

DatasetGAN

Ознобихин Арсений

Постановка задачи

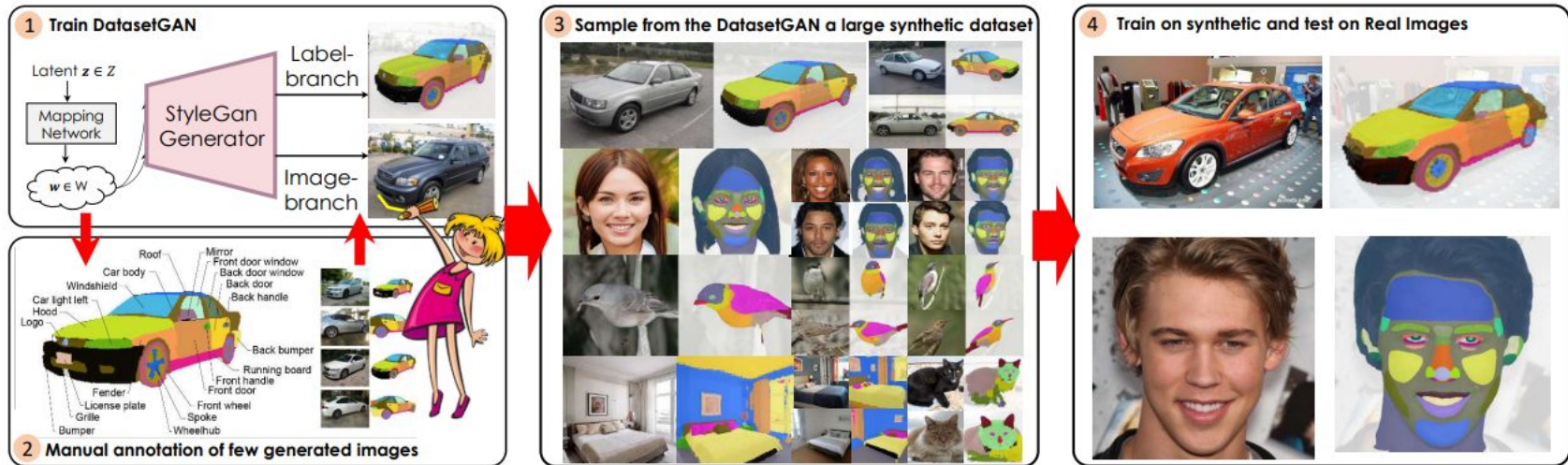
Проблема:

Хотим быстро получать размеченный датасет для pixel-wise задач с минимальными усилиями человека.

Решение:

Применим генеративные модели.

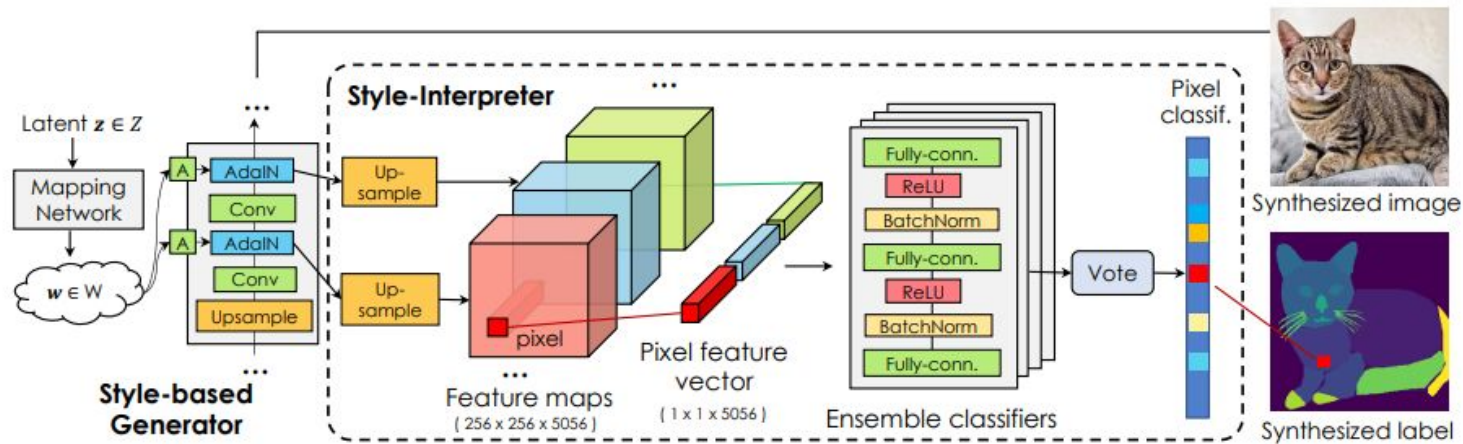
В чем идея?



Давайте попробуем создать автоматический генератор размеченных датасетов на основе StyleGAN.

В статье рассматриваются датасеты для pixel-wise сегментации и выделения точек. Для создания каждого датасета используется не более 40 размеченных изображений.

Извлечение маски сегментации. Style-interpreter

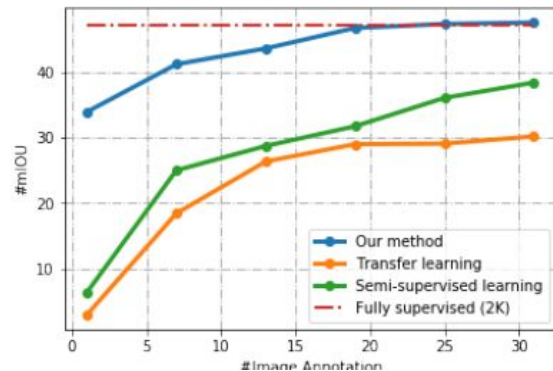


Вывод уровней с AdaIN апсемплится до максимального разрешения, разворачиваются в вектор, конкатенируются и подаются на вход в несколько MLP (в статье их 10), которые голосуют за маску пикселя. При этом, веса в MLP для всех пикселей одинаковые.

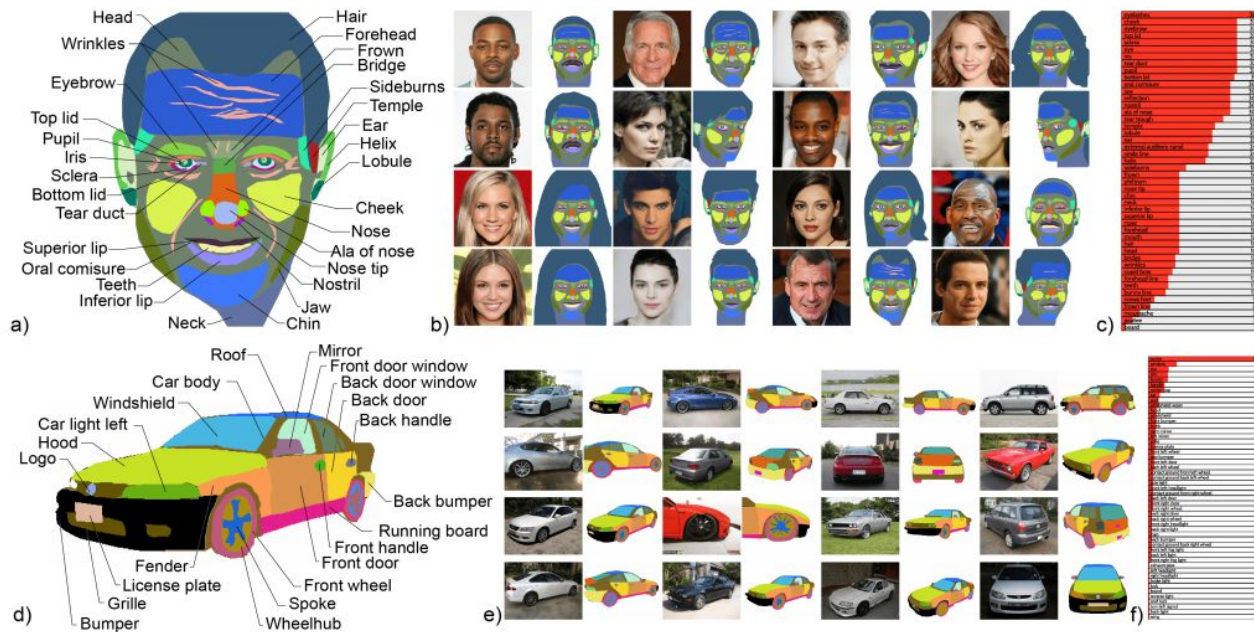
Обучение. Алгоритм

1. Обучается StyleGAN (на неразмеченном датасете).
2. StyleGAN генерирует несколько изображений, которые размечают вручную.
3. На размеченных изображениях обучают Style-interpreter.

Стоит понимать, что если учить Style-interpreter на всех пикселях изображения, то это выйдет крайне затратно. Поэтому для каждого изображения пиксели, на которых будет учиться Style-interpreter выбираются случайно (при этом гарантируется, что каждой значение маски будет представлено в такой выборке).



Обучение. Размеченный датасет



Примеры созданных датасетов



Результаты

Testing Dataset	ADE-Car-12	ADE-Car-5	Car-20	CelebA-Mask-8 (Face)	Face-34	Bird-11	Cat-16	Bedroom-19
Num of Training Images	16	16	16	16	16	30	30	40
Num of Classes	12	5	20	8	34	11	16	19
Transfer-Learning	24.85	44.92	33.91 \pm 0.57	62.83	45.77 \pm 1.51	21.33 \pm 1.32	21.58 \pm 0.61	22.52 \pm 1.57
Transfer-Learning (*)	29.71	47.22	\times	64.41	\times	\times	\times	\times
Semi-Supervised [41]	28.68	45.07	44.51 \pm 0.94	63.36	48.17 \pm 0.66	25.04 \pm 0.29	24.85 \pm 0.35	30.15 \pm 0.52
Semi-Supervised [41] (*)	34.82	48.76	\times	65.53	\times	\times	\times	\times
Ours	45.64	57.77	62.33 \pm 0.55	70.01	53.46 \pm 1.21	36.76 \pm 2.11	31.26 \pm 0.71	36.83 \pm 0.54

\times means that the method does not apply to this setting due to missing labeled data in the domain.

Testing Dataset	Car-20				CUB-Bird			
Metric	L2 Loss \downarrow	PCK th-15 \uparrow	PCK th-10 \uparrow	PCK th-5 \uparrow	L2 Loss \downarrow	PCK th-25 \uparrow	PCK th-15 \uparrow	PCK th-10 \uparrow
Transfer Learn.	4.4×10^{-4}	43.54	36.66	18.53	5.3×10^{-4}	23.17	18.21	12.74
Ours	2.4×10^{-4}	79.91	67.14	35.17	4.3×10^{-4}	60.61	46.36	32.00
Fully Sup.	\times	\times	\times	\times	3.2×10^{-4}	77.54	65.00	53.73

Table 1: Comparisons on Part Segmentation. (*) denotes In-domain experiment, where training and testing are conducted on the same dataset but a different split. Otherwise, training is conducted on our generated images. Note that In-domain setting does not apply to our approach, as we do not train StyleGAN on the provided datasets.

Table 2: Comparisons on Keypoint Detection. Our method leads to significantly better results than those obtained by baseline methods.

Источники

1. <https://doi.org/10.48550/arXiv.2104.06490>