

Распознавание редких дорожных знаков
с использованием синтетических
обучающих выборок

Владислав Игоревич Шахуро
научный руководитель: к. ф.-м. н. А. С. Конушин



26 ноября 2020 г.

Распознавание дорожных знаков на изображениях



Рассматривается задача выделения и классификации дорожных знаков на изображении

Системы помощи водителю и управления беспилотным автомобилем



Обновление карт высокого разрешения



Актуальные проблемы

Современные методы распознавания объектов используют методы машинного обучения. Качество распознавания зависит от размера и разнообразия обучающей выборки.

1. Выборку данных большого размера сложно собрать и разметить.
2. Есть редко встречающиеся или новые классы объектов.



Поэтому современные алгоритмы распознают только часть существующих дорожных знаков. Данная работа фокусируется на проблеме распознавания редких дорожных знаков.

Синтез обучающих данных

Одним из возможных решений проблемы отсутствия эталонных данных является синтез данных. Синтез данных успешно используется для обучения некоторых алгоритмов компьютерного зрения.

1. Методы компьютерной графики



вычисление оптического потока



выделение автомобилей

2. Нейросетевые методы генерации изображений — перспективная и мало исследованная область синтеза обучающих выборок

Цель и задачи работы

Цель работы — повышение качества распознавания редких дорожных знаков за счет использования синтетических обучающих данных.

Задачи:

1. Собрать и разметить выборку кадров с дорожными знаками широкого класса с полной разметкой встречающихся классов знаков.
2. Исследовать применимость современных нейросетевых методов к задаче генерации изображений дорожных знаков для обучения классификатора.
3. Разработать метод условной нейросетевой генерации изображений дорожных знаков для обучения классификатора дорожных знаков.
4. Разработать метод повышения визуального реализма синтетических изображений редких дорожных знаков.
5. Разработать метод классификации редких дорожных знаков.
6. Разработать метод синтеза обучающих выборок для повышения качества детектирования редких дорожных знаков.

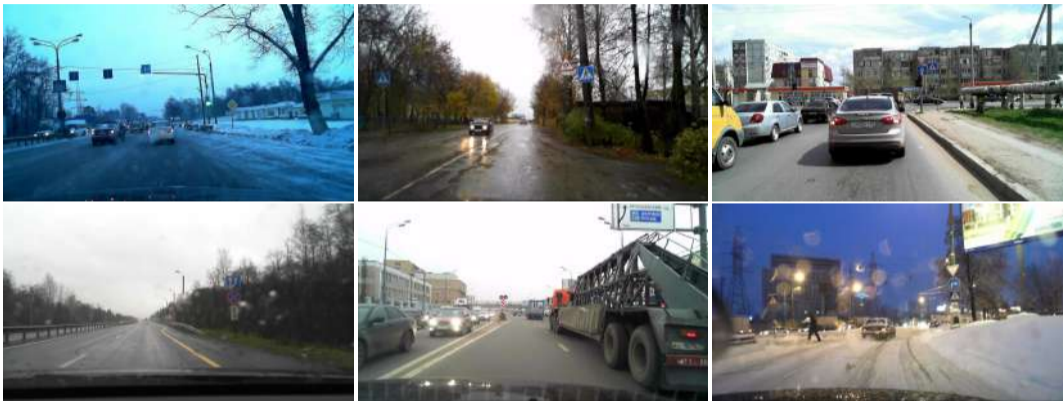
1. Выборка кадров с дорожными знаками

	GTSRB	GTSDDB	STS	BTSD	LISA	RTSD
Количество кадров	-	900	4000	25630	6610	179138
Количество знаков	43	43	7	108	47	205
Физических знаков	1728	1213	-	4565	-	15630
Изображений знаков	51839	1213	3488	13444	7855	104358

статистики публичных баз дорожных знаков

Существующие публичные выборки не подходят для обучения и оценки системы распознавания редких дорожных знаков (детектор и классификатор), поэтому была собрана база российских дорожных знаков (RTSD), подходящая для данной задачи.

Российская база автодорожных знаков

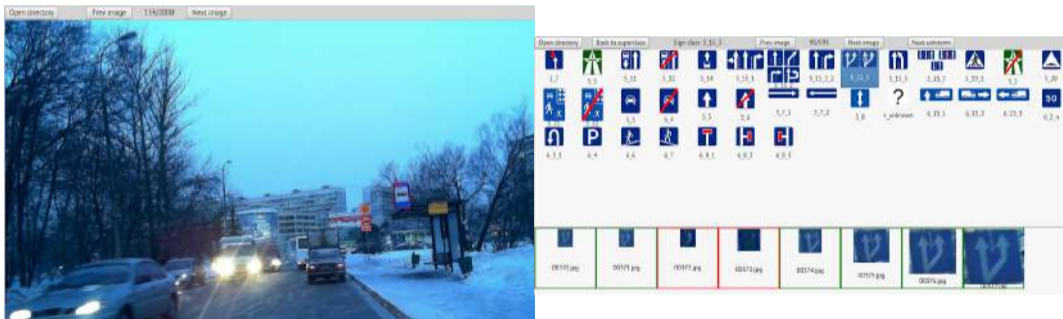


кадры из RTSD

получены с HD видеорегистраторов, установленных на автомобилях

Разметка данных

1. Ручная разметка с помощью 3 фрилансеров.
Трудоемкость — 500 человеко-часов.



2. Обучение детектора, анализ ошибок и доразметка выборки.
Добавлено 9 тысяч новых знаков.

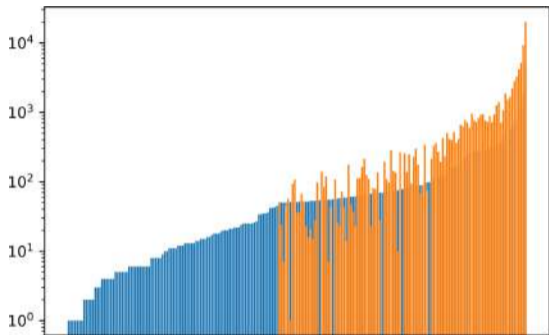
Разбиение выборки на тренировочную и тестовую части

Выборка делится в соотношении примерно 4 к 1.

В тестовую часть входят все 205 классов знаков.

Изображения одного экземпляра знака входят целиком либо в тестовую, либо в тренировочную часть.

В тренировочную часть вошло 106 классов. 99 классов — редкие знаки.



количество дорожных знаков
в **тренировочной** и **тестовой** части

2. Исследование применимости нейросетевых методов генерации изображений к дорожным знакам

Базовый метод

Генерация синтетических изображений дорожных знаков по иконке с помощью набора эвристических преобразований:



контраст



поворот



движение



размытие



фон

Получение более реалистичных синтетических изображений с помощью данного метода — очень трудоемкая задача.

Chigorin, Konushin. A system for large-scale traffic sign recognition and mapping. CMRT 2013

Порождающие состязательные сети

В последнее время активно развивается подход на основе порождающих состязательных нейросетей, которые позволяют генерировать реалистичные изображения в некоторых доменах (рукописные цифры, лица).

В работе исследуется метод обучения порождающих нейросетей на основе метрики Васерштейна:

$$W(p_r, p_g) = \max_{w \in \mathcal{W}} \mathbb{E}_{x \sim p_r(x)} [f_w(x)] - \mathbb{E}_{z \sim p(z)} [f_w(g_\theta(z))].$$

g_θ — нейросеть-генератор с настраиваемыми весами θ ,

f_w — нейросеть-критик с настраиваемыми весами w ,

x — реальные изображения,

$p_r(x)$ — распределение реальных изображений,

z — нормальный шум.

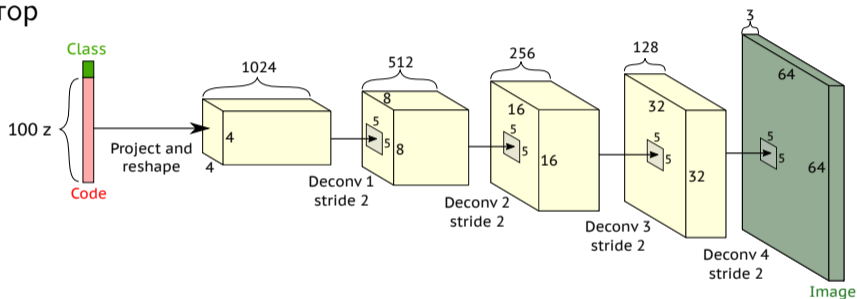
Синтетические выборки при обучении одного и нескольких поклассовых генераторов



В. Шахуро, А. Конушин. Синтез обучающих выборок для классификации знаков с помощью нейросетей. Компьютерная оптика 2018; 42(1)

3. Условная генерация изображений дорожных знаков с помощью порождающих нейросетей

Генератор



Критик

Сверточная нейросеть с двумя выходами: мера реалистичности и класс изображения

Odena et al. Conditional Image Synthesis With Auxiliary Classifier GANs. ICML 2017

Обучение

Функция потерь:

$$L = L_C + W(p_r, p_g) + L_R$$

Чтобы критик использовал метку класса, используется перекрестная энтропия:

$$L_C = E_{x \sim p_r(x)} \log p(c|x) + E_{z \sim p(z)} \log p(c|g_\theta(z))$$

При обучении на несбалансированной выборке каждому обучающему примеру с меткой y добавим вес $\frac{1}{k_p(y)}$, где k — количество объектов данного класса



сгенерированные
изображения

П. Хрушков, В. Шахуро, А. Конушин. Синтез изображений дорожных знаков с помощью условных порождающих противоборствующих нейросетей. ГрафиКон 2018

Экспериментальная оценка методов условной и безусловной генерации изображений дорожных знаков

	Тренировочная выборка			
	39 тыс. без размн.	215 тыс. без размн.	39 тыс. с размн.	215 тыс. с размн.
Реальные данные	96.6	—	98.4	—
WGAN синтетика	95.3	96.1	97.6	98.1
Условная WGAN	79.2	83.7	81.3	81.5
Синтетика по иконке	46.5	53.7	67.8	69.7

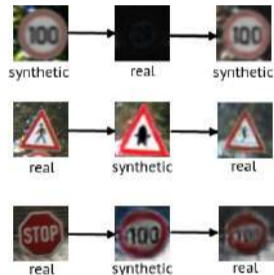
Экспериментальная оценка проводилась на базе немецких знаков GTSRB: 43 класса знаков, ~ 50 тыс. изображений для обучения и тестирования

В качестве классификатора использовалась сверточная нейросеть

4. Метод повышения визуального реализма синтетических изображений дорожных знаков

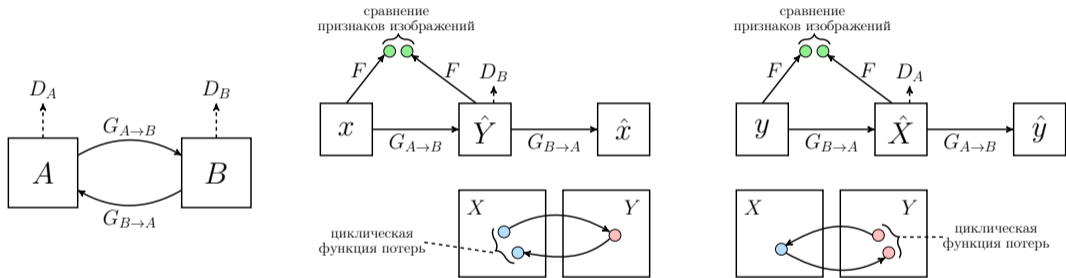


примеры простых синтетических изображений дорожных знаков



J.-Y. Zhu, T. Park et al. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. ICCV 2017

Схема предложенного алгоритма с сохранением класса объекта



Для сохранения класса объекта в функцию потерь добавляется сравнение нейросетевых признаков \bullet исходного и преобразованного изображения

Экспериментальная оценка

Тренировочная выборка	Точность классификации
GTSRB	98.2
GTSRB с заменой на прост. синт. «ограничение скорости»	87.7
GTSRB с заменой на GAN-синт. «ограничение скорости»	95.8
GTSRB с заменой на прост. синт. «синие круги»	93.8
GTSRB с заменой на GAN-синт. «синие круги»	94.4
GTSRB с заменой на прост. синт. «красные треугольники»	94.6
GTSRB с заменой на GAN-синт. «красные треугольники»	95.8

5. Классификация редких дорожных знаков

Базовый метод



Примеры изображений из выборок



Icon

Real

Synt

CGI

GAN

Экспериментальная оценка базового классификатора

Трен. выборка	Качество нейросети			Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	среднее	редкие	частые	среднее	редкие	частые	среднее	редкие	частые	среднее	редкие	частые
RTSD	88.9	0.0	94.9	71.5	41.7	73.5	73.0	61.1	73.8			
RTSD+Synt	91.6	43.2	94.9	75.8	58.4	76.9	76.1	64.5	76.8	81.6	57.6	83.3
RTSD+CGI	92.5	53.9	95.7	53.4	38.7	54.4	75.2	65.1	75.9	82.0	58.5	83.4
RTSD+GAN	92.8	53.8	95.4	77.0	61.3	78.1	74.8	65.8	75.4	82.1	59.6	83.7

Трен. выборка	Качество нейросети			Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	среднее	редкие	частые	среднее	редкие	частые	среднее	редкие	частые	среднее	редкие	частые
Synt	26.1	18.6	26.6	17.5	8.9	18.1	26.4	30.5	26.1	24.4	13.3	25.1
CGI	49.3	44.2	49.6	15.4	17.4	15.2	38.7	42.9	38.4	30.8	19.2	31.6
GAN	49.8	43.9	50.2	22.9	13.3	22.9	44.1	46.1	44.0	39.3	34.1	39.7

Предложенный метод классификации



Б. Фаизов, В. Шахуро, В. Санжаров, А. Конушин. Классификация редких дорожных знаков. Компьютерная оптика 2020; 44(2)

Архитектура предложенного классификатора

Wideresnet

```
graph TD; A[Wideresnet] --> B[Полносвязный слой, 850 нейронов  
Модифицированная сравнительная  
функция потерь  
Для извлечения признаков]; A --> C[Полносвязный слой, 205 нейронов  
Функция потерь: кросс-энтропия  
Для классификации нейросетью]; B --> D[Случайный лес, 1000 деревьев  
Для дискриминации на редкие  
и частые классы];
```

Полносвязный слой, 850 нейронов
Модифицированная сравнительная
функция потерь
Для извлечения признаков

Случайный лес, 1000 деревьев
Для дискриминации на редкие
и частые классы

Полносвязный слой, 205 нейронов
Функция потерь: кросс-энтропия
Для классификации нейросетью

Модифицированная сравнительная функция потерь

$$L(x_1, x_2, y, z) = z(1 - y)D^2(x_1, x_2) + zy(\max(0, m - D(x_1, x_2)))^2$$

$$D(x_1, x_2) = \|f(x_1) - f(x_2)\|_2$$

x_1, x_2 — изображения,

$D(x_1, x_2)$ — расстояние между вектор-признаками изображений,

$m > 0$ — гиперпараметр, регулирующий, насколько должны отличаться вектор-признаки разных классов,

y — бинарная переменная, равная 0, если x_1 и x_2 одного класса, и 1 в противном случае,

z — бинарная переменная, равная 0, если оба изображения редкого класса, и 1 в противном случае.

Экспериментальная оценка предложенного метода

Трен. выборка	Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	среднее	редкие	частые	среднее	редкие	частые	среднее	редкие	частые
RTSD+Synt	93.2	72.5	94.6	91.5	78.5	92.3	94.4	73.6	95.8
RTSD+CGI	74.3	71.9	74.4	91.8	78.3	92.7	88.6	60.5	90.5
RTSD+GAN	90.3	69.4	91.7	87.8	79.9	88.3	93.8	82.3	94.5

Точность бинарной классификации на частые и редкие знаки

Трен. выборка	Качество нейросети			Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	среднее	редкие	частые	среднее	редкие	частые	среднее	редкие	частые	среднее	редкие	частые
RTSD	88.9	0.0	94.9	71.5	41.7	73.5	73.0	61.1	73.8			
RTSD+Synt				92.4	55.7	94.9	92.0	68.5	93.5	93.3	57.3	95.8
RTSD+CGI				89.6	58.0	91.8	87.6	70.4	88.7	88.5	53.0	90.9
RTSD+GAN				93.1	70.7	94.6	92.1	71.7	93.1	93.5	70.2	95.1

Точность многоклассовой классификации предложенного метода

6. Метод синтеза обучающих выборок для детектора

Базовые методы: случайное размещение знаков



Synt

CGI

GAN

Улучшенный метод: вставка синтетических знаков вместо реальных



исходный кадр

inpainting

вставка знака

V. Shakhuro, B. Faizov, A. Konushin. Rare traffic sign recognition using synthetic training data. ICVIP 2019

Результаты экспериментов (детектор)

Тренир. выборка	все	част.	редк.
Real	89.08	89.20	86.02
Synt	13.90	13.85	14.83
CGI	10.70	10.63	13.23
InpaintSynt	55.23	56.26	55.26
InpaintGAN	62.12	62.36	54.26
Real + Synt	88.48	88.62	85.54
Real + CGI	88.56	88.72	85.72
Real + GAN	88.53	88.69	85.21
Real + InpaintSynt	88.61	88.71	86.63
Real + InpaintGAN	88.98	89.13	86.60

Использование синтетических данных повышает качество выделения редких знаков при незначительном уменьшении качества выделения частых знаков. Метрика качества — площадь под кривой точность-полнота.

Результаты экспериментов (детектор и классиф.)

Детектор и классификатор WideResNet учатся на одной выборке

Тренир. выборка	все	част.	редк.
Real	75.44	82.46	09.09
Synt	09.97	09.99	09.09
CGI	08.81	08.81	08.53
InpaintSynt	15.89	15.96	13.79
InpaintGAN	19.85	18.80	19.91
Real + Synt	80.22	84.38	33.96
Real + CGI	83.84	85.15	48.51
Real + GAN	84.45	85.68	49.83
Real + InpaintSynt	76.41	82.93	34.00
Real + InpaintGAN	84.99	85.86	51.36

Использование синтетических данных значительно повышает качество распознавания как частых, так и редких знаков.

Результаты экспериментов (детектор и классиф.)

Детектор и классификатор редких знаков с использованием индекса из синт. данных, улучшенных с помощью GAN

Тренир. выборка детектора	все	част.	редк.
Real	86.06	86.73	58.96
Real+GAN	85.14	85.89	58.57
Real+InpaintSynt	84.40	85.28	59.42
Real+InpaintGAN	86.11	86.62	60.13

Использование синтетических данных и предложенного метода классификации (задача 5) значительно повышает качество распознавания знаков.

Результаты

1. Собрана и размечена выборка российских дорожных знаков. Выборка подходит для обучения и оценки алгоритмов распознавания редких дорожных знаков.
2. Исследована применимость порождающих нейросетей к задаче генерации дорожных знаков.
3. Предложен метод условной генерации дорожных знаков на основе метрики Васерштейна.
4. Предложен нейросетевой метод улучшения синтетических дорожных знаков. Метод позволяет повысить точность классификации знаков.
5. Предложен метод классификации дорожных знаков, позволяющий классифицировать как частые, так и редкие классы дорожных знаков. Метод обучается на реальных и синтетических обучающих примерах.
6. Предложен метод синтеза обучающих выборок для детектора дорожных знаков. Метод позволяет повысить точность детектирования редких дорожных знаков.

Публикации по теме диссертационной работы

1. V. Shakhuro, A. Konushin. Russian traffic sign images dataset. Computer Optics 2016; 40(2)
2. В. Шахуро, А. Конушин. Синтез обучающих выборок для классификации знаков с помощью нейросетей. Компьютерная оптика 2018; 42(1)
3. П. Хрушков, В. Шахуро, А. Конушин. Синтез изображений дорожных знаков с помощью условных порождающих противоборствующих нейросетей. ГрафиКон 2018
4. V. Shakhuro, B. Faizov, A. Konushin. Rare traffic sign recognition using synthetic training data. ICVIP 2019
5. Б. Фаизов, В. Шахуро, В. Санжаров, А. Конушин. Классификация редких дорожных знаков. Компьютерная оптика 2020; 44(2)

Научная новизна

1. В диссертационной работе впервые исследована задача распознавания редких дорожных знаков на достаточно большой выборке с большим количеством классов.
2. В работе впервые подробно исследована применимость конкурирующих порождающих нейросетей к задаче генерации синтетических изображений дорожных знаков. Исследованные методы подробно протестированы на задаче классификации дорожных знаков.
3. В работе предложен новый метод решения задачи классификации дорожных знаков, позволяющий улучшить качество классификации редких дорожных знаков при неухудшении качества классификации частых дорожных знаков.
4. В работе предложен новый метод синтеза данных для обучения детектора дорожных знаков, повышающий точность детектирования редких дорожных знаков.