

ISSN 1816-0301 (Print)
ISSN 2617-6963 (Online)
УДК 004.931; 004.932

Поступила в редакцию 25.06.2018
Received 25.06.2018

Л. А. Ходасевич

*Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь*

ГЕНЕРАЦИЯ РЕАЛИСТИЧНЫХ ИЗОБРАЖЕНИЙ ДЛЯ ОБУЧЕНИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ В ЗАДАЧЕ НАВИГАЦИИ РОБОТА

Аннотация. На конкретном практическом примере рассмотрена и решена проблема формирования обучающей выборки для настройки нейросетевого детектора, предназначенного для распознавания дверей на цифровых изображениях помещений. Разработан метод генерации реалистичных синтетических данных, заключающийся в замене на цифровых изображениях априори известных объектов-мишеней новыми объектами, которые были получены путем проективного преобразования эталонных объектов. Метод предназначен для формирования обучающей выборки, необходимой для обучения и тестирования искусственных нейронных сетей, которые впоследствии применяются в системе управления мобильными роботами для решения задач автономной навигации. Эффективность предложенного метода была подтверждена экспериментально.

Ключевые слова: генерация реалистичных изображений, нейронная сеть, детектор, обучение, набор данных
Для цитирования. Ходасевич, Л. А. Генерация реалистичных изображений для обучения искусственных нейронных сетей в задаче навигации робота / Л. А. Ходасевич // Информатика. – 2018. – Т. 15, № 4. – С. 50–58.

L. A. Khodasevich

*The United Institute of Informatics Problems of the National Academy
of Sciences of Belarus, Minsk, Belarus*

REALISTIC IMAGES GENERATION FOR TRAINING ARTIFICIAL NEURAL NETWORKS IN ROBOT NAVIGATION PROBLEM

Abstract. The problem of obtaining training dataset for setting weights of neural network designed for indoor detection of doors is considered and solved on the particular example. A method for generating realistic synthetic data is developed. The method involves replacing a priori known target objects on digital images with new reference objects that were obtained by projective transformation of reference objects. The method is designed to obtain training dataset for training and testing of artificial neural networks, which will be used in the mobile robot control system to solve autonomous navigation problem. The effectiveness of the proposed method was confirmed experimentally.

Keywords: realistic images generation, neural network, detector, training, dataset

For citation. Khodasevich L. A. Realistic images generation for training artificial neural networks in robot navigation problem. *Informatics*, 2018, vol. 15, no. 4, pp. 50–58 (in Russian).

Введение. Важнейшим этапом решения задачи распознавания образов методами машинного обучения является подготовка обучающей выборки. Несмотря на наличие эффективного метода распознавания, при отсутствии хороших обучающих данных удовлетворительные результаты достигнуты не будут. В связи с этим при обучении моделей машинного обучения подготовка обучающей выборки является наиболее трудоемким процессом, особенно если они реализованы на основе искусственных нейронных сетей (ИНС). Современные ИНС, имеющие многослойную архитектуру, требуют не только больших вычислительных мощностей, но и колоссальных объемов обучающей выборки (десятки тысяч цифровых изображений для каждого из распознаваемых классов) [1]. Для оптимизации процесса формирования обучающей выборки создаются открытые базы данных миллионов цифровых изображений, которые содержат искомые графические объекты. Однако при решении конкретных практических задач

с использованием ИНС распространены ситуации, когда в базах данных отсутствуют изображения с распознаваемыми объектами. Эти случаи требуют использования других методов формирования обучающей выборки.

Существуют следующие подходы к формированию обучающей выборки:

- 1) на основе реальных изображений;
- 2) на основе синтетических изображений;
- 3) из синтетических данных, сгенерированных на основе реальных изображений.

При использовании первого подхода собираются реальные цифровые изображения интересующих объектов (в данном случае помещений с дверьми), которые проверяются на соответствие определенным требованиям, таким, например, как наличие дверей на изображениях, корректность выделения ограничивающих прямоугольников, если изображения были предварительно размечены, и т. п. Как правило, далее следует процесс разметки собранных изображений: для задачи классификации – приписывание метки класса каждому изображению, для задачи детектирования – выделение координат прямоугольников, ограничивающих объекты, и приписывание им метки класса. Очевидным недостатком этого подхода является трата большого количества времени на сбор и разметку обучающих данных. К его преимуществам следует отнести то, что обучающие данные будут максимально подобны данным, которые будут поступать на вход ИНС в процессе функционирования, и по ним в процессе тестирования можно будет судить об эффективности обученной ИНС на реальных сценах.

При формировании выборки вторым методом обучающие данные генерируются искусственно. Например, большинство объектов можно представить в виде геометрических примитивов с добавлением линий различной толщины для имитации фактуры и наложить на подходящее фоновое изображение. Затем к таким изображениям требуется применить, к примеру, преобразование яркости и контрастности, аффинное и проективное преобразования, добавление шума и т. д. Очевидным недостатком такого подхода является отдаленность синтетических изображений от реальных, вследствие чего ИНС хотя и будет показывать высокие результаты тестирования на синтетических данных, но при работе с реальными данными результаты распознавания будут низкими. Достоинством такого подхода является то, что по сравнению с предыдущим методом время, необходимое на формирование обучающей выборки, значительно уменьшается.

Третий подход представляет собой комбинацию первых двух. Для генерации обучающих данных за основу берутся реальные изображения. На них показаны распознаваемые объекты, класс которых априори известен (далее – объекты-мишени). Затем эти объекты замещают эталонными объектами, к которым применяются пространственные преобразования для получения наибольшего соответствия контуров эталонного объекта и объекта-мишени, причем типы применяемых преобразований определяются путем анализа существующего набора реальных изображений. Такой подход совмещает достоинства двух предыдущих.

Необходимость использования того или иного подхода зависит от решаемой задачи. Если реальных изображений много и легко получить для них разметку, то для формирования обучающей выборки используется первый подход; если реальные изображения отсутствуют, – второй подход. В случае, когда реальные изображения присутствуют в небольшом количестве и (или) не покрывают все возможные типы и конфигурации объекта, то оптимальным будет использование третьего подхода для формирования обучающей выборки. Третий подход хорошо себя зарекомендовал при решении задач распознавания различных типов объектов на изображениях. Так, в работах [2, 3] приводятся результаты сравнения обучения ИНС на выборке, сформированной из реальных изображений, и на выборке, сформированной из реальных и синтетических изображений. Второй набор обучающих данных позволил увеличить качество распознавания объектов.

Постановка задачи. Цель работы – сформировать обучающую выборку для настройки параметров ИНС при решении задачи детектирования дверей на изображениях внутри помещений. Эта ИНС будет применяться в системе технического зрения автономного мобильного робота, описанного в работе [4]. Указанный мобильный робот должен строить маршрут в заранее неизвестном помещении, опираясь на расположение и порядок следования дверей.

Детектор – это ИНС, реализующая алгоритм генерации ограничивающих прямоугольников, в которых потенциально располагаются искомые объекты (например, алгоритм выборочного поиска [5] или ИНС RPN [6]), и классификатор, который применяется для классификации областей, ограниченных сгенерированными прямоугольниками. Для проведения экспериментов был выбран классический детектор Faster R-CNN [6] на основе сверточной нейронной сети Inception V2 [7].

В сети Интернет представлено большое количество баз данных, содержащих наборы изображений для решения задач классификации, детектирования, сегментации, анализа сцены и т. д. (например, ImageNet (URL: <http://www.image-net.org>), MNIST [8], COCO dataset [9]). Для задачи детектирования объектов требуются размеченные изображения, т. е. изображения, на которых искомые объекты выделены ограничивающими прямоугольниками. Координаты ограничивающих прямоугольников хранятся в отдельном файле, называемом аннотацией. Аннотации могут записываться в различных форматах. В работе используется широко известный формат PASCAL VOC, который применяется в том числе и в базе данных ImageNet.

Необходимость разметки изображений значительно усложняет поиск и формирование обучающей выборки. Из-за специфики детектируемых объектов, необходимых для решения задачи навигации мобильного робота, размеченные изображения были найдены только в базе данных ImageNet. Обучающая выборка содержала 1000 изображений дверей, сфотографированных под прямым углом (рис. 1). Из них разметку имели только 700 изображений. Остальные были размечены вручную с помощью открытого программного инструмента LabelImg [10], который использует формат PASCAL VOC для чтения и записи аннотаций к изображениям.

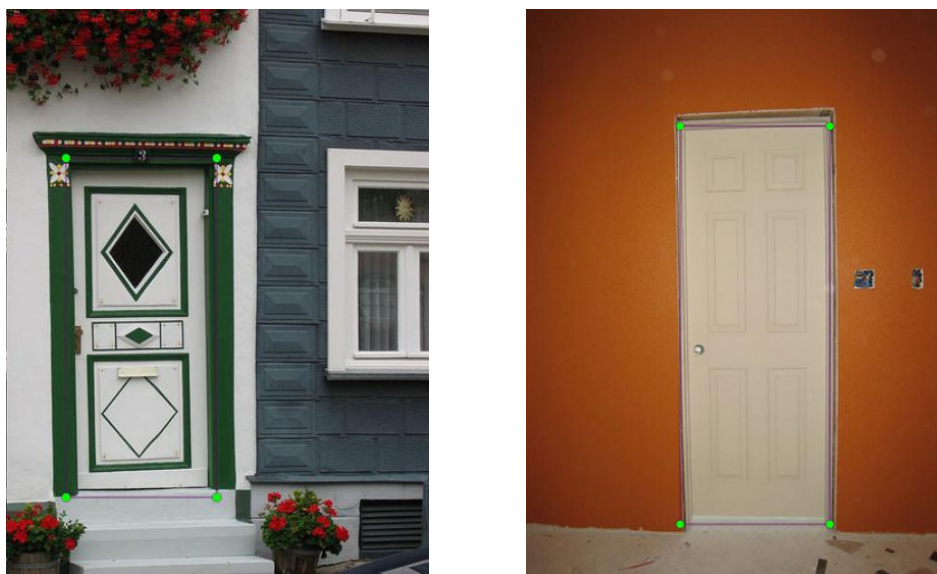


Рис. 1. Размеченные обучающие изображения из базы данных ImageNet

Проблема обучения детектора на данных, представленных на рис. 1, состоит в том, что в процессе выполнения роботом заложенной программы он двигается вдоль коридоров и стен, в результате двери чаще располагаются не прямо, а слева или справа от направления взгляда, что искажает их первоначальную геометрическую форму. Стала очевидной необходимость использования третьего подхода к формированию обучающей выборки – генерации синтетических данных на основе реальных изображений.

Генерация синтетических обучающих данных на основе реальных изображений. Для генерации изображений дверей, расположенных слева и справа от направления взгляда, к каждому исходному изображению дважды было применено проективное преобразование, при котором прямые линии остаются прямыми, однако параллельные линии не обязательно являются таковыми. При использовании проективного преобразования изображение сжимается, сверху и снизу образуются «пустоты», которые автоматически заполняются черным цветом.

Из-за этого детектор может обучиться распознавать лишь двери, расположенные на таком черном фоне. Для устранения данной проблемы черный цвет сверху и снизу на преобразованных изображениях был закрасен усредненным цветом, вычисленным по наборам цветов верхней и нижней частей изображения соответственно (рис. 2).



Рис. 2. Результат применения проективного преобразования для дверей, расположенных слева, и для дверей, расположенных справа

Результаты проведенных экспериментов показали, что детектор, обученный на данных рис. 2, способен распознавать лишь двери на изображениях с такими же пропорциями, как у изображений из обучающей выборки, т. е. двери, занимающие значительную площадь изображения, у которого ширина приблизительно в два-три раза меньше длины. Для устранения этого недостатка преобразованные изображения были наложены на размытые по Гауссу с размером ядра (31, 31) и стандартным отклонением 10 исходные изображения, причем изображения накладывались со смещением от центра влево или вправо в зависимости от типа преобразования (рис. 3).

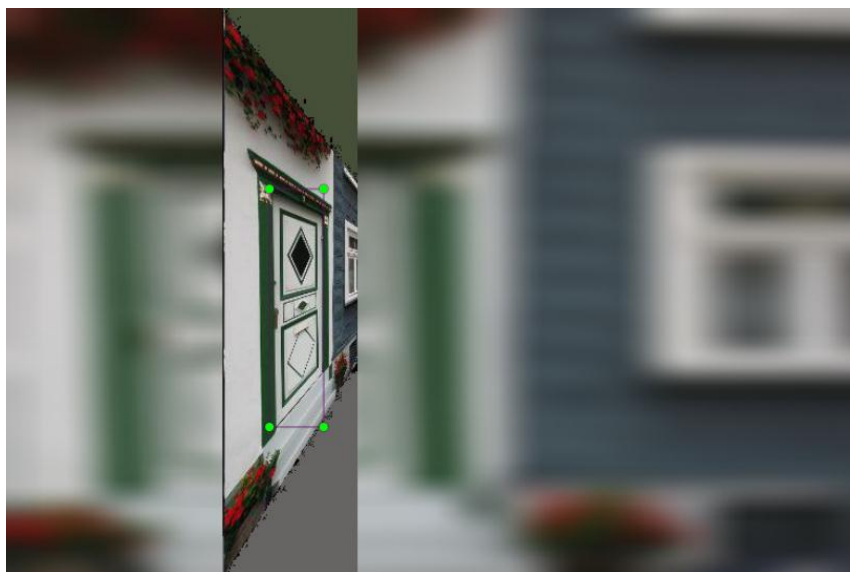


Рис. 3. Преобразованное обучающее изображение

Тем не менее обучение детектора на новых данных не дало удовлетворительных результатов, поэтому возникла необходимость подготовить более правдоподобные синтетические данные.

Генерация реалистичных синтетических данных. Для генерации реалистичных синтетических данных за основу был взят набор видеопоследовательностей, на которых запечатлено движение вдоль коридора. Для реализации метода генерации размеченных обучающих данных из видеопоследовательностей следует выполнить следующие действия:

- 1) разбить видеопоследовательность на набор последовательных кадров;
- 2) для первого кадра вручную выделить прямоугольники, ограничивающие двери, которые будут служить шаблонами для распознавания дверей;
- 3) для оставшихся кадров применить метод поиска объектов по шаблону для генерации ограничивающих прямоугольников;
- 4) при необходимости вручную подкорректировать границы ограничивающих прямоугольников.

Было использовано 11 видеопоследовательностей, причем на одной видеопоследовательности при смене кадров типы дверей не меняются. Это может привести к тому, что детектор обучится распознавать только те типы дверей, которые представлены на используемом наборе видеопоследовательностей. Для решения данной проблемы была предложена идея замены дверей, которые присутствуют на изображениях коридоров и для которых найдены ограничивающие прямоугольники (далее – двери-мишени), на другие двери различных типов и дизайна (далее – эталонные двери). Для того чтобы контуры эталонных дверей совпадали с контурами дверей-мишеней, к ним требуется применить проективное преобразование. Для вычисления матрицы проективного преобразования необходимо знать координаты четырех точек (x_{01}, y_{01}) , (x_{02}, y_{01}) , (x_{01}, y_{02}) , (x_{02}, y_{02}) исходного изображения и координаты четырех точек (x_1, y_1) , (x_2, y_2) , (x_1, y_3) , (x_2, y_4) целевого изображения. Координаты точек (x_{01}, y_{01}) , (x_{02}, y_{01}) , (x_{01}, y_{02}) , (x_{02}, y_{02}) известны из значений ширины w и высоты h изображений эталонных дверей: $x_{01}=0$, $y_{01}=0$, $x_{02}=w$, $y_{02}=h$. Также из координат ограничивающих прямоугольников x_{\min} , y_{\min} , x_{\max} , y_{\max} известны координаты точек (x_1, y_1) , (x_1, y_3) целевого изображения для дверей, расположенных слева: $x_1=x_{\min}$, $y_1=y_{\min}$, $y_3=y_{\max}$, и координаты двух точек (x_2, y_2) , (x_2, y_4) целевого изображения для дверей, расположенных справа: $x_2=x_{\max}$, $y_2=y_{\min}$, $y_4=y_{\max}$. Координаты оставшихся двух точек целевого изображения необходимо найти. На рис. 4 красным цветом отмечены известные координаты, зеленым – неизвестные.

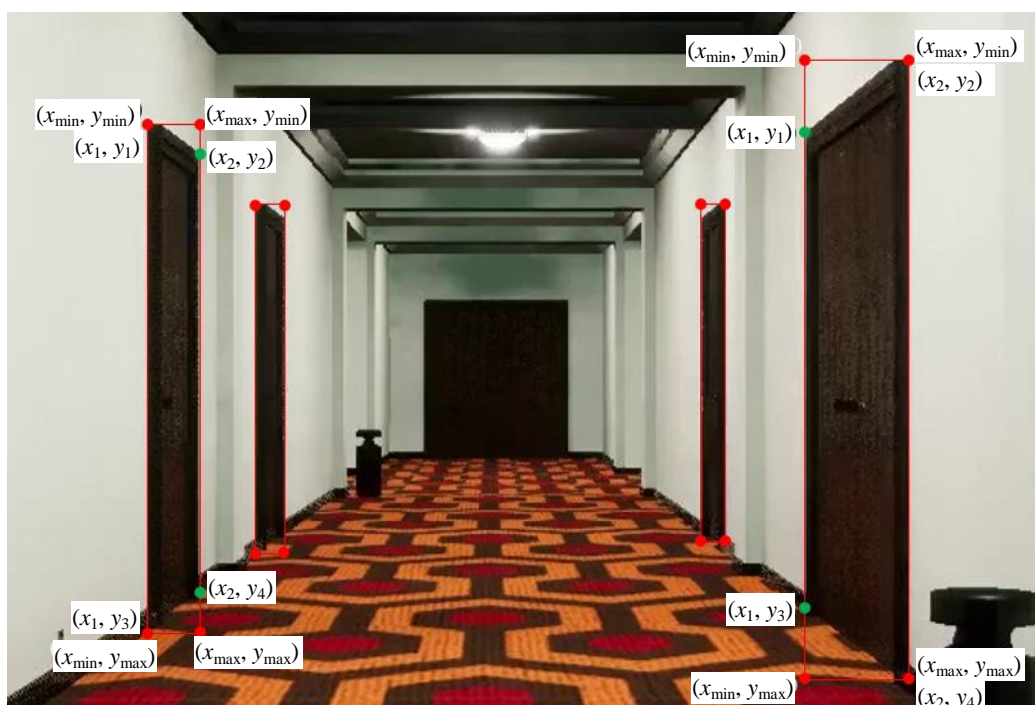


Рис. 4. Проблема нахождения целевых точек для вычисления матрицы проективного преобразования

Предлагаются три способа нахождения координат двух неизвестных точек:

1. Применение проективного преобразования с постоянными углами с изменением при этом ширины и высоты изображений эталонных дверей по ширине и высоте ограничивающего дверь-мишень прямоугольника. Использование данного метода оправдано, так как видеочамера на борту автономного мобильного робота, для навигации которого применяется обучаемый детектор, расположена на постоянной высоте и движется по центру коридоров, вследствие чего двери расположены под одним углом к полю зрения автономного мобильного робота. Недостаток метода заключается в том, что контуры эталонных дверей не будут в точности совпадать с контурами дверей-мишеней.

2. Использование метода выделения и анализа контуров для поиска горизонтальных линий контуров дверей-мишеней. Данный метод следует применять не ко всему изображению, а к части изображения внутри ограничивающего прямоугольника в том случае, когда цвет дверей-мишеней контрастирует с цветом стен. Действительно, на большинстве видеопоследовательностей стены светлого цвета, а двери-мишени – темного. Однако напольное покрытие в коридорах чаще всего имеет темный цвет, вследствие чего двери сливаются с напольным покрытием и становится невозможно выделить контуры дверей-мишеней снизу. Решением этой проблемы является использование в качестве нижнего контура зеркального отображения верхнего. Недостаток данного метода такой же, как и у предыдущего: контуры эталонной двери не будут совпадать с контурами двери-мишени.

3. Использование свойств линейной перспективы, при которой горизонтальные линии контуров дверей-мишеней одинаковой высоты при перспективном искажении находятся на одной прямой линии. Для вычисления параметров этой линии достаточно наличия известных координат точек (x_{\min}, y_{\min}) двух ограничивающих прямоугольников, расположенных слева, и координат точек (x_{\max}, y_{\min}) двух ограничивающих прямоугольников, расположенных справа, для верхних горизонтальных линий и соответственно координат точек (x_{\min}, y_{\max}) двух ограничивающих прямоугольников, расположенных слева, и (x_{\max}, y_{\max}) двух ограничивающих прямоугольников, расположенных справа, для нижних горизонтальных линий. Тогда точки (x_2, y_2) , (x_2, y_4) , (x_1, y_1) , (x_1, y_3) пересечения таких четырех нижних и верхних горизонтальных линий дверей-мишеней, расположенных справа и слева, с вертикальными линиями дверей-мишеней (известными из координат ограничивающих прямоугольников) будут недостающими точками для вычисления матриц проективного преобразования.

Опишем алгоритм вычисления координат точек пересечения двух прямых, заданных общим уравнением.

Пусть известны координаты двух точек (x_{10}, y_{10}) и (x_{20}, y_{20}) , которые лежат на прямой, заданной уравнением

$$Ax + By + C = 0. \quad (1)$$

Тогда параметры A, B, C прямой находятся следующим образом:

$$A = y_{10} - y_{20}, \quad B = x_{20} - x_{10}, \quad C = x_{10}y_{20} - x_{20}y_{10}. \quad (2)$$

Пусть заданы две прямые по формуле (1) с параметрами A_1, B_1, C_1 и A_2, B_2, C_2 . Тогда координаты точки пересечения прямых $M_0 = (x_0, y_0)$ вычисляются по формулам

$$x_0 = \frac{B_2C_1 - B_1C_2}{A_2B_1 - A_1B_2}, \quad y_0 = \frac{A_1C_2 - A_2C_1}{A_2B_1 - A_1B_2}. \quad (3)$$

При $A_2B_1 - A_1B_2 = 0$ прямые либо совпадают, либо параллельны. Если $A_1C_2 - A_2C_1 = 0$ и $B_2C_1 - B_1C_2 = 0$, то прямые совпадают; если $A_1C_2 - A_2C_1 \neq 0$ и $B_2C_1 - B_1C_2 \neq 0$, то параллельны.

Третий метод не имеет недостатков двух предыдущих, однако на него накладываются некоторые ограничения: для вычисления параметров линии необходимо наличие двух ограничивающих прямоугольников на одной стороне изображения.

Если при записи видеопоследовательности камера перемещалась ровно и не тряслась, то ограничения можно снять следующим образом:

1) в множестве изображений, полученных из видеопоследовательности, найти изображение, на котором присутствуют две двери на одной стороне. По ограничивающим двери прямоугольникам найти параметры линий, а затем использовать эти параметры при вычислении неизвестных координат двух точек для остальных изображений;

2) для каждого изображения, на котором недостаточно ограничивающих прямоугольников для вычисления параметров линий, взять координаты ограничивающего прямоугольника предыдущего или следующего изображения. Если таких прямоугольников несколько, то следует брать тот из них, у которого расстояние от соответствующих координат ограничивающего прямоугольника текущего изображения наибольшее.

Для вычисления координат неизвестных точек использовался третий метод, так как он позволяет получить наиболее реалистичные изображения.

При наложении на изображения, полученные из видеопоследовательностей, были использованы 28 различных изображений эталонных дверей. Для вставки в каждый ограничивающий прямоугольник изображение эталонной двери выбиралось случайным образом по равномерному распределению.

Перед применением проективного преобразования изображения эталонных дверей конвертировались в формат RGBA. «Пустоты», образующиеся при использовании преобразования, заполнялись цветом (0, 0, 0, 0), т. е. прозрачным цветом. Вставка осуществлялась следующим образом: каждый пиксел изображения эталонной двери проверялся на прозрачность. Если прозрачность пиксела равнялась 255, то соответствующий пиксел исходного изображения заменялся пикселом изображения эталонной двери; если 0, то никакие изменения не производились.

В результате были сгенерированы 2180 реалистичных изображений, на которых изображены 5713 дверей (рис. 5).

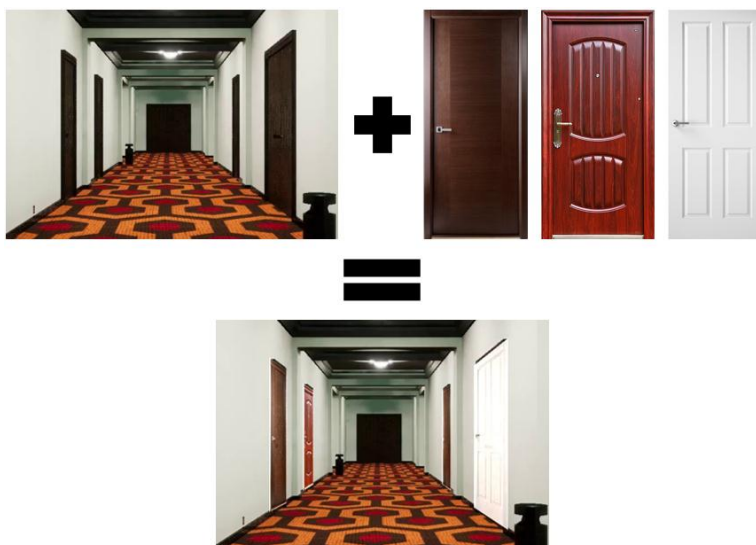


Рис. 5. Получение реалистичных синтетических данных путем замены четырех объектов-мишеней эталонными объектами

Проведение экспериментов. Для проведения численных экспериментов использовались изображения базы данных ImageNet, преобразованные изображения базы данных ImageNet, сгенерированные реалистичные изображения и фотографии коридоров. Данные были разделены на обучающую и тестовую выборки в соотношении четыре к одному.

Обучение детектора Faster R-CNN Inception V2 проводилось на компьютере с двумя графическими процессорами NVIDIA GeForce 1080 и занимало несколько часов. Параллельно с процессом обучения сети запускался процесс тестирования. Когда точность детектирования объектов на тестовой выборке достигала своего наибольшего значения, а затем начинала падать, обучение останавливалось.

Детектор обучался распознавать четыре класса: `straightDoor` – двери, расположенные прямо; `leftDoor` – двери, расположенные слева; `rightDoor` – двери, расположенные справа; `openDoor` – открытые двери.

По итогам тестирования лучшие результаты показало обучение детектора на сгенерированных правдоподобных синтетических данных (таблица).

Численные характеристики обученного детектора Faster R-CNN

Класс	Средняя точность
<code>leftDoor</code>	0,958
<code>rightDoor</code>	0,967
<code>openDoor</code>	0,775
<code>straightDoor</code>	0,885
Общая	0,896

Средняя точность детектирования класса открытых дверей значительно ниже средней точности детектирования остальных классов, так как обучающих примеров открытых дверей значительно меньше, чем остальных (около 200 против более 1000).

Заключение. В работе рассмотрена проблема формирования обучающей выборки для обучения ИНС на примере задачи детектирования дверей на изображениях реальных помещений. В ходе выполнения работы для формирования обучающей выборки был использован подход, при котором синтетические данные генерируются путем модификации реальных изображений. Предложен метод генерации обучающих данных для настройки параметров детектора Faster R-CNN Inception V2. В результате обученный детектор с достаточно высокой точностью (88,5–96,7 %) распознает двери, расположенные прямо, слева и справа, и открытые двери. Проведенные эксперименты позволяют сделать вывод о том, что использование синтетических данных для обучения нейронных сетей оправдано и позволяет получить результаты, сопоставимые с обучением на реальных данных.

Список использованных источников

1. What is the best multi-stage architecture for object recognition? / Y. LeCun [et al.] // The 12th Intern. Conf. on Computer Vision, Kyoto, 27 Sept. – 4 Oct. 2009. – Kyoto, 2009. – P. 2146–2153.
2. Training deep face recognition systems with synthetic data [Electronic resource] / A. Kortylewski [et al.] // Cornell University Library. – 2018. – Mode of access: <https://arxiv.org/pdf/1802.05891.pdf>. – Date of access: 16.04.2018.
3. Чигорин, А. Классификация автодорожных знаков на основе сверточной нейросети, обученной на синтетических данных / А. Чигорин, Б. Моисеев // Труды 22 Междунар. конф. по компьютерной графике и машинному зрению ГрафиКон'2012, Москва, 1–5 окт. 2012 г. – М., 2012. – С. 284–287.
4. Прокопович, Г. А. Разработка системы технического зрения для сервисного мобильного робота / Г. А. Прокопович // Третий Всерос. науч.-практ. семинар «Беспилотные транспортные средства с элементами искусственного интеллекта», Иннополис, 22–23 сент. 2015 г. – Иннополис : Ун-т Иннополис, 2016. – С. 127–136.
5. Selective search for object recognition [Electronic resource] / J. R. R. Uijlings [et al.]. – Mode of access: <http://www.huppelen.nl/publications/selectiveSearchDraft.pdf>. – Date of access: 28.04.2018.
6. Faster R-CNN: towards real-time object detection with region proposal networks [Electronic resource] / S. Ren [et al.] // Cornell University Library. – 2015. – Mode of access: <https://arxiv.org/pdf/1506.01497.pdf>. – Date of access: 15.05.2018.
7. Rethinking the inception architecture for computer vision [Electronic resource] / C. Szegedy [et al.] // Cornell University Library. – 2015. – Mode of access: <https://arxiv.org/pdf/1512.00567.pdf>. – Date of access: 15.05.2018.
8. LeCun, Y. The MNIST database [Electronic resource] / Y. LeCun, C. Cortes, C. J. C. Burges. – Mode of access: <http://yann.lecun.com/exdb/mnist/>. – Date of access: 16.03.2018.
9. Microsoft COCO: Common objects in context [Electronic resource] / T. Lin [et al.]. – 2018. – Mode of access: <http://cocodataset.org/#home>. – Date of access: 16.03.2018.
10. Tzutalin / labelImg [Electronic resource]. – 2015. – Mode of access: <https://github.com/tzutalin/labelImg>. – Date of access: 18.06.2018.

References

1. LeCun Y., Jarrett K., Kavukcuoglu K., Ranzato M. What is the best multi-stage architecture for object recognition? *The 12th International Conference on Computer Vision, Kyoto, 27 September – 4 October 2009*. Kyoto, 2009, pp. 2146–2153.
2. Kortylewski A., Schneider A., Gerig T., Egger B., Morel-Foster A., Vetter T. *Training deep face recognition systems with synthetic data*. Available at: <https://arxiv.org/pdf/1802.05891.pdf> (accessed 16.04.2018).

3. Chigorin A., Moiseev B. Klassifikatsiya avtorozhnykh znakov na osnove svertochnoi neiroseti, obuchennoi na sinteticheskikh dannyykh [Road signs classification based on convolutional neural network that trained on synthetic data]. Trudy 22 Mezhdunarodnoj konferencii po komp'yuternoj grafike i mashinnomu zreniju GrafiKon'2012 [*Proceedings of the 22nd International Conference on Computer Graphics and Computer Vision GraphiCon 2012, Moscow, 1–5 October 2012*]. Moscow, 2012, pp. 284–287 (in Russian).
4. Prokopovich G. A. Razrabotka sistemy tehnikeskogo zreniya dlja servisnogo mobil'nogo robota [Development of technical vision system for service mobile robot]. Tretiy Vserossiyskiy nauchno-prakticheskiy seminar "Bespilotnye transportnye sredstva s elementami isskustvennogo intellekta" [*The Third All-Russian Scientific and Practical Seminar "Unmanned Vehicles with Artificial Intelligence Elements", Innopolis, 22–23 September 2015*]. Innopolis, Universitet Innopolis, 2016, pp. 127–136 (in Russian).
5. Uijlings J. R. R., Sande K. E. A., Gevers T., Smeulders A. W. M. *Selective search for object recognition*. Available at: <http://www.huppelen.nl/publications/selectiveSearchDraft.pdf> (accessed 28.04.2018).
6. Ren S., He K., Girshick R., Sun J. *Faster R-CNN: Towards real-time object detection with region proposal networks*. Available at: <https://arxiv.org/pdf/1506.01497.pdf> (accessed 15.05.2018).
7. Szegedy C., Vanhoucke V., Ioffe S., Shlens J. *Rethinking the inception architecture for computer vision*. Available at: <https://arxiv.org/pdf/1512.00567.pdf> (accessed 15.05.2018).
8. LeCun Y., Cortes C., Burges C. J. C. *The MNIST database*. Available at: <http://yann.lecun.com/exdb/mnist/> (accessed 16.03.2018).
9. Lin T., Patterson G., Ronchi M. R., Cui Y., ..., Dollar P. *Microsoft COCO: Common objects in context*. Available at: <http://cocodataset.org/#home> (accessed 16.03.2018).
10. *Tzatalin / labelImg*. Available at: <https://github.com/tzatalin/labelImg> (accessed 18.06.2018).

Информация об авторе

Ходасевич Любовь Александровна – стажер младшего научного сотрудника лаборатории робототехнических систем, Объединенный институт проблем информатики НАН Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь).

E-mail: liubov.hodasevich@gmail.com

Information about the author

Liubov A. Khodasevich – Trainee of Junior Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganova Str., 220012, Minsk, Republic of Belarus).

E-mail: liubov.hodasevich@gmail.com