

## Классификация редких дорожных знаков

Б.В. Фаизов<sup>1</sup>, В.И. Шахуро<sup>1</sup>, В.В. Санжаров<sup>1,3</sup>, А.С. Конушин<sup>1,2</sup>

<sup>1</sup> МГУ имени М.В. Ломоносова, Москва, Россия,

<sup>2</sup> НИУ Высшая школа экономики, Москва, Россия,

<sup>3</sup> РГУ нефти и газа (НИУ) имени И.М. Губкина, Москва, Россия

### Аннотация

В работе исследуется возможность применения нейронных сетей для классификации объектов, которых мало или совсем нет в обучающей выборке, на примере задачи классификации редких дорожных знаков. Рассматриваются нейронные сети, обученные с использованием сравнительной функции потерь и её модификациями, а также методы генерации синтетических выборок для задач классификации. В качестве базового метода используется индексирование классов объектов при помощи нейросетевых признаков. Проводится сравнение классификаторов, обученных при помощи трёх видов синтетических выборок, а также их смесей с реальными данными. Предлагается метод классификации редких дорожных знаков, использующий нейросетевой дискриминатор редких и частых знаков. Проведённая экспериментальная оценка показала, что предложенный метод позволяет классифицировать редкие дорожные знаки без существенной потери качества на частых знаках.

**Ключевые слова:** классификация дорожных знаков, синтетические обучающие выборки, нейронные сети, распознавание изображений, трансформации изображений, композиции нейросетей.

**Цитирование:** Фаизов, Б.В. Классификация редких дорожных знаков / Б.В. Фаизов, В.И. Шахуро, В.В. Санжаров, А.С. Конушин // Компьютерная оптика. – 2020. – Т. 44, № 2. – С. 236-243. – DOI: 10.18287/2412-6179-CO-601.

**Citation:** Faizov BV, Shakhuro VI, Sanzharov VV, Konushin AS. Classification of rare traffic signs. Computer Optics 2020; 44(2): 236-243. DOI: 10.18287/2412-6179-CO-601.

### Введение

Для решения задач компьютерного зрения в последнее время всё чаще используются нейросети. Основным требованием нейросетевых подходов является наличие достаточного числа обучающих примеров. Подготовка таких наборов данных сама по себе является большой проблемой, поскольку требует сначала кропотливой работы по сбору изображений, а затем больших временных затрат на разметку данных. Размечать данные дорого, при этом каких-то данных при разметке мы можем вообще не найти. В итоге возникают проблемы при классификации объектов, отсутствующих в обучающих выборках. В то же время нужно, чтобы модель научилась правильно определять класс объекта и не давала уверенных предсказаний на заведомо неверных примерах. Для решения проблемы отсутствующих в выборке классов на помощь приходят синтетические данные. Использование плохих искусственных данных может даже привести к падению качества на классах, которые были в изначальной выборке.

Рассмотрим задачу распознавания дорожных знаков. Она является нерешённой для большого количества классов знаков. Также имеется много редких классов знаков. Для систем, распознающих дорожные знаки, важна способность к быстрой адаптации в случае появления нового вида дорожных знаков. При появлении нового класса знаков трудно собрать достаточное количество реальных обучающих примеров.

При наличии синтетически сгенерированных изображений этого знака можно быстро внедрить в систему возможность классификации нового вида знака. Простой нейросетевой классификатор редкие классы на тестовой выборке будет причислять к неверным классам. Качественные синтетические выборки позволяют решить проблему с отсутствующими классами и со сборкой данных с примерами этих знаков.

В данной работе рассматривается задача классификации редких дорожных знаков на основе метода поиска аномалий [1] и метода ближайших соседей, использующего нейросетевые признаки [2]. Предложенный подход по-разному выполняет классификацию редких и частых знаков в выборке. Для этого используется дискриминатор, различающий эти два типа объектов. Метод позволяет классифицировать редкие дорожные знаки, не ухудшая при этом качество классификации частых дорожных знаков. По сравнению с ним качество распознавания на редких знаках базового классификатора оставляет желать лучшего. Обучение нейронной сети для получения нейросетевых признаков сравнительной функцией потерь *contrastive loss* повышает качество на редких знаках, однако этот результат получается в ущерб качеству на частых знаках.

### 1. Обзор существующих работ

#### 1.1. Методы классификации изображений

Существующие актуальные методы классификации изображений можно условно разделить на

нейросетевые и комбинированные. Первая группа методов, нейросетевая, позволяет получать очень высокое качество классификации изображений. Быстрое развитие таких методов началось после победы в соревновании ILSVRC2012 свёрточной нейросети AlexNet в задаче классификации изображений [3]. С тех пор появилось большое количество разных архитектур. Основное направление их развития было связано с увеличением количества слоёв в архитектуре. При этом решалась проблема возрастающей увеличительной сложности новых архитектур. Для повышения эффективности создаются новые виды слоёв. Были представлены архитектуры ResNet [4] и Inception [5]. Эти архитектуры стали основой для создания сетей с сотнями и даже тысячами слоёв. В данных архитектурах происходило увеличение глубины нейросетей. В работе [2] предложена нейросетевая архитектура WideresNet, в которой было сделано увеличение ширины нейросети для классификации. Благодаря этому стало возможно параллельно производить вычисления для различных слоёв на одном уровне. Повысилась эффективность и качество нейросети. Однако общий недостаток всех таких методов – это большая требовательность к размеру коллекции и количеству примеров каждого класса.

Ко второй группе относятся комбинированные методы классификации, использующие нейросетевые признаки изображения совместно с другими методами классификации. При этом совсем не обязательно, чтобы нейросеть обучалась на том же наборе данных, на которых она будет работать. Было замечено [6], что нейросети, обученные на коллекции ImageNet, для классификации изображений извлекают достаточно общие признаки, применимые в самых разных задачах. Также данный подход позволяет решать задачу классификации при небольшом количестве примеров каждого класса. К примеру, из изображений тренировочной выборки извлекаются признаки. Далее уже на них обучается достаточно общий метод классификации изображений – например, метод ближайших соседей. Такой подход позволяет избавиться от проблемы фиксированного количества классов, поскольку не обязательно переучивать нейросеть, что требует значительных вычислительных затрат. Достаточно заново обучить метод ближайших соседей. Метод также позволяет решить задачи классификации при небольшом размере коллекции. Данный подход для классификации редких знаков был использован нами в качестве базового.

Существуют подходы, строящие «центроиды» – наилучшие представители каждого класса [7]. В них так же, как и в предыдущем способе, извлекаются признаки из примеров обучающей выборки. А далее градиентным методом получают вектора признаков – наилучших представителей каждого из классов. Далее классификация происходит с помощью определения того, к centroidу какого класса ближе всего находится пример. Ос-

новной проблемой данного подхода является то, что centroidы могут быть расположены очень близко в пространстве признаков. Классификация между такими классами будет очень некачественной.

Авторы ещё одного метода предлагают строить линейное преобразование признаков изображений, извлекаемых нейросетью [8]. Такое преобразование пытается разделить в пространстве признаков как можно дальше друг от друга объекты разных классов и сделать как можно ближе друг к другу объекты одного класса. Для обучения преобразования используется триплетная функция потерь.

Следующий подход основан на идее, что можно обучать нейросеть специально для извлечения признаков (так как обычно обучается сеть для классификации изображений, а далее отбрасывается её последний слой). Для этого была введена сравнительная функция потерь [9]. Эта функция зависит от признаков, извлекаемых нейросетью. В процессе обучения сразу два изображения прогоняются через нейросеть. Если они одного класса, то расстояние между их признаками минимизируется, а если разных классов, то, наоборот, становится больше.

### *1.2. Методы определения редких классов*

В предложенном методе используется дискриминатор редких и частых классов, присутствующих в выборке. Нужно сразу обозначить, что есть две разные мотивации решения этой задачи. Первая – определение выбросов в тестовой выборке, то есть тех изображений, про которые классификатор не может и не должен ничего сказать. Вторая – поиск тех изображений, с которыми классификатор должен уметь работать, но не обучился из-за недостаточного количества данных. Именно вторая причина нас интересует в данной статье.

Первым подходом для решения задачи дискриминации изображений на редкие и частые классы было установление порога уверенности классификатора. Если он меньше заданного уровня, то полагается, что изображение из неизвестного класса. Однако данный метод показывает плохие результаты, поскольку нейросети обычно оказываются слишком уверенными в тех изображениях, которые они не должны уметь классифицировать.

Для решения этой проблемы было предложено использовать внешние данные в процессе обучения. Также была предложена функция ошибки, учитывающая внешние данные [10]. Она обучала нейросеть выдавать неуверенные классификации на классах, которые она не должна знать, и «прижимала» выходы нейросети к равномерному распределению. В итоге авторы получили значительный прирост в качестве метода.

В другом подходе [1] с использованием немного трансформированной сравнительной функции потерь обучается нейросеть для извлечения признаков. Так-

же в процессе обучения используются дополнительные внешние данные из другой коллекции. В итоге признаки, извлекаемые нейросетью, становятся хорошо линейно разделимы для разных классов. Для изображений из неизвестных классов признаки также неявно кластеризуются. В итоге полученная нейросеть пригодна как для задачи определения, известен ли это класс, так и для введения новых классов без обучения нейросети заново. Метод показал свою эффективность на различных коллекциях, в том числе и на дорожных знаках. Мы будем использовать такие признаки для классификации дорожных знаков.

## 2. Описание метода классификации редких дорожных знаков

### 2.1. Генерация синтетических данных

В данной работе были сгенерированы три вида синтетических данных. Во всех случаях для генерации использовались иконки дорожных знаков, которые являются «идеальными» примерами дорожных знаков.

Первый вид данных – простая синтетика (synt) – получался путём применения к пиктограмме знака преобразований со случайными параметрами: поворот, сдвиг, изменение контраста, гауссовское размытие, размытие движения, наложение на фон.

Второй вид данных – фотореалистичные дорожные знаки (sgi) – получался путём рендеринга трёхмерных моделей дорожных знаков на столбах поверх реальных изображений. Рендеринг проводился с помощью движка Hydra [11].

Третий вид данных – улучшенные фотореалистичные дорожные знаки (sgi-gan). Современные методы позволяют обучать нейросетевые генераторы изображений дорожных знаков из случайного шума [12]. В данной работе мы обучаем нейросеть для преобразования синтетических изображений из коллекции sgi в более качественные. Нейросеть-преобразователь синтетических дорожных знаков в реальные обучалась с помощью дополнительной нейросети-преобразователя и циклической функции потерь, описанной в [13]. Метод обучения использует две нейросети для двух доменов изображений А (синтетические фотореалистичные дорожные знаки) и В (реальные изображения дорожных знаков). Первая нейросеть преобразует изображения из домена А в домен В (из синтетических в реальные изображения), вторая – из В в А (из реальных в синтетические изображения). Для обучения требуется два набора изображений из доменов А и В. Во время обучения требуется циклическая согласованность: последовательное применение двух нейросетей из домена А в В и обратно в А должно давать результат, эквивалентный исходному изображению. Для того, чтобы нейросеть при преобразовании сохраняла класс дорожного знака, мы используем дополнительное слагаемое в функции потерь из [14], штрафующее за разницу вектор-признаков исходного и преобразованного изображения. Вектор-признаки – это значения активаций нейронов предпоследнего слоя обученного классификатора дорожных знаков. Схема обучения нейросетей-преобразователей показана на рис. 1.

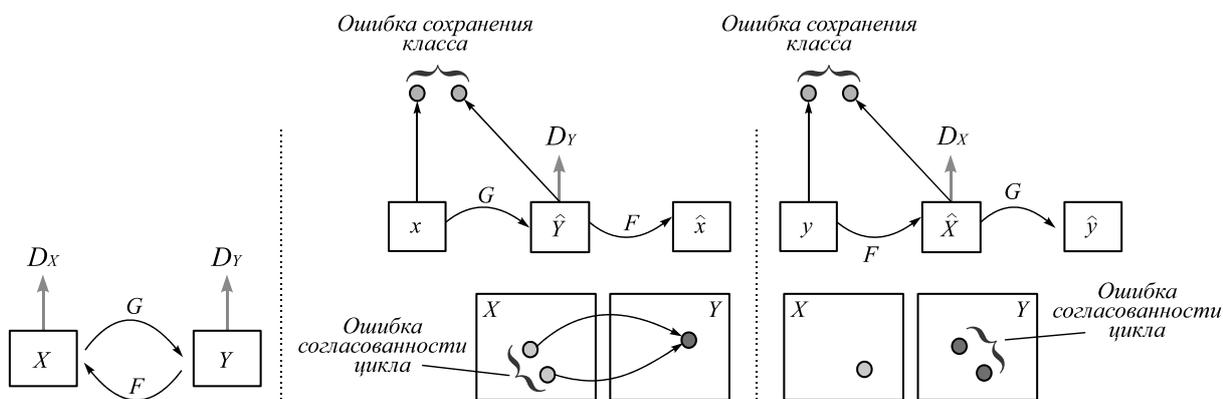


Рис. 1. Схема генерации изображений, улучшенных cyclegan, с сохранением класса дорожного знака

### 2.2. Базовый метод классификации

В качестве базового метода классификации дорожных знаков, отсутствующих в исходной коллекции RTSD, был использован метод, при котором сначала обучается нейросетевой классификатор WideresNet. Далее у него отбрасывается последний слой для классификации. После полученная нейросеть используется для извлечения признаков из изображений. Для классификации дорожных знаков на признаках из изображений обучается метод ближайших соседей. Признаки, извлечённые из изображений

нейросетью, перед подачей на вход методу ближайших соседей, нормируются по каждой координате. Это повышает точность работы метода ближайших соседей. Среднее и стандартное отклонение каждого признака рассчитываются на обучающей выборке, а далее они же используются на тестовой. Общая схема работы этого метода показана на рис. 2.

### 2.3. Улучшение нейросетевых признаков

Как будет показано в экспериментальной части, полученное базовое качество на редких знаках в лучшем случае будет сильно проигрывать качеству на ча-

стных знаков. Для повышения качества классификации мы рассмотрели методы улучшения пространственного распределения нейросетевых признаков. Базовый метод, использующий softmax-классификатор, не даёт гарантий на то, что признаки объектов разных классов будут находиться достаточно далеко друг от друга в признаковом пространстве. Поэтому мы добавили в функцию потерь дополнительное слагаемое, способствующее пространственному распределению признаков. В качестве функции потерь в признаковом пространстве мы использовали сравнительную функцию потерь. Она имеет вид (1, 2):

$$L(x_1, x_2, y) = \frac{1}{2}(1 - y)D(x_1, x_2) + \frac{1}{2}y \max(0, m - D(x_1, x_2)), \quad (1)$$

$$D(x_1, x_2) = \|f(x_1) - f(x_2)\|_2. \quad (2)$$

Тут  $x_1, x_2$  – это два изображения;  $f(x)$  – вектор-признак, который выдает нейросеть на входном изображении  $x$ ;  $D(x_1, x_2)$  – это функция, возвращающая евклидово расстояние между векторами признаков;  $m$  – это числовой гиперпараметр – порог, показывающий, насколько сильно должны отличаться векторы-признаки;  $y$  – бинарная переменная, которая равна 0, если  $x_1$  и  $x_2$  одного класса, и равна 1, если они разного класса.

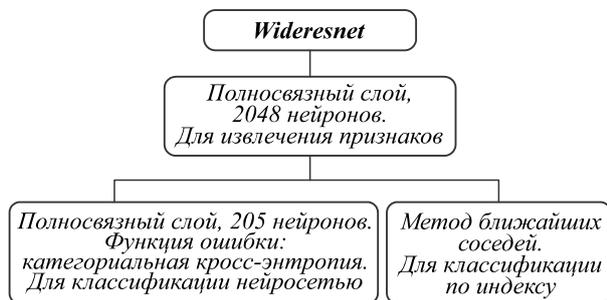


Рис. 2. Схема базового классификатора

Далее аналогично базовому методу извлеченные признаки масштабируются и на них обучается метод ближайших соседей. Общая схема работы этого метода показана на рис. 3.



Рис. 3. Схема сети для получения улучшенных признаков

#### 2.4. Предложенный метод классификации

В результате экспериментов будет показано, что для редких знаков лучше всего производить классификацию методом ближайших соседей, обученным на нейросетевых признаках, а для частых знаков – по выходу нейросетевого слоя для классификации изображения. Поэтому сначала нужно определить, к какому классу относится изображение – частому или редкому. Для этого оно подаётся на вход описанному нами далее дискриминатору редких и частых знаков. А далее в зависимости от типа знака выполняется классификация соответствующим образом наилучшей нейросетью.

#### 2.5. Метод разделения редких и частых знаков

Идея нашего метода дискриминации редких и частых знаков основана на подходе, описанном в [1]. В нейросеть WideresNet к выходу классификатора мы добавили еще один выход, который выдает обученные признаки. Функцией ошибки нового выхода стала модифицированная сравнительная функция потерь, которая имеет вид (3):

$$L(x_1, x_2, y) = \frac{1}{2}z(1 - y)D(x_1, x_2) + \frac{1}{2}zy \max(0, m - D(x_1, x_2)). \quad (3)$$

В этой функции потерь в сравнении с (1) добавлена переменная  $z$ , которая равна 0, если оба изображения из редких классов, и 1, если класс хотя бы одного изображения частый.

В процессе обучения на вход нейросети подаётся батч из 64 изображений. Далее извлекаются векторы признаков каждого изображения. Функция ошибки на новом выходе нейросети считается как среднее её значение для каждой пары изображений в батче. На старом выходе нейросети для классификации в качестве функции ошибки используется категориальная кросс-энтропия.

Добавление в функцию ошибки такой переменной  $z$  позволяет избавиться от влияния пар из редких классов на функцию ошибки.

Далее слой для классификации удаляется из нейросети (он использовался только для обучения, чтобы в процессе обучения нейросеть имела больше данных о решаемой задаче). На извлечённых признаках бинарной классификации обучается решающий лес из 1000 деревьев. Он должен выдавать 1, если входное изображение редкое, и 0, если оно частое. Общая схема работы этого метода показана на рис. 4.

### 3. Экспериментальная оценка

#### 3.1. База изображений и данные для обучения

В качестве реальных данных была взята база изображений RTSD [15], состоящая из 205 классов, из которых 99 встречаются только в тестовой выборке и полностью отсутствуют в тренировочной.

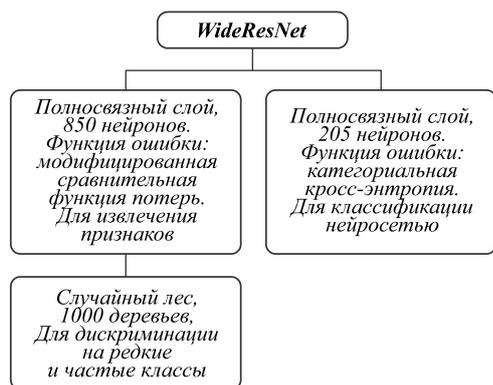


Рис. 4. Схема сети для классификации двух видов знаков

В качестве синтетических данных были взяты простые иконки дорожных знаков, synt – простая синтетика, cgi – данные, полученные трассировкой лучей, и cgi-gan – данные, полученные улучшением данных cgi.

Для обучения и тестирования нейросети использовались следующие наборы данных:

- rtsd – только реальные данные;
- rtsd+ <название синтетической коллекции> – реальные данные и соответствующие синтетические данные; для индекса из синтетических данных используется соответствующая синтетика;
- icons – иконки дорожных знаков.

Иконка	Реальная картинка	synt	cgi	cgi-gan	Иконка	Реальная картинка	synt	cgi	cgi-gan

Рис. 5. Примеры изображений из различных выборок, используемых для обучения и оценки качества классификатора

Для обучения метода ближайших соседей на извлечённых признаках использовались разные данные: иконки для каждого из классов дорожных знаков, одно изображение из тестовой выборки (далее оно удалялось из неё на этапе тестирования), 10 синтетических картинок (из тех, на которых обучалась нейросеть) для каждого класса. При использовании синтетических картинок параметр  $k$ -метода ближайших соседей равен 5, а в остальных случаях – 1.

Примеры изображений из разных выборок приведены на рис. 5.

### 3.2. Результаты разных подходов

При обучении на реальных данных без использования синтетических получаем качество классификации нейросети 94,9% на частых классах и 0% на редких. При добавлении синтетических данных в обучающую выборку качество нейросети повышается до 95,7% на частых знаках (при использовании данных cgi) и 53,9% на редких знаках. Таким образом, само по себе использование синтетических данных позволяет повысить качество классификации редких знаков.

Базовый метод, при котором используются нейросетевые признаки и метод ближайших соседей, позволяет без использования реальных данных для индекса улучшить качество на редких знаках в лучшем случае до 59,6% в ущерб качеству на частых знаках – 83,7%. Индекс из реальных данных позволяет улучшить качество на редких знаках уже до 65,8%, однако на частых знаках качество упало ещё больше. Полные результаты – в табл. 1.

Также было проверено, можно ли совсем отказаться от реальных данных и использовать только синтетические. Результаты эксперимента показаны в табл. 2. Как видно, от реальных данных отказаться нельзя.

Кроме того проверялось, насколько наличие редких классов при обучении нейросети ухудшает качество работы на частых классах и каким бы было качество, если бы нейросеть обучалась только на частых данных. Результаты эксперимента показаны в табл. 3. Как видно, качество на частых знаках немного подросло (в пределах 1%), но на редких знаках качество заметно упало.

Улучшенные нейросетевые признаки (при обучении их сравнительной функцией потерь) позволяют без использования реальных данных для индекса улучшить качество на редких знаках до 73,7%; на частых знаках качество также стало чуть лучше – 86,8%. Индекс из реальных данных в эксперименте уже не дал прироста в качестве. Результаты эксперимента показаны в табл. 4.

Предложенный нами подход позволил добиться улучшения качества на редких знаках без ущерба качеству на частых. Сначала обрабатывает дискриминатор изображения на редкие и частые классы. Результаты его работы показаны в табл. 7. Как видно, без использования реальных данных для индекса в лучшем случае получается добиться 70,2% точности определения редкого класса и 95,1% точности определения частого класса.

Далее для классификации редких знаков используются нейросетевые признаки, улучшенные сравни-

тельной функцией потерь, и метод ближайших соседей. А для классификации частых знаков используется ответ нейросети (выход её softmax-слоя).

Мы рассмотрели два подхода. В первом из них для дискриминации и классификации знаков используются две очень похожие, но разные нейросети. Ре-

зультаты – в табл. 5. Во втором мы попробовали для дискриминации редких и частых знаков использовать нейросетевые признаки классификатора. В таком случае качество немного упало. Но заметно выросла эффективность работы – не нужно прогонять изображение через две нейросети. Результаты – в табл. 6.

Табл. 1. Точность работы базового метода

	Качество нейросети			Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые	средняя	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 + cgi-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

Табл. 2. Точность работы при обучении нейронной сети только на синтетических данных

	Качество нейросети			Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые
rtsd	88,9	0,0	94,9	71,5	41,7	73,5	73,0	61,1	73,8			
rtsd + 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
rtsd + 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
rtsd + cgi-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

Табл. 3. Точность работы при обучении нейронной сети только на частых классах

	Качество нейросети			Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые
rtsd	88,9	0,0	94,9	71,5	41,7	73,5	73,0	61,1	73,8			
rtsd + synt	89,8	0,0	95,9	76,3	54,1	77,8	75,9	59,5	76,9	80,9	55,9	82,6
rtsd + cgi	89,9	0,0	96,0	75,2	45,7	77,2	75,6	61,9	76,5	81,5	56,2	83,5
rtsd + cgi-gan	90,3	0,0	96,5	74,1	51,8	75,6	76,9	57,8	78,2	82,6	58,4	84,2

Табл. 4. Точность работы при обучении с использованием сравнительной функции потерь и улучшенными признаками

	Качество нейросети			Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые
rtsd	88,9	00,0	94,9	71,5	41,7	73,5	73,0	61,1	73,8			
rtsd + synt	93,7	52,0	96,5	83,9	63,1	85,4	82,3	68,2	83,2	87,8	61,3	89,6
rtsd + cgi	92,5	57,1	94,9	67,7	46,9	69,1	77,1	66,6	77,8	56,4	40,0	57,5
rtsd + cgi-gan	94,4	65,1	96,4	84,0	71,7	84,8	80,5	71,9	81,0	86,0	73,7	86,8

Табл. 5. Точность работы предложенного метода, реализованного двумя нейросетями

	Качество нейросети			Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые
rtsd	88,9	00,0	94,9	71,5	41,7	73,5	73,0	61,1	73,8			
rtsd + synt				91,9	59,2	94,1	92,5	70,2	93,6	92,8	57,5	95,2
rtsd + cgi				84,6	48,8	87,0	90,5	69,7	91,8	87,8	47,6	90,5
rtsd + cgi-gan				90,4	67,8	91,9	92,2	71,0	93,2	92,4	70,5	93,9

Табл. 6. Точность работы предложенного метода, реализованного одной нейросетью

	Качество нейросети			Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые	средняя	редкие	частые
rtsd	88,9	00,0	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 + cgi-gan				93,1	70,7	94,6	92,1	71,7	93,1	93,5	70,2	95,1

Табл. 7. Точность работы дискриминатора редких и частых классов

	Качество на индексе из иконок			Качество на индексе из тестовой выборки			Качество на индексе из синтетики		
	средняя	редкие	частые	средняя	редкие	частые	средняя	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 + cgi-gan	90,3	69,4	91,7	87,8	79,9	88,3	93,8	82,3	94,5

### 3.3. Выводы

Из результатов можно сделать следующие выводы:

1. Добавление синтетических данных к реальным позволяет получить прирост в качестве работы нейросетевого классификатора.
2. Улучшенные при помощи cuselegan синтетические данные позволяют приблизить данные к реальным. Такие данные показали наилучшую точность работы классификатора.
3. Обучить качественный классификатор только на синтетических данных не получится.
4. Использование сравнительной функции потерь помогает улучшить нейросетевые признаки и сделать классы более разделимыми в признаковом пространстве.
5. При помощи правильно обученных нейросетевых признаков и при наличии только синтетических примеров редких классов можно обучить дискриминатор редких и частых классов с хорошим качеством.
6. Предложенная нами комбинированная архитектура позволяет улучшить качество классификатора на редких знаках без ущерба качеству на частых.

### Заключение

В данной статье были исследованы методы повышения качества работы нейросетевых классификаторов на данных, которые либо не представлены в коллекции, либо представлены в очень маленьком количестве. В качестве примера рассмотрена задача классификации дорожных знаков. Было произведено сравнение различных синтетических коллекций дорожных знаков. Экспериментально установлено, что трансформированная сравнительная функция потерь позволяет улучшить нейросетевые признаки и сделать их более репрезентативными. Был предложен метод классификации дорожных знаков, который улучшает качество на редких знаках и совсем незна-

чительно ухудшает качество на частых знаках. Для реализации метода был разработан нейросетевой дискриминатор редких и частых классов знаков.

### Благодарности

Работа В.В. Санжарова по фотореалистичному синтезу дорожных знаков выполнена при поддержке гранта РФФИ 18-31-20032 мол\_а\_вед «Физически-корректное моделирование освещения и синтез изображений на массивно-параллельных вычислительных системах в приложениях искусственного интеллекта», работа Б.В. Фаизова, В.И. Шахуро и А.С. Конушина по улучшению дорожных знаков с помощью нейросетей и классификации редких знаков поддержана грантом РНФ 17-71-20072 «Нейробайесовские методы в задачах машинного обучения, масштабируемой оптимизации и компьютерного зрения».

### Литература

1. **Masana, M.** Metric learning for novelty and anomaly detection [Electronical Resource] / M. Masana, [et al.] // arXiv preprint. – URL: <https://arxiv.org/abs/1808.05492> (request date 3.10.2019).
2. **Zagoruyko, S.** Wide residual networks [Electronical Resource] / S. Zagoruyko, N. Komodakis // arXiv preprint. – URL: <https://arxiv.org/abs/1605.07146> (request date 3.10.2019).
3. **Krizhevsky, A.** ImageNet classification with deep convolutional neural networks / A. Krizhevsky, I. Sutskever, G.E. Hinton // Advances in Neural Information Processing Systems. – 2012. – Vol. 25, Issue 2. – P. 1097-1105.
4. **He, K.** Deep residual learning for image recognition / K. He, [et al.] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2016. – P. 770-778.
5. **Szegedy, C.** Going deeper with convolutions / C. Szegedy, [et al.] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. – 2015. – P. 1-9.
6. **Sharif Razavian, A.** CNN features off-the-shelf: an astounding baseline for recognition / A. Sharif Razavian, [et al.] // Proceedings of the IEEE Computer Society Confer-

- ence on Computer Vision and Pattern Recognition Workshops. – 2014. – P. 806-813.
7. **Huang, G.** Centroid networks for few-shot clustering and unsupervised few-shot classification [Electronical Resource] / G. Huang, H. Larochelle, S. Lacoste-Julien // arXiv preprint. – URL: <https://arxiv.org/abs/1902.08605> (request date 3.10.2019).
  8. **Sankaranarayanan, S.** Triplet probabilistic embedding for face verification and clustering / S. Sankaranarayanan, [et al.] // 2016 IEEE 8<sup>th</sup> International Conference on Biometrics: Theory, Applications and Systems (BTAS). – 2016. – P. 1-8.
  9. **Hadsell, R.** Dimensionality reduction by learning an invariant mapping / R. Hadsell, S. Chopra, Y. LeCun // 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06). – 2006. – Vol. 2. – P. 1735-1742.
  10. **Hendrycks, D.** Deep anomaly detection with outlier exposure [Electronical Resource] / D. Hendrycks, M. Mazeika, T.G. Dietterich // arXiv preprint. – URL: <https://arxiv.org/abs/1812.04606> (request date 3.10.2019).
  11. HydraAPI [Electronical Resource]. – URL: <https://github.com/Ray-Tracing-Systems/HydraAPI> (request date 3.10.2019).
  12. **Шахуро, В.И.** Синтез обучающих выборок для классификации дорожных знаков с помощью нейросетей / В.И. Шахуро, А.С. Конушин // Компьютерная оптика. – 2018. – Т. 42, № 1. – С. 105-112. – DOI: 10.18287/2412-6179-2018-42-1-105-112.
  13. **Zhu, J.Y.** Unpaired image-to-image translation using cycle-consistent adversarial networks / J.Y. Zhu, [et al.] // Proceedings of the IEEE International Conference on Computer Vision. – 2017. – P. 2223-2232.
  14. **Huang, R.** Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis / R. Huang, [et al.] // Proceedings of the IEEE International Conference on Computer Vision. – 2017. – P. 2439-2448.
  15. **Шахуро, В.И.** Российская база изображений автодорожных знаков / В.И. Шахуро, А.С. Конушин // Компьютерная оптика. – 2016. – Т. 40, № 2. – С. 294-300. – DOI: 10.18287/2412-6179-2016-40-2-294-300

### Сведения об авторах

**Фаизов Борис Владимирович**, 1998 года рождения, студент МГУ имени М.В. Ломоносова. Научные интересы: компьютерное зрение, машинное обучение, программирование. E-mail: [boris.faizov@graphics.cs.msu.ru](mailto:boris.faizov@graphics.cs.msu.ru).

**Шахуро Владислав Игоревич**, 1993 года рождения, в 2015 году окончил МГУ имени М.В. Ломоносова. Окончил аспирантуру НИУ ВШЭ по компьютерным наукам. Научные интересы: обработка изображений, компьютерное зрение, машинное обучение, программирование. E-mail: [vlad.shakhuro@graphics.cs.msu.ru](mailto:vlad.shakhuro@graphics.cs.msu.ru).

**Санжаров Вадим Владимирович**, 1987 года рождения, в 2010 году окончил РГУ нефти и газа (НИУ) имени И.М. Губкина. Работает ведущим программистом лаборатории компьютерной графики и мультимедиа на ВМК МГУ имени М.В. Ломоносова и старшим преподавателем в РГУ нефти и газа (НИУ) имени И.М. Губкина. Научные интересы: фотореалистичная компьютерная графика, синтез изображений на GPU. E-mail: [yadim.sanzharov@graphics.cs.msu.ru](mailto:yadim.sanzharov@graphics.cs.msu.ru).

**Конушин Антон Сергеевич**, 1980 года рождения, в 2002 году окончил МГУ имени М.В. Ломоносова. В 2005 году защитил кандидатскую диссертацию в ИПМ имени М.В. Келдыша РАН. Работает доцентом на ВМК МГУ имени М.В. Ломоносова. Научные интересы: компьютерное зрение, машинное обучение. E-mail: [anton.konushin@graphics.cs.msu.ru](mailto:anton.konushin@graphics.cs.msu.ru).

ГРНТИ: 28.23.15

Поступила в редакцию 22 июля 2019 г. Окончательный вариант – 11 октября 2019 г.

---

# Classification of rare traffic signs

B.V. Faizov<sup>1</sup>, V.I. Shakhuro<sup>1</sup>, V.V. Sanzharov<sup>1,3</sup>, A.S. Konushin<sup>1,2</sup>

<sup>1</sup>Lomonosov Moscow State University, Moscow, Russia,

<sup>2</sup>NRU Higher School of Economics, Moscow, Russia,

<sup>3</sup>Gubkin RSU of Oil and Gas

## Abstract

The paper studies the possibility of using neural networks for the classification of objects that are few or absent at all in the training set. The task is illustrated by the example of classification of rare traffic signs. We consider neural networks trained using a contrastive loss function and its modifications, also we use different methods for generating synthetic samples for classification problems. As a basic method, the indexing of classes using neural network features is used. A comparison is made of classifiers trained with three different types of synthetic samples and their mixtures with real data. We propose a method of classification of rare traffic signs using a neural network discriminator of rare and frequent signs. The experimental evaluation shows that the proposed method allows rare traffic signs to be classified without significant loss of frequent sign classification quality.

**Keywords:** traffic sign classification, synthetic training samples, neural networks, image recognition, image transforms, neural network compositions.

**Citation:** Faizov BV, Shakhuro VI, Sanzharov VV, Konushin AS. Classification of rare traffic signs. *Computer Optics* 2020; 44(2): 237-244. DOI: 10.18287/2412-6179-CO-601.

**Acknowledgements:** This work was supported by the Russian Science Foundation under RSF grant 18-31-20032 ("Physically correct lighting modeling and image synthesis on massively parallel computing systems in applications of artificial intelligence") and the Russian Science Foundation under RSF grant 17-71-20072 ("Deep Bayesian Methods in Machine Learning, Scalable Optimization and Computer Vision Problems").

## References

- [1] Masana M, et al. Metric learning for novelty and anomaly detection. Source: (<https://arxiv.org/abs/1808.05492>).
- [2] Zagoruyko S, Komodakis N. Wide residual networks. Source: (<https://arxiv.org/abs/1605.07146>).
- [3] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 2012; 25(2): 1097-1105.
- [4] He K, et al. Deep residual learning for image recognition. *IEEE Conf Comput Soc Comput Vis Pattern Recogn* 2016: 770-778.
- [5] Szegedy C, et al. Going deeper with convolutions. *IEEE Conf Comput Soc Comput Vis Pattern Recogn* 2015: 1-9.
- [6] Sharif Razavian A, et al. CNN features off-the-shelf: an astounding baseline for recognition. *IEEE Conf Comput Soc Comput Vis Pattern Recogn Workshops* 2014: 806-813.
- [7] Huang G, Larochelle H, Lacoste-Julien S. Centroid networks for few-shot clustering and unsupervised few-shot classification. Source: (<https://arxiv.org/abs/1902.08605>).
- [8] Sankaranarayanan S, et al. Triplet probabilistic embedding for face verification and clustering. *2016 IEEE 8<sup>th</sup> Int Conf Biometrics: Theory Appl Syst (BTAS)* 2016: 1-8.
- [9] Hadsell R., Chopra S., LeCun Y. Dimensionality reduction by learning an invariant mapping. *2006 IEEE Conf Comput Soc Comput Vis Pattern Recogn (CVPR'06)* 2006; 2: 1735-1742.
- [10] Hendrycks D, Mazeika M, Dietterich TG. Deep anomaly detection with outlier exposure. Source: (<https://arxiv.org/abs/1812.04606>).
- [11] <https://github.com/Ray-Tracing-Systems/HydraAPI>
- [12] Shakhuro VI, Konushin AS. Image synthesis with neural networks for traffic sign classification. *Computer Optics* 2018; 42(1): 105-112. DOI: 10.18287/2412-6179-2018-42-1-105-112.
- [13] Zhu JY, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proc IEEE Int Conf Comput Vis* 2017: 2223-2232.
- [14] Huang R, et al. Beyond face rotation: Global and local perception gan for photorealistic and identity preserving frontal view synthesis. *Proc IEEE Int Conf Comput Vis* 2017: 2439-2448.
- [15] Shakhuro VI, Konushin AS. Russian traffic sign images dataset. *Computer Optics* 2016; 40(2): 294-300. DOI: 10.18287/2412-6179-2016-40-2-294-300.

---

## Authors' information

**Boris Vladimirovich Faizov**, (b. 1998), student at Lomonosov Moscow State University. Research interests are computer vision, machine learning and programming. E-mail: [boris.faizov@graphics.cs.msu.ru](mailto:boris.faizov@graphics.cs.msu.ru).

**Vladislav Igorevich Shakhuro**, (b. 1993), graduated from Lomonosov Moscow State University in 2015. Research interests are image processing, computer vision, machine learning, and programming. E-mail: [vlad.shakhuro@graphics.cs.msu.ru](mailto:vlad.shakhuro@graphics.cs.msu.ru).

---

---

**Vadim Vladimirovich Sanzharov**, (b. 1987), graduated from Gubkin Russian State University of Oil and Gas in 2010. He is currently a lead software engineer at Graphics and Media Lab at Lomonosov Moscow State University and a senior lecturer at Gubkin Russian State University of Oil and Gas. Research interests include photo-realistic computer graphics and rendering on GPUs. E-mail: [vadim.sanzharov@graphics.cs.msu.ru](mailto:vadim.sanzharov@graphics.cs.msu.ru) .

**Anton Sergeevich Konushin**, (b. 1980), graduated from Lomonosov Moscow State University in 2002. In 2005 he successfully defended his PhD thesis in M.V. Keldysh Institute for Applied Mathematics RAS. He is currently associate professor at Lomonosov Moscow State University. Research interests are computer vision and machine learning. E-mail: [anton.konushin@graphics.cs.msu.ru](mailto:anton.konushin@graphics.cs.msu.ru) .

---

*Received July 22, 2019. The final version – October 11, 2019.*

---