



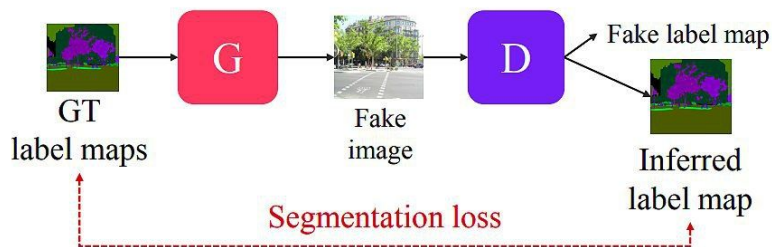
# **Generator Knows What Discriminator Should Learn in Unconditional GANs**



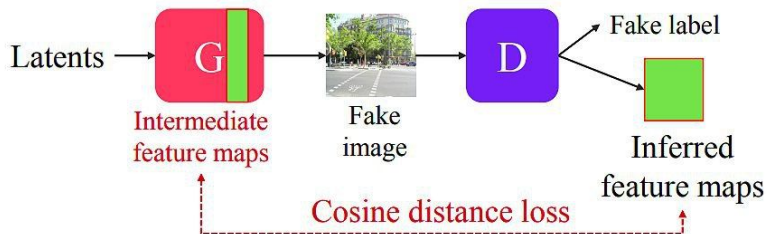
# Идея

- Используем GAN для генерации изображений
- Хотим улучшить качество путем изменения работы дискриминатора
- Вдохновляемся методом из условного генерирования
- Модифицируем и применяем

# Условная и безусловная генерация



**OASIS (Conditional image generation)**



**GGDR (Ours, Unconditional image generation)**



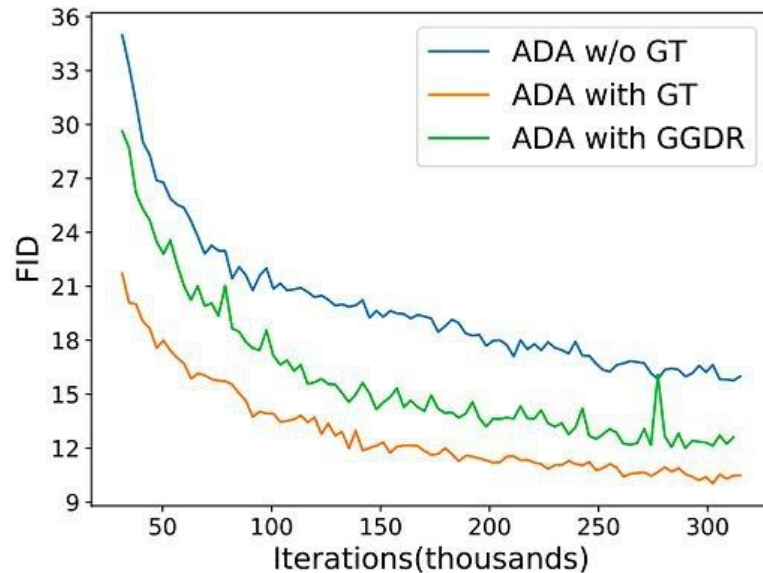
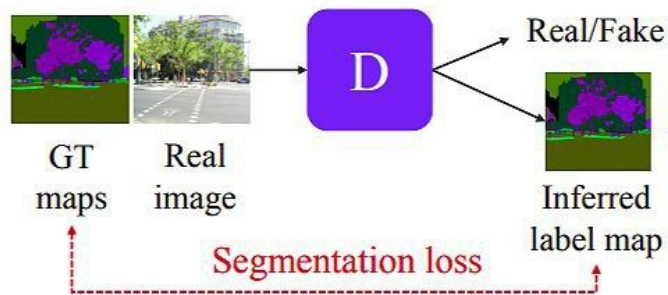
# План

- Показать, что использование карт семантических признаков в дискриминаторе позволяет добиться лучшего качества для безусловной генерации
- Показать, что карты признаков генератора являются альтернативой карте семантических признаков
- Предложение дискриминатора, регулируемого картой признаков генератора
- Демонстрация его эффективности

# Добавление семантических карт в обучение дискриминатора

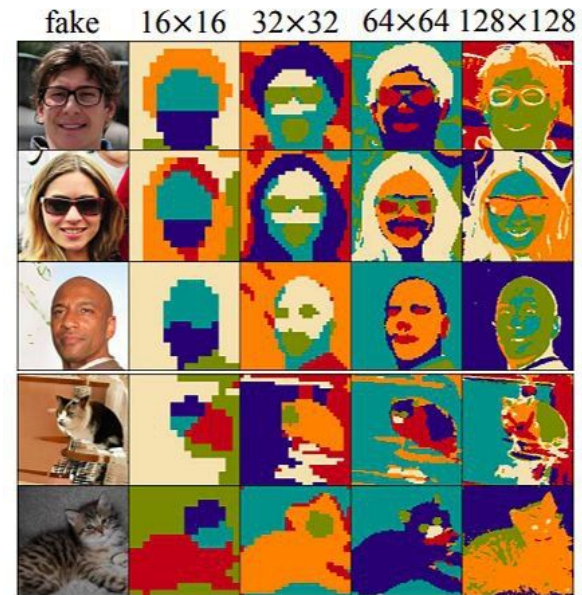
Генерация: StyleGan2 + adaptive discriminator  
augmentation на датасете ADE20K

Модифицированная задача дискриминатора:



# Визуализация карт признаков генератора

- при использовании карт семантических признаков дискриминатор опирается на осмысленные представления
- карты признаков генератора также их содержат
- можем использовать их



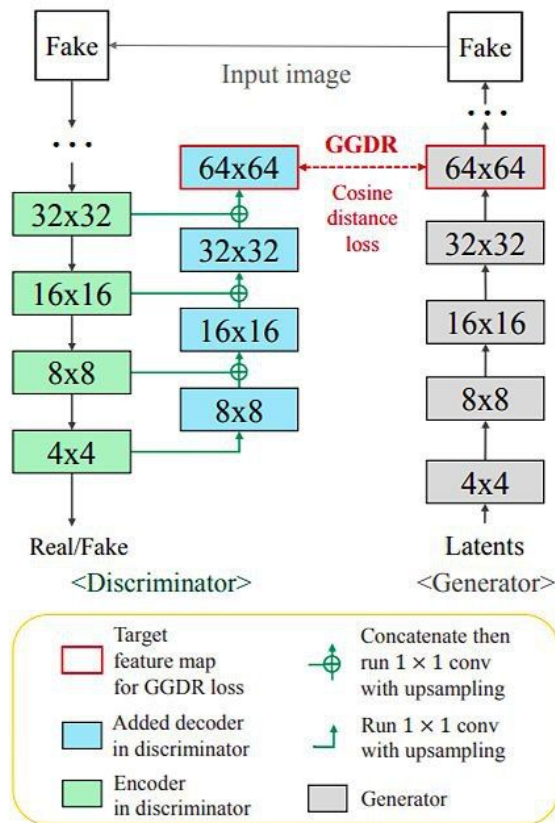
(a) Generator feature map visualization

# Обучение

$$\min_G \max_D \mathcal{L}_{adv}(G, D) = \mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

$$\max_D \mathcal{L}_{ggdr}(G, D) = - \mathbb{E}_{\mathbf{z} \sim p(\mathbf{z})} \left[ 1 - \frac{F^t(G(\mathbf{z})) \cdot G(\mathbf{z})^t}{\|F^t(G(\mathbf{z}))\|_2 \cdot \|G(\mathbf{z})^t\|_2} \right]$$

$$\mathcal{L}_{total} = \mathcal{L}_{adv} + \lambda_{reg} \mathcal{L}_{ggdr}$$



## Результаты

**Table 1.** FID scores of ours and comparison methods on FFHQ. We run three training for each data and show their means and standard deviations. The numbers are largely brought from ADA [26] and we follow their evaluation protocol.

| FFHQ      | 2k                | 10k              | 140k             |
|-----------|-------------------|------------------|------------------|
| PA-GAN    | 56.49±7.28        | 27.71±2.77       | 3.78±0.06        |
| WGAN-GP   | 79.19±6.30        | 35.68±1.27       | 6.43±0.37        |
| zCR       | 71.61±9.64        | 23.02±2.09       | 3.45±0.19        |
| AR        | 66.64±3.64        | 25.37±1.45       | 4.16±0.05        |
| StyleGAN2 | 78.80±2.31        | 30.73±0.48       | 3.66±0.10        |
| +GGDR     | 70.59±5.16        | 24.44±0.63       | <b>3.14±0.03</b> |
| ADA       | <b>16.49±0.65</b> | 8.29±0.31        | 3.88±0.13        |
| +GGDR     | 18.28±0.77        | <b>6.11±0.15</b> | 3.57±0.10        |

**Table 2.** FID scores of ours and comparison methods on CIFAR-10. We run three training for mean and standard deviations. We brought the numbers of diffusion models from [54].

| CIFAR-10  | FID              | IS                |
|-----------|------------------|-------------------|
| ProGAN    | 15.52            | 8.56±0.06         |
| AutoGAN   | 12.42            | 8.55±0.10         |
| StyleGAN2 | 8.32±0.09        | 9.21±0.09         |
| ADA       | 2.92±0.05        | 9.83±0.04         |
| FSMR      | 2.90             | 9.68              |
| DDPM      | 3.17±0.05        | 9.46±0.11         |
| NCSN++    | 2.2              | 9.89              |
| ADA+GGDR  | <b>2.15±0.02</b> | <b>10.02±0.06</b> |



# Результаты

**Table 3.** Comparison on FFHQ, LSUN Cat, LSUN Horse and LSUN Church. Our method improves StyleGAN2 [29] in large datasets in terms of FID and recall. P and R denote precision and recall. Lower FID and higher precision and recall mean better performance. The bold numbers indicate the best FID, P, R for each dataset.

| Method        | FFHQ        |             |             | LSUN Cat    |             |             | LSUN Horse  |             |             | LSUN Church |             |             |
|---------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|               | FID↓        | P↑          | R↑          | FID         | P           | R           | FID         | P           | R           | FID         | P           | R           |
| UT [4]        | 6.11        | <b>0.73</b> | 0.48        | -           | -           | -           | -           | -           | -           | 4.07        | <b>0.71</b> | 0.45        |
| Polarity [19] | -           | -           | -           | 6.39        | <b>0.64</b> | 0.32        | -           | -           | -           | 3.92        | 0.61        | 0.39        |
| StyleGAN2     | 3.71        | 0.69        | 0.44        | 7.98        | 0.60        | 0.27        | 3.62        | 0.63        | 0.36        | 3.97        | 0.59        | 0.39        |
| +GGDR         | <b>3.14</b> | 0.69        | <b>0.50</b> | <b>5.28</b> | 0.58        | <b>0.38</b> | <b>2.50</b> | <b>0.64</b> | <b>0.43</b> | <b>3.15</b> | 0.61        | <b>0.46</b> |

**Table 4.** Comparison on AFHQ Cat, Dog, Wild and Landscape. Our method improves ADA [26] in small datasets in terms of FID and recall. P and R denote precision and recall. The bold numbers indicate the best FID, P, and R of the models.

| Method       | AFHQ Cat    |             |             | AFHQ Dog    |             |             | AFHQ Wild   |             |             | Landscape    |             |             |
|--------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|
|              | FID↓        | P↑          | R↑          | FID         | P           | R           | FID         | P           | R           | FID          | P           | R           |
| FastGAN [40] | 4.69        | <b>0.78</b> | 0.31        | 13.09       | 0.75        | 0.38        | 3.14        | 0.76        | 0.20        | 16.44        | <b>0.77</b> | 0.16        |
| ContraD [23] | 3.82        | -           | -           | 7.16        | -           | -           | 2.54        | -           | -           | -            | -           | -           |
| ADA          | 3.55        | 0.77        | 0.41        | 7.40        | 0.76        | 0.48        | 3.05        | 0.76        | 0.13        | 13.87        | 0.72        | 0.20        |
| +GGDR        | <b>2.76</b> | 0.74        | <b>0.52</b> | <b>4.59</b> | <b>0.79</b> | <b>0.53</b> | <b>2.06</b> | <b>0.80</b> | <b>0.27</b> | <b>10.38</b> | 0.69        | <b>0.29</b> |

# Результаты

| Target         | FID         | Activation   | FID         |
|----------------|-------------|--------------|-------------|
| None           | 7.98        | Linear       | <b>5.28</b> |
| $8 \times 8$   | 7.57        | leaky ReLU   | 5.43        |
| $16 \times 16$ | 6.56        | Kernel size  | FID         |
| $32 \times 32$ | 5.98        | $1 \times 1$ | 5.28        |
| $64 \times 64$ | <b>5.28</b> | $3 \times 3$ | <b>5.25</b> |

(a) Target size

(b) Decoder design

| Method   | # params | time(s) |
|----------|----------|---------|
| Baseline | 4.87M    | 5.60    |
| + GGDR   | 5.05M    | 6.05    |
|          | (+3.7%)  | (+8.0%) |

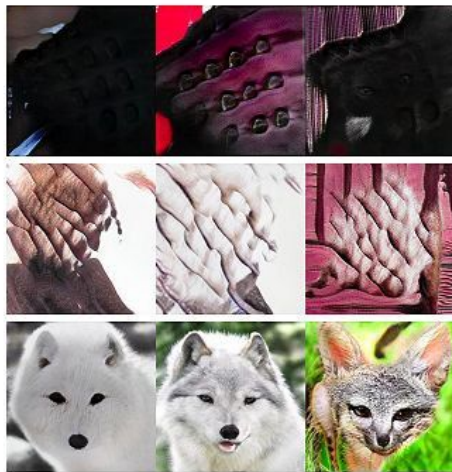
(c) Calculation costs

# Примеры



# Примеры

ADA

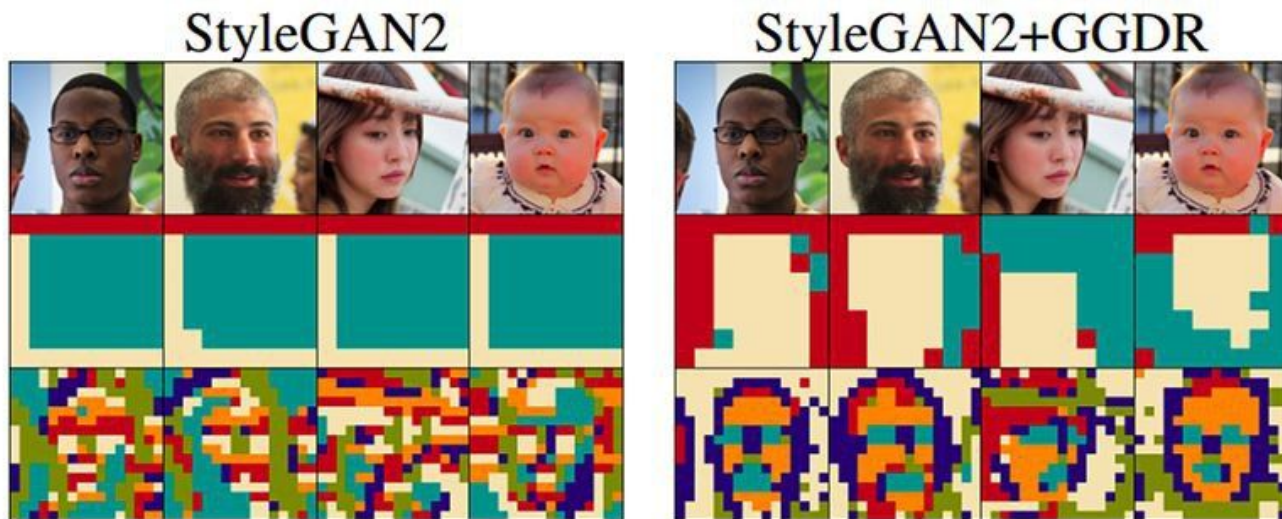


ADA+GGDR



(a) Worst sample comparison

## Карты признаков энкодера в дискриминаторе



(b) Comparison of the the encoder feature maps in the discriminators

# Generator Knows What Discriminator Should Learn in Unconditional GANs

рецензент-исследователь

# Что происходит в области

Изначально была ОНА -  
первая статья про GAN

Справка для  
интересующихся:

Вышла в 2014 году на конфе  
NIPS исследователями из  
Монреальского университета





# Что происходит в области

## GAN Basics

Generative Adversarial Networks with codes

| GAN model      | Title  | conference     | tensorflow code            | pytorch code   |
|----------------|--|----------------|----------------------------|--|
| DCGAN          | Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks | ICLR 2016      |                            | <a href="#">pytorch example</a>                              |
| WGAN           | Wasserstein GAN  |                |                            | <a href="#">official</a>                                     |
| WGAN-GP        | Improved Training of Wasserstein GANs  | NIPS 2017      | <a href="#">official</a>   | <a href="#">unofficial</a>                                   |
| BEGAN          | BEGAN: Boundary Equilibrium Generative Adversarial Networks                                  | arXiv 2017     | <a href="#">unofficial</a> | <a href="#">unofficial</a>                                   |
| pix2pix        | Image-to-Image Translation with Conditional Adversarial Networks                             | CVPR 2017      |                            | <a href="#">official</a>                                     |
| CycleGAN       | Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks              | ICCV 2017      |                            | <a href="#">official</a>                                     |
| GauGAN         | Semantic Image Synthesis with Spatially-Adaptive Normalization                               | CVPR 2019 oral |                            | <a href="#">official</a>                                     |
| ProgressiveGAN | Progressive Growing of GANs for Improved Quality, Stability, and Variation                   | ICLR 2018      | <a href="#">official</a>   | <a href="#">unofficial1</a> ,<br><a href="#">unofficial2</a> |
| StyleGAN       | A Style-Based Generator Architecture for Generative Adversarial Networks                     | CVPR 2019      | <a href="#">official</a>   | <a href="#">unofficial1</a> ,<br><a href="#">unofficial2</a> |

Дальше пошел ряд статей и раз-два в год выходили статьи, которые предлагали достаточно полезные улучшения, что их выделило сообщество



# Что происходит в области

А так на пике своей “популярности” на тему ганов выходило 100+ статей в год (2020 для примера). Из них большая часть была посредственной.

Топ статьи попадали на конференции

NIPS, ICML - Если у тебя general улучшение

CVPR - с уклоном в CV

(для примера списка: [github.com/ysbsb/awesome-gan-papers](https://github.com/ysbsb/awesome-gan-papers) , в основном отбираются по конфам, а там уже барьером выступают судьи)

# Так что же из себя представляет эта статья

- Вышла в июле этого года (на arXiv она пролежала меньше полугода!)
- В октябре прошла конференция ECCV 2022 (Европейская про CV), куда эта статья прошла (списки публиковались еще в начале лета)
- Конференция считается “хорошим выбором для публикации”, но на ступеньку ниже CVPR если говорить про Computer Vision

## Что можно сказать про статью

- Показали космические показатели (FID в некоторых местах улучшается до 20% по сравнению со StyleGan2 / StyleGan3)

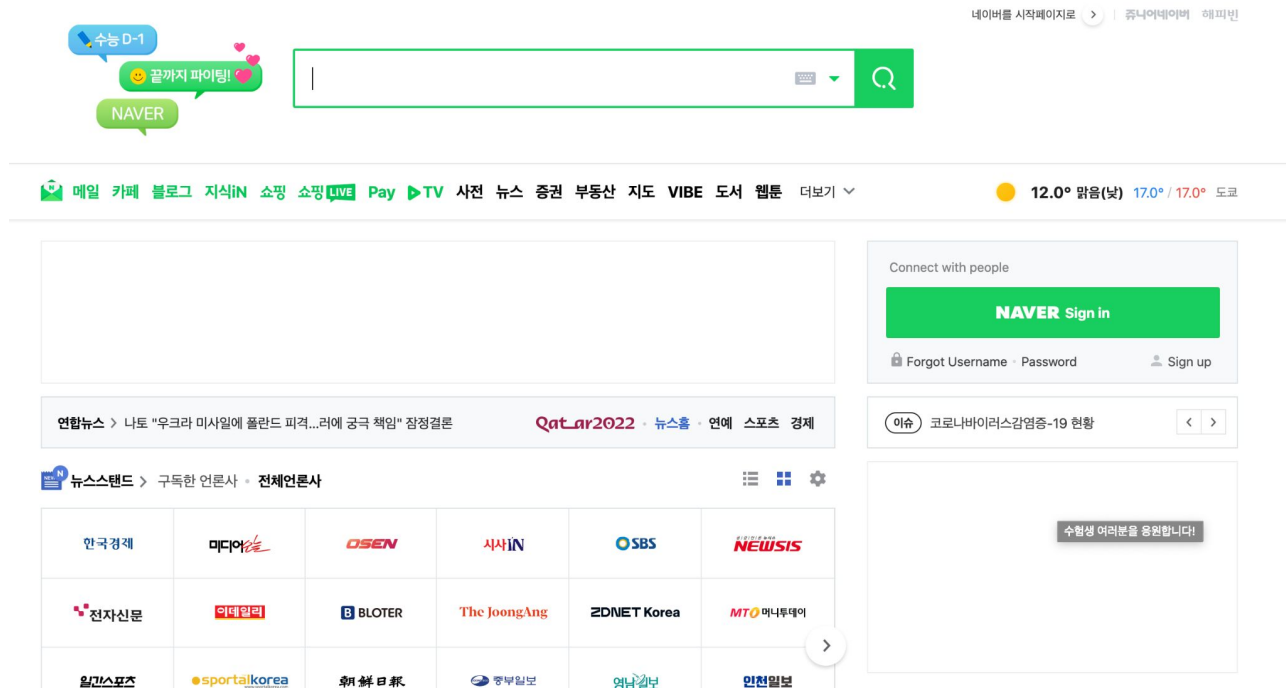
| Method        | FFHQ        |             |             |
|---------------|-------------|-------------|-------------|
|               | FID↓        | P↑          | R↑          |
| UT [4]        | 6.11        | <b>0.73</b> | 0.48        |
| Polarity [19] | -           | -           | -           |
| StyleGAN2     | 3.71        | 0.69        | 0.44        |
| +GGDR         | <b>3.14</b> | 0.69        | <b>0.50</b> |

# Что можно сказать про статью

Что было не супер (сравниваю со state of the art статьями на тему)

- Ablation studies слабый. Пробовали добавлять GGDR только к StyleGan2 и ADA.
- Не совсем понятно насколько хорошо давать информацию о внутренних представлениях дискриминатора -> генератору

# Что за люди это делали



Коротко - R&D team  
корейского Яндекса  
(поисковика под  
названием Naver)

Работа сразу с  
мыслями о статье.  
Наука на конвейере  
так сказать

Что за люди это делали



Чуть глубже - лаба подразделения Naver -> Naver Ai Labs

Сложно оценить размер вклада, которые сделал каждый из участников, но насколько понимаю распределение +- следующее

# Что за люди это делали

Основные лица:

Gayoung Lee (выложила код, больше всех рабочего опыта)

У последующих 3х есть опыт работы с ГАНами, так что идея должна была идти от них:

Junho Kim (давно работает в различных R&D)

Yunjey Choi (предположительно тимлид)

Hyunsu Kim (совсем молодой 2021 masters)

Начальник лабы:

Jung-Woo Ha (в 2018 году был соавтором большой статьи по ганам - StarGan, которая уже была сильно принята интернационально)

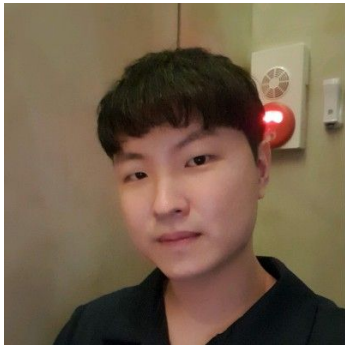
???:

Seonghyeon Kim (Исследователь из соседней лабы: работает в другом отделении корпорации. Не совсем понятен его вклад в работу, но он тоже здесь есть. Увлекается ганами, так что можно попытаться объяснить горизонтальной мобильностью)

# Что за люди это делали

Фоточки (вот они слева направо):

Hyunsu Kim



Junho Kim



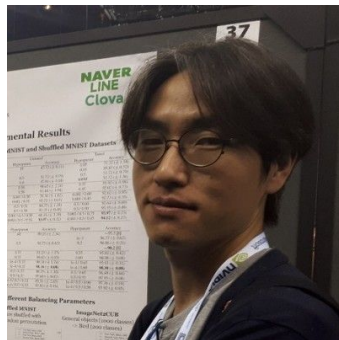
Yunjey Choi



Gayoung Lee



Jung-Woo Ha



Seonghyeon Kim



# Что эта статья вообще забыла у нас на НИИС?

Или почему эта статья олицетворяет очень плохой феномен в современной науке (Массовое производство статей. Действительно хорошие статьи теряются в серийниках). Почему я считаю, что рано ее выбрали на НИИС:

Молодые еще не показавшие себя авторы

Супер новая статья без какого либо взрыва в области

С тем же успехом можно было взять любую статью из 1640 в этом году на ECCV

Есть вероятность, что через год вспомнят про нее только такие же серийные статьи в разделе “Related Work”