

Рецензия на статью **DatasetGAN: Efficient Labeled Data Factory with Minimal Human Effort**

Гришанин Виктор

Содержание и вклад

В статье рассматривается новый подход к генерации размеченных изображений для задачи сегментации. Авторы предлагают модель DatasetGAN, основанную на модели StyleGAN, генерирующей изображения, с дополнительной обработкой векторов признаков для каждого пикселя и последующей классификацией для создания итоговой разметки сегментов изображения. Такой подход требует малое количество размеченных данных для обучения и позволяет генерировать большое количество качественно размеченных изображений, что решает нехватку тренировочных данных для задачи сегментации.

Сильные стороны

- Актуальность

Нехватка размеченных данных для обучения моделей сегментации изображений - достаточно актуальная проблема. Таким моделям нужно очень много тренировочных данных для получения хорошего качества сегментации. Разметка всех этих изображений вручную занимает очень много времени. DatasetGAN позволяет решить эту весомую проблему - с помощью этой модели и небольшого количества размеченных данных можно получить практически бесконечное количество размеченных данных для дальнейшей тренировки моделей сегментации.

- Новизна + качество

Ранее никто не использовал StyleGAN и свойства векторов признаков итогового изображения для генерации размеченных сегментированных изображений. Этот подход требует значительно меньше данных, чем существующие fully supervised методы, и показывает качество не сильно хуже, чем эти методы. Учитывая, что метод новый и еще не до конца исследованный, его наверняка еще можно оптимизировать для получения лучших результатов.

Слабые стороны

- Плохое описание алгоритма

В статье не хватает подробностей и обоснованности при описании некоторых частей модели. В частности, очень мало сказано про классификатор для пикселей - к примеру, нет подробного теоретического или практического обоснования, почему сам ансамбль включает в себя 10 моделей.

- Применимость

С одной стороны, если есть StyleGAN, обученный на данных, сегментация которых необходима в задаче, то использование DatasetGAN с основой из такого StyleGAN весьма правильное и хорошее решение в случае нехватки размеченных данных. С другой стороны, если в задаче требуется разметка некоторого типа изображений, но нет такого StyleGAN, который был бы обучен генерировать подобные изображения, то задача сводится к обучению StyleGAN с нуля, а уже потом преобразованию в DatasetGAN. В данном случае возникает вопрос - стоит ли тратить много времени и ресурсов на обучения StyleGAN для генерации необходимого типа изображений или проще разметить все вручную?

- Качество

Описанная модель обучена всего лишь на 5 типах изображений и для большинства из них размечает не самое хорошее число объектов. Более того, можно было бы сравнить результаты описанного подхода с конкурирующей моделью, кратко описанной в статье. Стоит понимать, что качество взятого StyleGAN очень сильно влияет на решение задачи генерации размеченных данных. Для использования сгенерированных изображений в качестве тренировочной выборки, они, очевидно, должны быть достаточно реалистичными, и плохо обученный StyleGAN негативно повлияет на итоговые данные, пусть даже разметка будет хорошей.

Насколько хорошо написана статья

В целом, статья не является сложной для понимания, общий алгоритм понятен, для всего есть иллюстрации. Однако есть нехватка подробностей при описании отдельных частей модели, а также есть небольшие опечатки.

Воспроизводимость

Часть методов, используемых в модели, достаточно хорошо описана в статье, чтобы без проблем её реализовать. Однако есть некоторые методы, описанные практически на поверхностном уровне, что затрудняет их реализацию. К статье приложен репозиторий с имплементацией описанной модели, что упрощает реализацию непонятных методов.

Дополнительные комментарии

- В 3.1 и 3.2 $\{S^0, \dots, S^k\}$ должны быть получены из $k+1$ Adaln слоя

Оценка - 6: Marginally above the acceptance threshold.

Уверенность - 3: You are fairly confident in your assessment.