

mipt2024s-5-belkov-alexey

Report 20.05.2024 Белков Алексей Сергеевич M01-301Б

Задача оптимизации генератора Задача оптимизации генератора представляет собой задачу оптимизации черного ящика, для её решения понадобится следующая информация

- Пространство оптимизируемых параметров
- Оценка качество работы генератора
- Метод оптимизации

Разбор генератора

На данный момент имеется генератор кодов, генератор искажений и набор аугментаций
Генерация кода происходит следующим образом:

- Генерируется изображение кода
- Применяется perspective transform (далее PP) со случайными параметрами
- Применяются фиксированный набор аугментаций (но каждая из них уже случайным образом воздействует на изображение)
- Полученное изображение кода вставляется в заранее поданное изображение фона

На выходе получаем *синтезированное* изображение

В качестве пространства параметров на котором будет выполняться оптимизация было выбрано:
Параметры распределения для генерации параметров PP + пространство параметров аугментаций

(Так же сюда можно добавить и сами аугментации, для выявления чрезмерно нереалистичных или совсем неэффективных аугментационных преобразований, об этом далее)

Введя пространство оптимизируемых параметров, введём метод оценки качества генерации

Изначально такой метрикой было попарное MSE при генерации батча изображений, с помощью неё даже получалось повысить разнообразие генерируемой выборки, но такая метрика никак не свидетельствует о реалистичности результатов генерации

Для оценки качества генерации выбор пал на две метрики

- [Fréchet inception distance \(FID\)](#)
- [Inception score \(IS\)](#)

- FID позволяет сравнить близость распределения сгенерированной выборки (распределения которое моделирует генератор) с распределением реальных данных
- IS можно интерпретировать как оценку репрезентативности и разнообразия генерации

Как выполнялась оценка

Будем получать глубокие эмбединги изображений из модели детекции обученной `model_ab` от [belkov-arseniy](#) либо из модели сегментации `model_t` от [Tiniakov-A-D](#)

Первичная оценка:

- Пусть имеется датасет D (для модели эксперимента нам важно лишь что это размеченный для детекции датасет изображений объектов с кодами на них)
- Разобьём его (стратифицированно по типу кодов) на два равномоощных, неперсекающихся датасета $D1$, $D2$
- Для $D1$ предиктнем боксы, и внутри этих боксов проведем генерацию кода, т.е. перекроем существующий реальный код сгенерированным
- $D2$ не меняется
- Получим эмбединги $E1$ для $D1$ и $E2$ для $D2$ (т.е. $E1$ - множество полученное получением эмбединга каждого измененного изображения в $D1$, $E2$ аналогично, с отличием лишь только в том что $D2$ не был подвержен изменениям)
- Посчитаем FID на основе $E1$ и $E2$, посчитаем IS

Более сильный способ оценки:

- Как и в предыдущем методе из датасета D получим $D1$ и $D2$, также вставим в картинки $D1$ сгенерированные коды
- Обучим модель на $D1$ и отвалидируем на $D2$

В данном случае мы устраиваем классическую проверку обобщаемости модели на реальные данные при обучении на синтезированных

Так же в последнем способе можно разбивать на датасеты разных и не изменять часть картинок в $D1$ генерацией

Метод оптимизации

Будем использовать байесовскую оптимизацию через библиотеку Optuna, оптимизируемый функционал - FID из первичной оценки, сам алгоритм оптимизации будет устроен так:

- Optuna выбирает параметры для итерации

- С этими параметрами проводим процедуру первичной оценки (разбиение датасета заранее фиксировано)
- Оценка FID
- Оптимизационный шаг Optuna

Miscellaneous результаты

Удалось собрать 4 публичных датасета:

1. <https://www.kaggle.com/datasets/jonathanimmanuel/barcode-and-qr/data>
2. <https://www.kaggle.com/datasets/whoosis/barcode-detection-annotated-dataset/data>
3. <https://www.kaggle.com/datasets/kniazandrew/ru-goods-barcodes>
4. <https://www.kaggle.com/datasets/bd2927/ideal-barcode-mask-alignment>

(1) - (3) - детекция, (4) - сегментация

Эксперименты

На данный момент от IS пришлось отказаться, нет понимания как применять его для детекционных и сегментационных датасетов

Эксперименты с первичной оценкой и оптимизацией проведены на открытом датасете (3)