

# Алгоритм пост-обработки синтетических объектов для аугментации изображений



Алексей Калинов

## Постановка задачи и данные

- Российская база автодорожных знаков (RTSD) [1] содержит 105509 изображений знаков, распределённых по 205 классам.
- Набор данных очень не сбалансирован. В обучающей выборке всего 106 классов.
- Перед обучением моделей необходимо добавлять новые данные для отсутствующих классов.
- Сгенерированные синтетические данные отличаются от реальных, поэтому необходима предобработка.

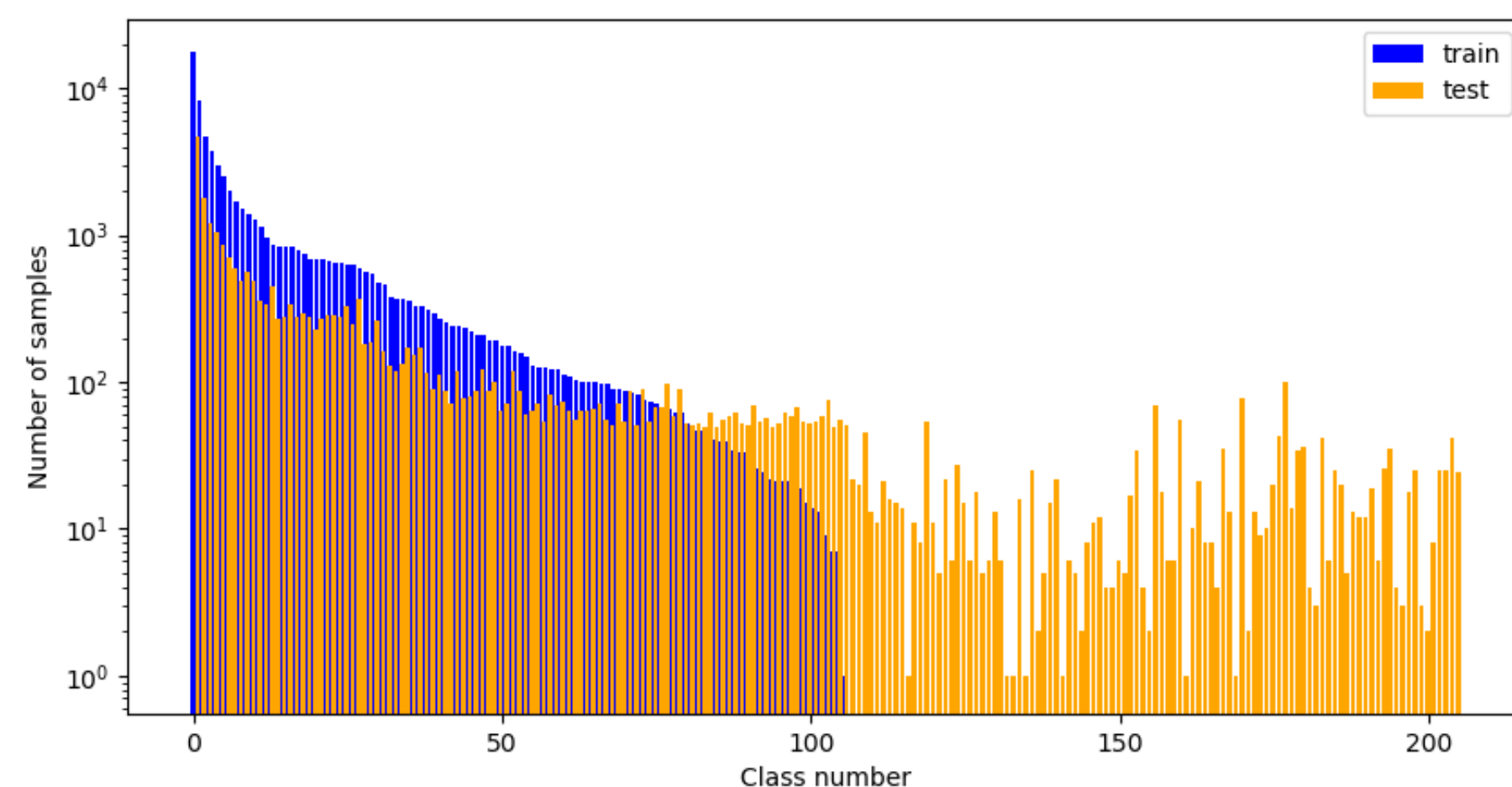


Рис. 1: Распределение изображений по классам в RTSD. Синим указаны объекты в обучающей выборке, оранжевым в тестовой.

## Обучение

- Большой разброс в количестве классов влияет на обучение CycleGAN; ромбовидные знаки переходят в круглые, белые в синие.
- Решение: уменьшить количество объектов в непропорционально больших классах.

## Методы решения задачи

- Проблему неравномерности классов можно попытаться решить путём генерации знаков с помощью компьютерной графики.
- Для улучшения качества искусственных картинок будем решать задачу непарного перевода из синтетического домена в реальный и обратно с помощью CycleGAN[2] и улучшения MUNIT[3].

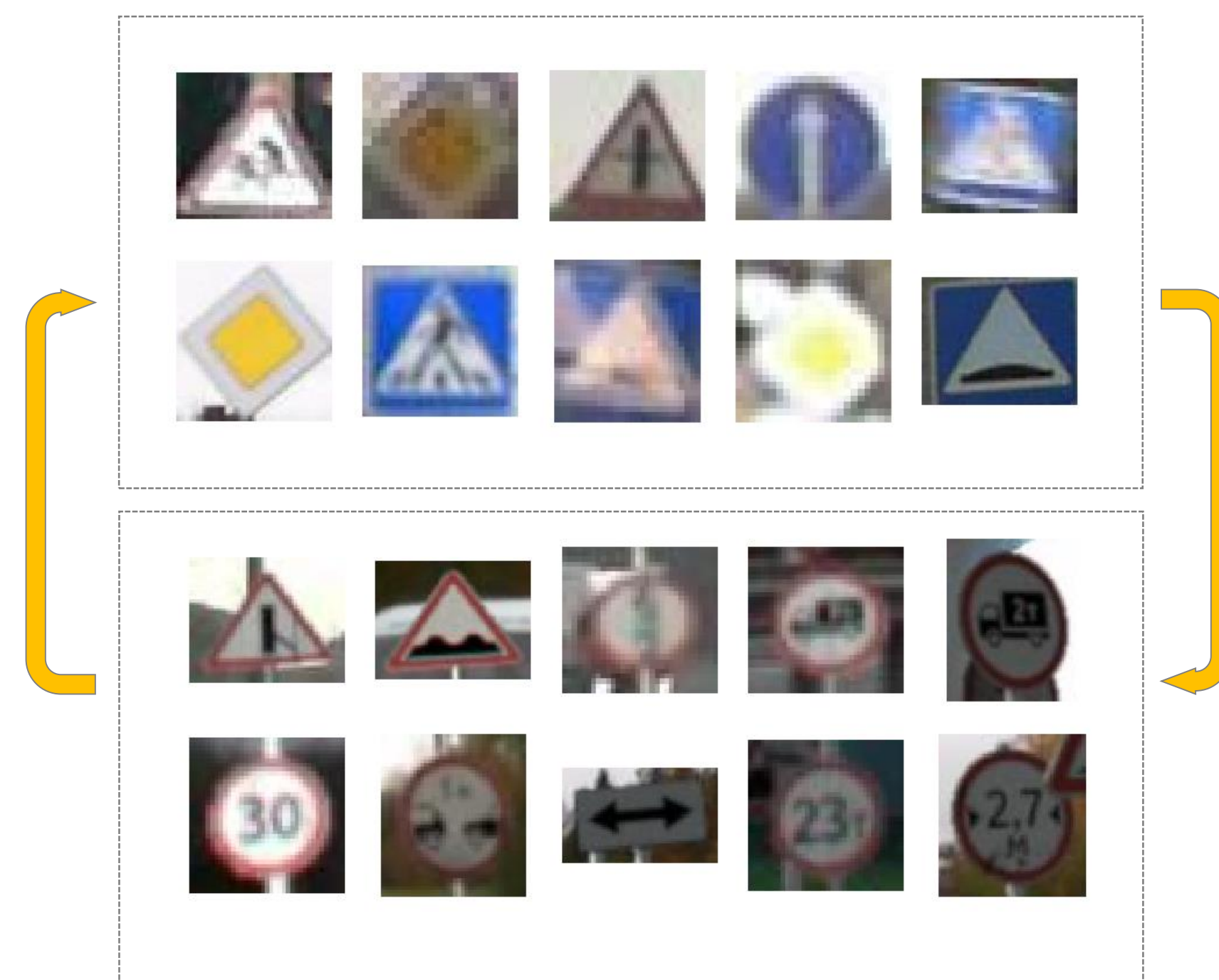


Рис. 2: Примеры реальных и синтетических изображений и схематический непарный перевод.

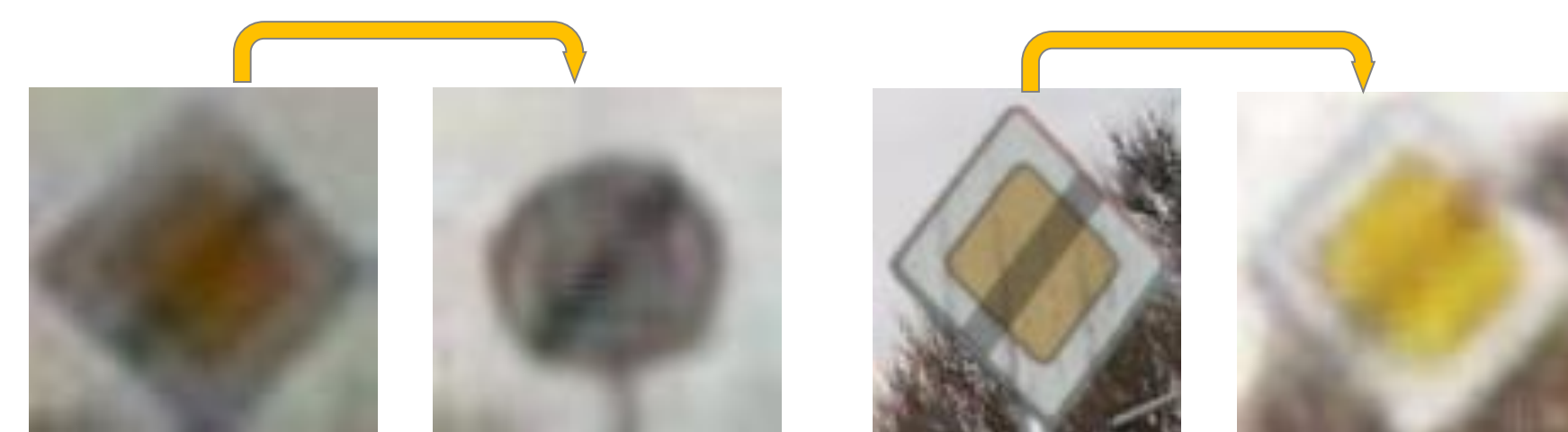


Рис. 3-6: Перевод до и после балансировки классов.

## Оценка качества

- Сравниваются результаты классификации моделями, обученными на улучшенных датасетах RTSD.
- Классификатор — WideResNet с глубиной блока 8 и двукратным увеличением числа фильтров.
- Основная метрика: доля правильных ответов (accuracy); измеряется на трех подмножествах:
  - Полное подмножество (**Full**).
  - Подмножество классов обучающей выборки (**Train classes**).
  - Подмножество классов, которых нет в обучающей выборке (**Test classes**).

## Текущие результаты

Доля правильных ответов при классификации знаков

	Full	Train classes	Test classes
Baseline	89.4%	95.4%	0.0%
CG Baseline	<b>92.9%</b>	<b>96.6%</b>	39.4%
Real2Synt	44.0%	44.1%	42.5%
Synt2Real	92.4%	95.3%	<b>50.7%</b>

## Выводы

- Больше данных, даже синтетических — лучше качество.
- Предварительная обработка искусственных данных, приводящая их к исходному домену, способна улучшить качество.

[1] Шахуро В., Конушин А. Российская база изображений автодорожных знаков. Компьютерная оптика, 2016

[2] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Computer Vision (ICCV), 2017 IEEE International Conference on, 2017.

[3] X. Huang, M.-Y. Liu, S. Belongie, and J. Kautz. Multimodal unsupervised image-to-image translation. In Proceedings of the European Conference on Computer Vision (ECCV), pages 172–189, 2018.