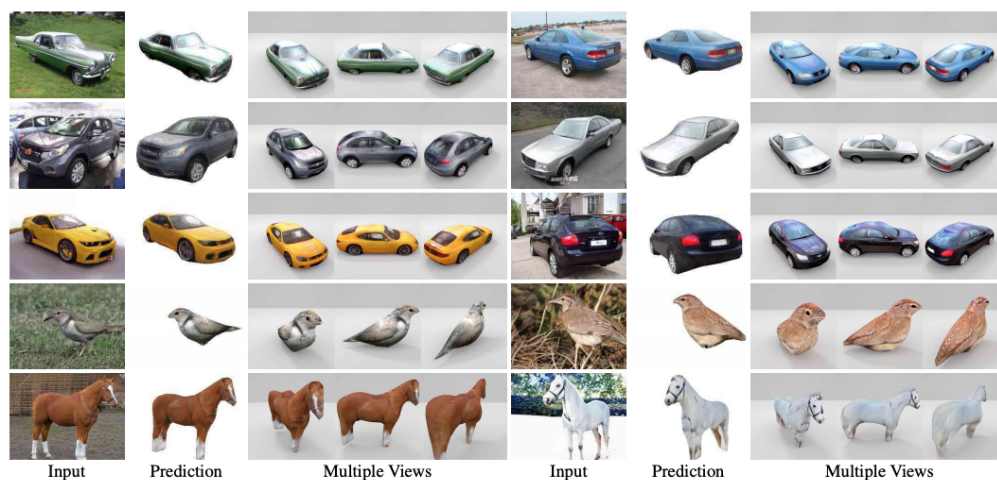


Работа опубликована в апреле 2021 года и принята на CVPR 2021 как орал. У статьи две версии, которые вышли с разницей в неделю.

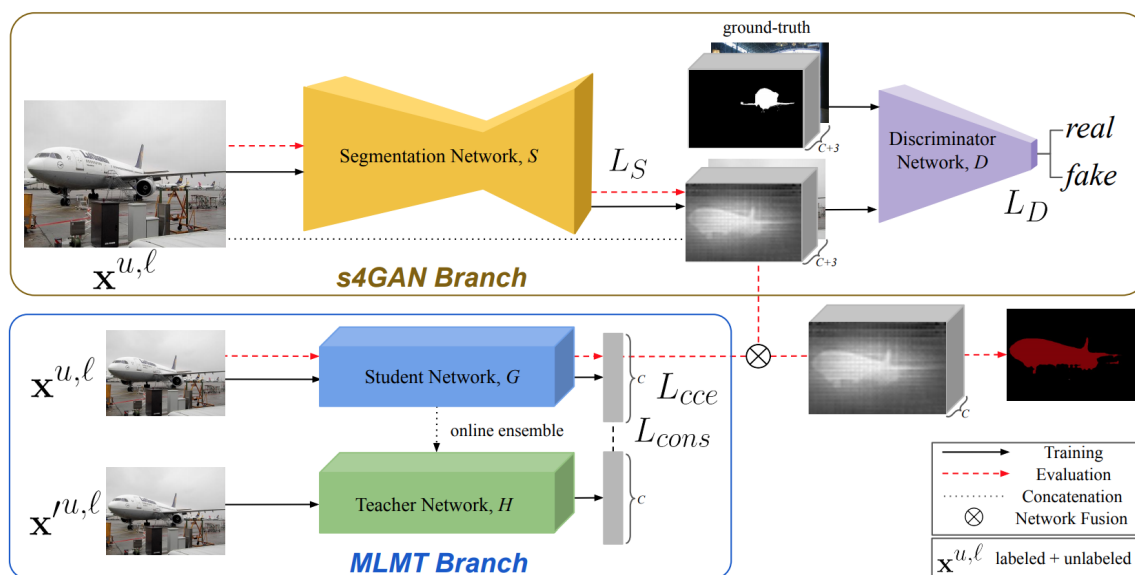
Авторы статьи: Yuxuan Zhang, Huan Ling, Jun Gao, Kangxue Yin, Jean-Francois Lafleche, Adela Barriuso, Antonio Torralba, Sanja Fidler, которые являются сотрудниками NVIDIA, MIT, University of Toronto, University of Waterloo и Vector Institute. Основной вклад внесли Yuxuan Zhang и Huan Ling. Для Yuxuan Zhang это третья работа, за полгода до нее он так же совместно с Huan Ling и другими опубликовал статью о [дифференцируемом рендеринге 3D моделей из двумерных изображений](#), где StyleGAN использовался для генерации синтетических изображений с разных ракурсов из исходного. Эта работа была опубликована на ICLR 2021.



Результаты работы модели предыдущей статьи авторов

Huan Ling хорошо известен благодаря моделям для сегментации изображений. Таким, как Polygon-RNN++ и Curve-GCN. Так что для основных вкладчиков в рассматриваемую статью работа является логичным продолжением предыдущих наработок.

Прорывом этой работы является возможность генерации данных с ручной разметкой всего десятка - сотни экземпляров, что делает модель более близкой к unsupervised методам, чем к классическим semi-supervised решениям данной задачи. В качестве бейзлайна авторы статьи брали работу [Semi-Supervised Semantic Segmentation with High- and Low-level Consistency](#) 2019 года. Особенностью этой работы является лучшее качество модели на меньших количествах размеченных данных, по сравнению с предыдущими работами, которые значительно теряют в качестве при их уменьшении. В статье используются модели Semi-Supervised Semantic Segmentation GAN (s4GAN) и Multi-Label Mean Teacher (MLMT). s4GAN используется для улучшения низкоуровневых деталей, а MLMT выполняет semi-supervised мультиклассификацию, чтобы использовать информацию на уровне класса для удаления ложноположительных прогнозов с карты сегментации. В исходной работе карты признаков из StyleGAN используются для построения репрезентативных векторов каждого пикселя изображения, после чего ансамбль из MLP классификаторов учится по вектору пикселя определять класс сегментации.



Архитектура Semi-Supervised Semantic Segmentation

Статья цитируется в работах, использующих синтетические данные, сгенерированные генеративно-состязательными нейросетями. Среди прочих можно упомянуть [Semantic Segmentation with Generative Models](#), в которой используется semi-supervised подход для сегментации изображений, ключевой особенностью этой работы является высокая обобщающая способность моделей, т.е. возможность менять контекст задачи. В работе используется модифицированный StyleGAN, который на выходе дополнительно выдает маски сегментаций, кроме того, градиенты из дискриминатора в этом GAN не передаются в область синтеза изображений генератора. Сравнений с DatasetGAN авторы не приводят. Из конкурентов существует работа [Segmentation in Style](#), которая позволяет генерировать карты сегментации изображений совсем без какой-либо первичной разметки. Они показывают state-of-the-art результат на наборе данных CelebA-Mask8, обгоняя DatasetGAN, однако по другим наборам данным результаты не предоставлены, из чего можно заключить, что данная работа не является полноценным конкурентом исходной.

В индустрии работа может быть использована для генерации 3d моделей городов, детектирования опухолей на медицинских снимках и других подобных задачах.