DatasetGAN

Шинкарев Елисей БПМИ 211

DatasetGAN*: an automatic procedure to generate massive datasets of high-quality semantically segmented images requiring minimal human effort*

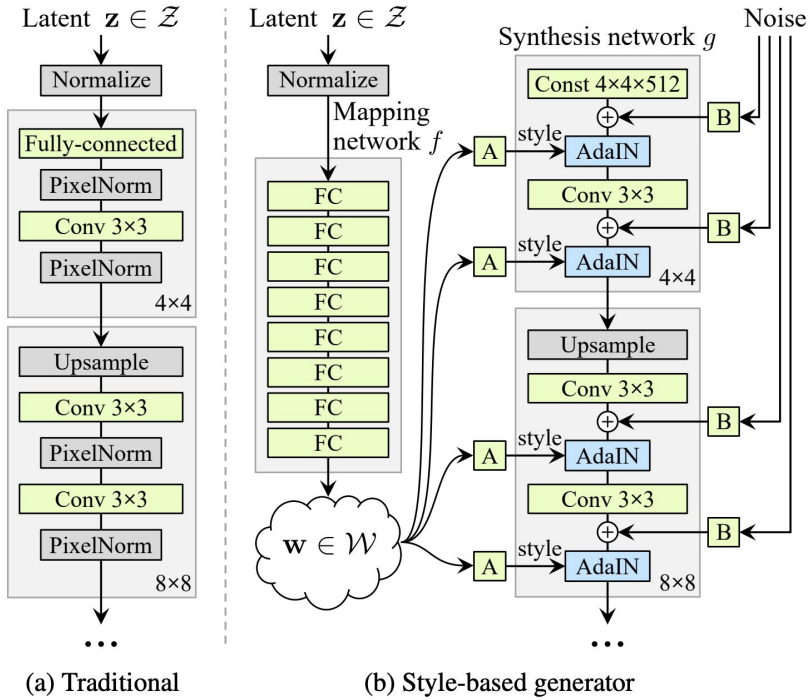
Training the decoder only needs a few labeled examples to generalize to the rest of the latent space, resulting in an infinite annotated dataset generator*!*

StyleGAN

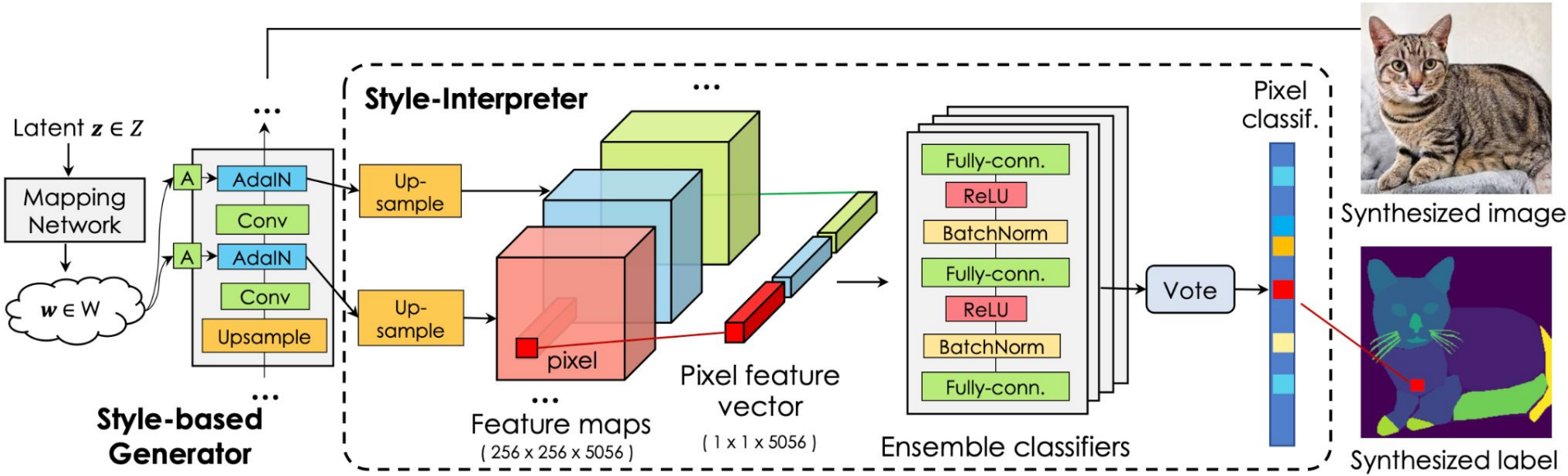
- Основная идея - вектор w отвечает за стиль, передаваемый в изображения

- Шум, пропускаемый через B отвечает за вариативность получаемых объектов

$$\mathrm{AdaIN}(\mathbf{x}_i, \mathbf{y}) = \mathbf{y}_{s,i} \frac{\mathbf{x}_i - \mu(\mathbf{x}_i)}{\sigma(\mathbf{x}_i)} + \mathbf{y}_{b,i}$$

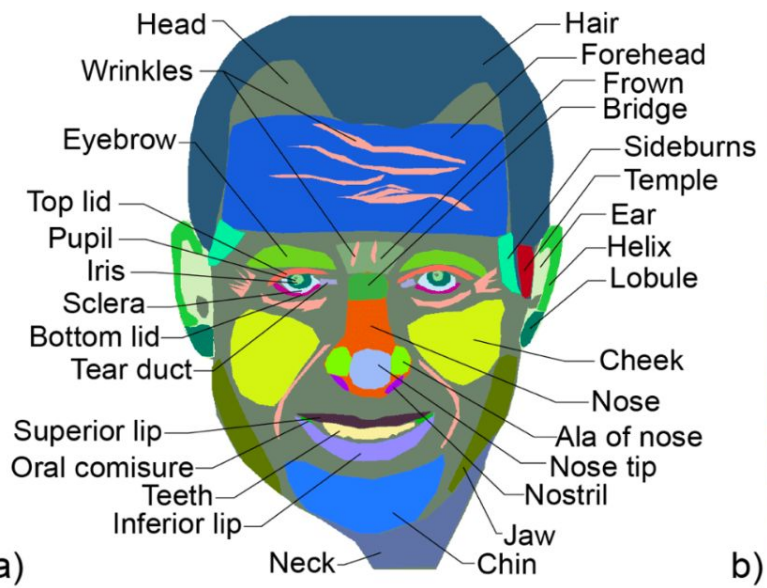


DatasetGAN



- MLP classifier - обучается на ручной разметке сгенерированных картинок. Cross-entropy loss.
- AdaIN теперь не константное преобразование вектора, а Style-Interpreter

Figure 3: Small human-annotated face datasets



a) shows an example of segmentation mask and associated labels

b) shows the full collection of training images (GAN samples)

c) shows the list of annotated parts and the number of instances in the dataset

c)

eyelashes	38
cheek	32
eyebrow	32
top lid	32
sclera	32
eye	32
iris	32
tear duct	32
pupil	32
botton lid	31
oral comisure	30
jaw	30
reflection	30
nostril	30
ala of nose	30
tear trough	28
temple	26
lobule	25
ear	25
external auditory canal	24
smile line	24
helix	23
sideburns	18
frown	16
philtrum	16
nose tip	16
chin	16
neck	16
inferior lip	16
superior lip	16
nose	16
forehead	16
mouth	16
hair	16
head	16
bridge	15
wrinkles	15
cupid bow	14
forehead line	13
teeth	13
bunny line	12
crows feet	9
frown line	8
moustache	5
goatee	3
beard	1

Дивергенция Йенсона-Шеннона

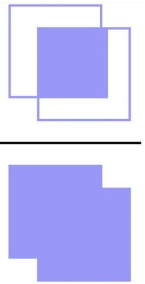
$$JS(p,q) = \frac{1}{2}\left(KL(p(x)||\frac{p(x)+q_\theta(x)}{2}) + KL(q_\theta(x)||\frac{p(x)+q_\theta(x)}{2})\right)$$

- симметрична
- ограничена $0 \leq JS(P,Q) \leq \ln(2)$
- можно превратить в расстояние $\sqrt{JS(p,q)}$

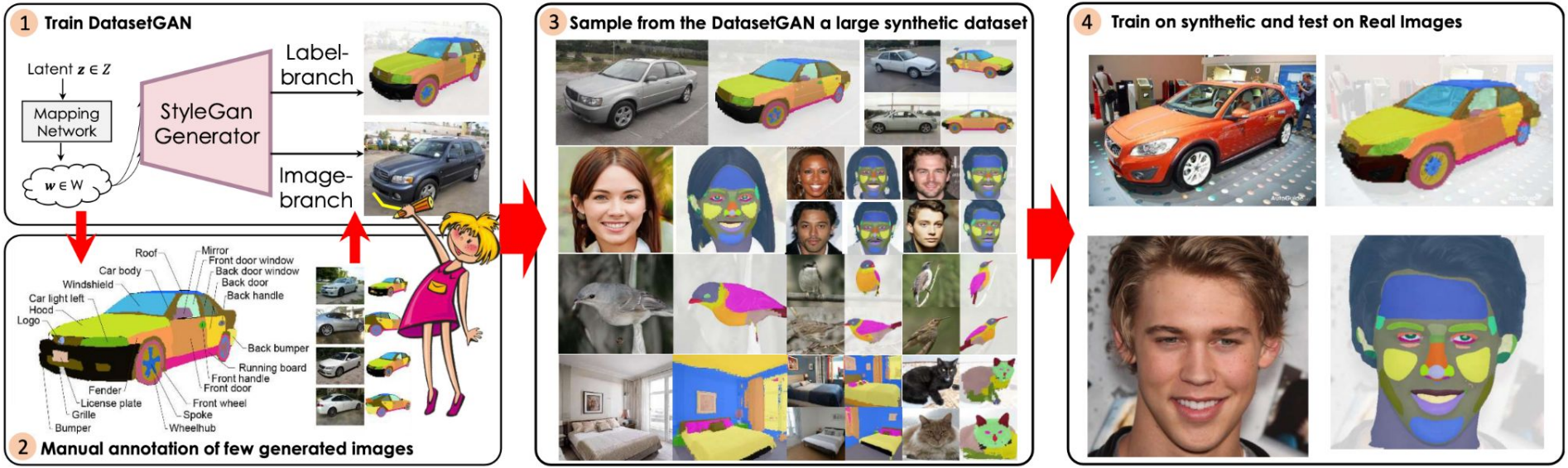
Filtering Ratio	0%	5%	10%	20%
mIOU	44.60	44.89	**45.64**	45.18

Table 4: **Ablation study of the filtering ratio.** We filter out the most uncertain synthesized Image-Annotation pairs. Result shown are reported on ADE-Car-12 test set, using the generated dataset of size 10k. We use 10% in other experiments.

mean Intersection over Union: $\text{IoU} = \dfrac{\text{Area of Overlap}}{\text{Area of Union}}$



Summarization



Results

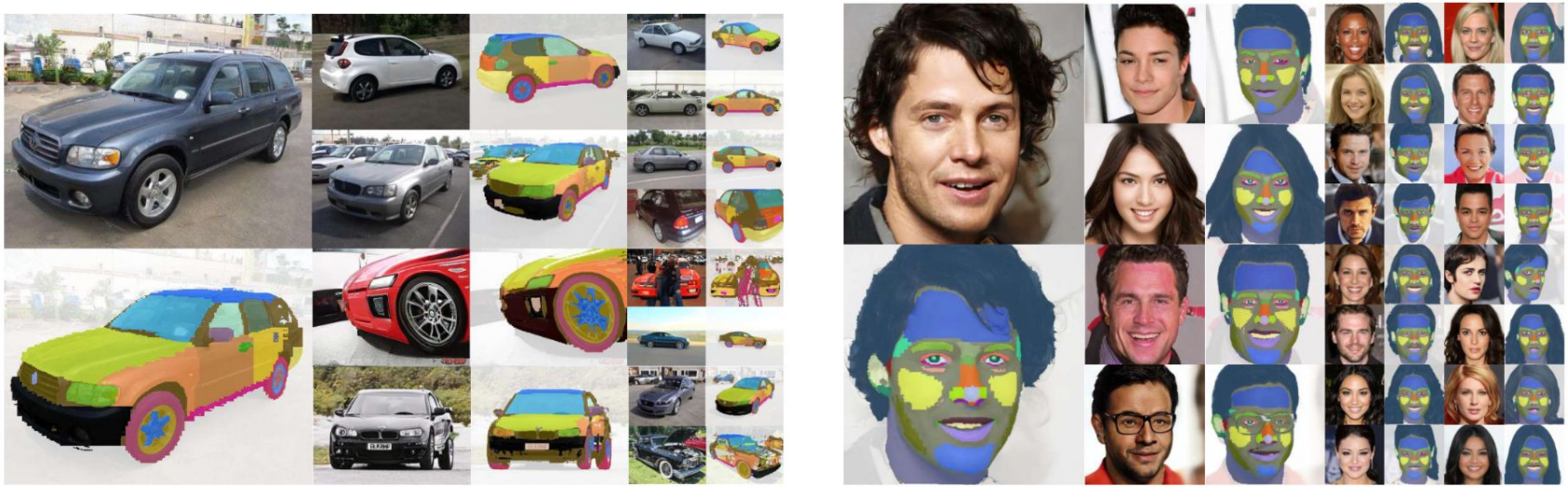


Figure 4: Examples of synthesized images and labels from our DATASETGAN for faces and cars. StyleGAN backbone was trained on CelebA-HQ (faces) on 1024 × 1024 resolution images, and on LSUN CAR (cars) on 512 × 384 resolution images. DATASETGAN was trained on 16 annotated examples.

Results

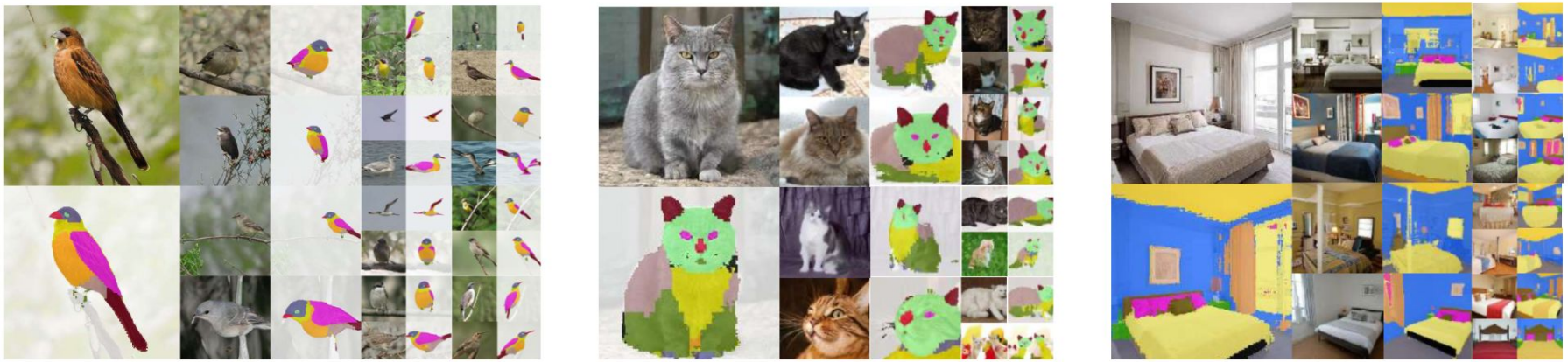


Figure 5: Examples of synthesized images and labels from our DATASETGAN for birds, cats, bedrooms. StyleGAN was trained on NABirds (1024×1024 images), LSUN CAT (256×256), and LSUN Bedroom (256×256). DATASETGAN was trained on 30 annotated bird examples, 30 cats, and 40 bedrooms.

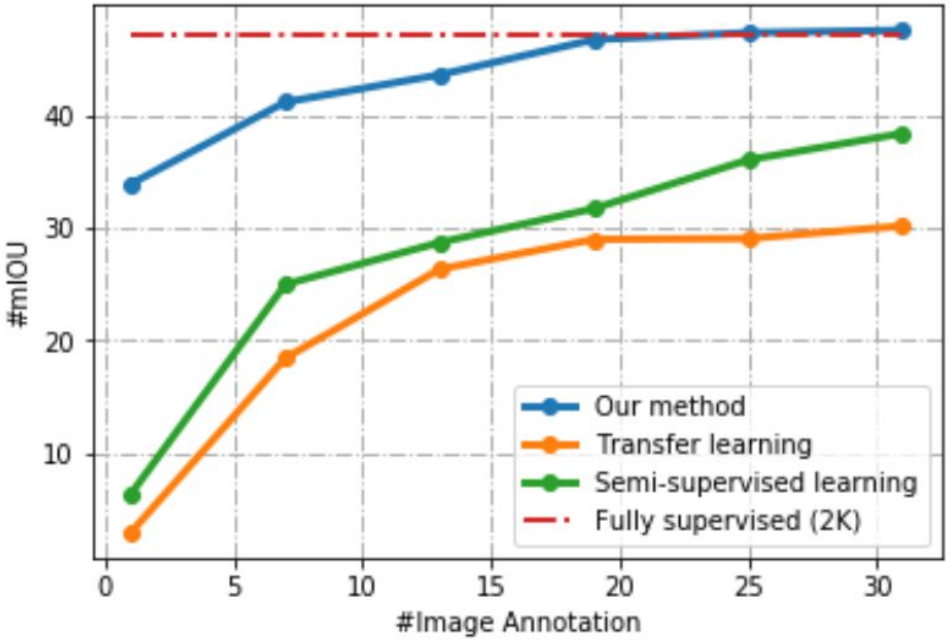


Figure 6: **Number of training examples vs. mIOU**. We compare to baselines on ADE-Car-12 testing set. The red dash line represents the fully supervised method which exploits 2.6k training examples from ADE20k.

Transfer learning - MS-COCO (Microsoft common objects in context)

Deeplab-V3, with ImageNet pre-trained ResNet151

Testing Dataset	Car-20				CUB-Bird			
Metric	L2 Loss ↓	PCK th-15 ↑	PCK th-10 ↑	PCK th-5 ↑	L2 Loss↓	PCK th-25 ↑	PCK th-15 ↑	PCK th-10 ↑
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.	✗	✗	✗	✗	3.2×10^{-4}	77.54	65.00	53.73

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 ± 0.57	62.83	45.77 ± 1.51	21.33 ± 1.32	21.58 ± 0.61	22.52 ± 1.57
Transfer-Learning (*)	29.71	47.22	✗	64.41	✗	✗	✗	✗
Semi-Supervised [41]	28.68	45.07	44.51 ± 0.94	63.36	48.17 ± 0.66	25.04 ± 0.29	24.85 ± 0.35	30.15 ± 0.52
Semi-Supervised [41] (*)	34.82	48.76	✗	65.53	✗	✗	✗	✗
Ours	**45.64**	**57.77**	**62.33 ± 0.55**	**70.01**	**53.46 ± 1.21**	**36.76 ± 2.11**	**31.26 ± 0.71**	**36.83 ± 0.54**

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

(*) denotes In-domain experiment, where training and testing are conducted on the same dataset but a different split