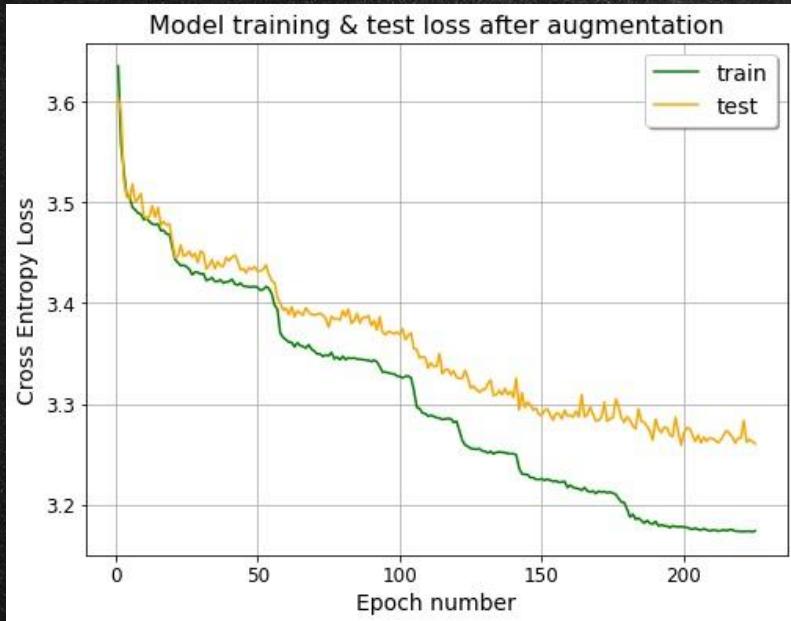
Аугментация: второй этап

- + вырезание кругов по центру
- + нормировка RGB-значений
- + случайные повороты
- + случайные отражения (по вертикали и горизонтали)
- другие цветовые пространства (YCbCr, HSV из cv2)
- другие нормировки значений



Модели: сверточная сеть

Архитектура: 3 блока сверточных слоев + полно связанный слой
(блоки Conv2d → ReLU → MaxPooling → BatchNormalization)



Преимущества: всегда обучается, не очень большое число параметров, мало переобучается с аугментацией.

Недостатки: медленно обучается, требуется много эпох для обучения, невысокая точность предсказания, делает ошибки на схожих классах.

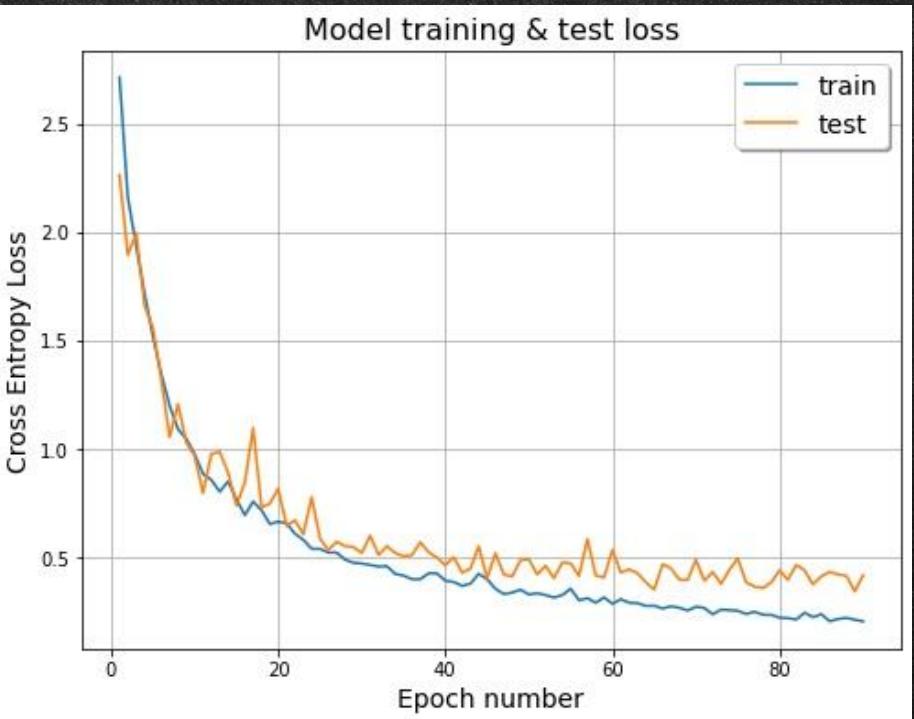
Результаты: преодоление baseline, максимально около 60% accuracy.

Модели: MobileNet v2

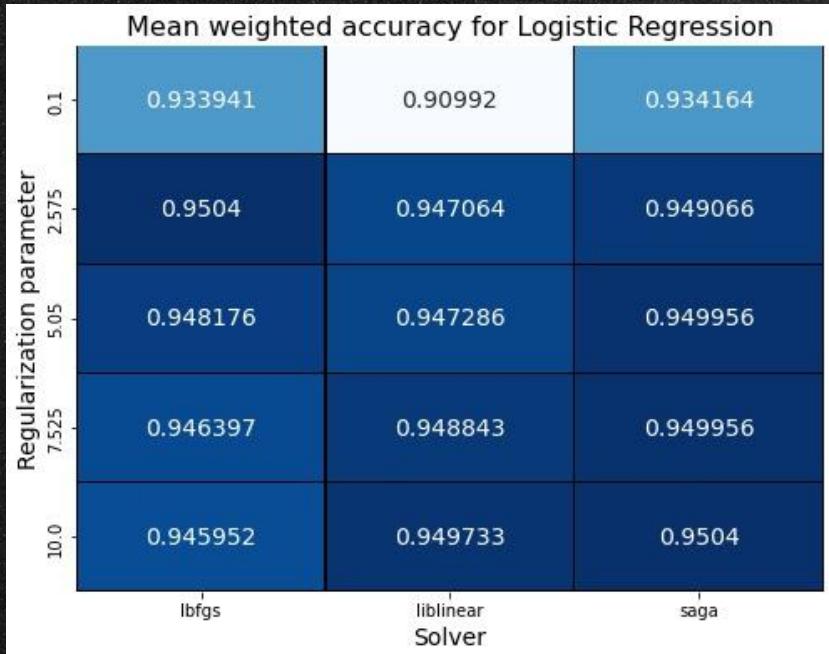
Преимущества: требует меньшего количества эпох обучения, не очень большое число параметров, в итоге выше точность классификации.

Недостатки: "скачки" loss-функции, модель быстро перестает обучаться, иногда склеивает похожие классы.

Результаты: порядка 72% accuracy.



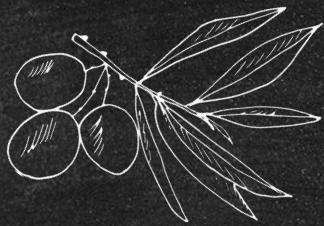
Модели: MobileNet + LogReg



Проблема: полученная нейросеть
“склеивает” похожие классы.

Решение: обучить на выходах сети
модели логистическую регрессию.

Результаты: около 80% accuracy.



Итоговая схема алгоритма классификации пиццы

Подготовка данных

1. Вручную удаляются смазанные картинки, а также изображения с предположительной неправильной меткой
2. Тестовый датасет размечается вручную.



Аугментация

1. Выделение главного круга на изображении
2. Масштабирование
3. Случайные повороты, отражения, нормировка

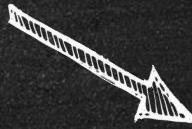
Обучение MobileNet v2

1. Разделение датасета на тренировочную выборку и валидационную.
2. Обучение MobileNet
3. Вероятности классов на тестовой выборке



Обучение регрессии

1. Обучение логистической регрессии на результатах работы модели нейросети
2. Предсказание классов
3. Анализ результатов



Pizza labels & predictions cross tabulation

Матрица выходов полученной модели

Итог: значения сконцентрированы вдоль главной диагонали, значит, большинство видов пиццы модель различает достаточно успешно.

Проблема: некоторые классы всё ещё могут “склеиться” вследствие небольших различий и дисбаланса классов в обучающей выборке.

Решения: сгенерировать примеры редких классов изображений, или обучить отдельные модели на них.

Спасибо за
внимание!

