

Распознавание и локализация дорожных знаков

В. В. Ковалев¹, Е. А. Масленникова², Д. М. Филатов³, А. О. Федоркова⁴

Санкт-Петербургский государственный электротехнический университет
«ЛЭТИ» им. В.И. Ульянова (Ленина)

¹vladimir.kovalev.2012@bk.ru, ²katerinamaslennikova1494@gmail.com,
³dmfilatov@etu.ru, ⁴anastasiikanata@gmail.com

Аннотация. Доклад рассматривает проблемы распознавания дорожных знаков посредством компьютерного зрения и нейронных сетей, также уделено внимание вопросу локализации распознанных знаков относительно некоторого объекта управления. Рассматривается вопрос генерации изображений для обучения нейронной сети на реальных, смешанных и искусственных данных, с целью сравнения подходов к обучению нейронной сети.

Ключевые слова: распознавание дорожных знаков; компьютерное зрение; нейронные сети; генерация изображений; локализация объекта

I. ВВЕДЕНИЕ

С момента появления первых автомобилей мир неустанно менялся: промышленные гиганты формировали новые городские ландшафты, делая даже самые отдаленные уголки земли доступными для жизни и передвижения. Автомобили стали незаменимыми помощниками в перевозке людей и грузов, что способствовало развитию транспортной инфраструктуры. Все дороги рано или поздно стали приобретать элементы регулирования движения автомобилей – от простых знаков до целых светофоров.

Параллельно с технологическим прогрессом шагала вперед и вычислительная мощь компьютеров. Появились новые технологии, среди которых и искусственный интеллект (ИИ), способный анализировать огромные массивы данных и делать на их основе точные прогнозы. Хотя до эры полностью автономных автомобилей, управляемых единым ИИ, еще далеко, уже видны первые шаги в этом направлении в виде систем ADAS (англ. Advanced Driver Assistance Systems) – интеллектуальных систем помощи водителям. Основная функция этих систем – повышение безопасности на дорогах, что в конечном итоге должно привести к снижению числа дорожно-транспортных происшествий. [1].

Такие системы часто оснащены функциями предупреждения о съезде с полосы, опасности столкновения и многими другими. Рассмотрим подробнее задачу обнаружения дорожных знаков на изображении с камеры.

II. СПОСОБЫ РАСПОЗНАВАНИЯ ДОРОЖНЫХ ЗНАКОВ

Существует два основных направления при решении задачи обнаружения и распознавания объектов с помощью камеры.

A. Машинное зрение

Первое направление – это использование машинного зрения, которое основывается на анализе определенных шаблонов и геометрических форм, таких как круги,

квадраты или треугольники. Подход включает в себя распознавание цвета объекта и оценку его размеров в пикселях на изображении.

На рис. 1 приведен пример работы такого алгоритма:



Рис. 1. Результат работы программы машинного зрения на примере опознавания дорожного знака «пешеходный переход» слева и знака «движение запрещено» справа

Из рис. 1 видно, что система способна обнаружить контур знака, но нельзя утверждать, что данный результат свидетельствует о корректном определении самого знака.

Отдельно отметим достоинства и недостатки такого подхода.

Достоинства: простота в реализации; адаптация алгоритма под все существующие классы дорожных знаков; возможность не только обнаружения, но и базового распознавания объектов. [2].

Недостатки: чувствительность к изменениям погодных условий; вероятность ошибочного распознавания объектов, схожих по форме с дорожными знаками. [2].

B. Нейронные сети

Второе направление – это использование нейронных сетей, которые представляют собой более сложные и мощные системы, способные анализировать изображения с высокой точностью. Нейронные сети могут использовать различные архитектуры или их ансамбли для более детального анализа изображений, позволяя не только обнаруживать объекты, но и выделять их на изображении рамками, указывая дополнительные параметры, например, как показано на рис. 2:



Рис. 2. Результат работы программы, использующей нейронную сеть для распознавания дорожных знаков

Из рис. 2 видно, что нейронная сеть успешно идентифицировала дорожный знак круглой формы с вероятностью распознавания в 82%. Однако стоит отметить, что система обнаружила знак только в левой части кадра, в то время как аналогичный знак, расположенный справа, остался без внимания, что говорит уже о некачественном результате работы нейронной сети.

Для реализации нейронной сети как элемента программы компьютерного зрения существуют библиотеки, среди которых особенно выделяются PyTorch и TensorFlow. Данные платформы предлагают широкий спектр предварительно обученных моделей, которые можно адаптировать под конкретные задачи. Так называемый «зоопарк моделей» включает в себя разнообразие архитектур, предназначенных для обработки различных типов данных и решения множества задач, начиная от классификации изображений (определение общих характеристик изображения без выделения конкретных объектов) и заканчивая сложными задачами распознавания объектов (локализация и идентификация конкретных объектов на изображении, с возможностью подсчета их количества).

Каждый из этих подходов имеет свои уникальные характеристики и применяется в зависимости от конкретных задач и требований к точности распознавания. При обучении нейронных сетей важным показателем является метрика точности, которая оценивается через сравнение с валидационным набором данных и рассчитывается с помощью формул 1 и 2, учитывающих правильность и ошибочность распознавания объектов:

$$mAP = \frac{1}{|classes|} \sum_{c \in classes} \frac{|TP_c|}{|FP_c| + |TP_c|} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

где mAP – это среднее значение точности для всех классов, Precision – точность определения одного класса, TP – количество верно определенных распознаваний, а FP – количество неверно определенных распознаваний. [3].

III. РАСПОЗНАВАНИЕ ДОРОЖНЫХ ЗНАКОВ

При подробном рассмотрении вопросов, связанных с задачей распознавания объектов, принято решение использовать архитектуру YOLOv5, которая имеет варианты своей реализации для встраиваемых систем, а также вполне хорошо работает с изображениями для поиска конкретных объектов.

A. Сбор и обработка данных

Для создания системы, которая помогала бы водителям или даже управляла автомобилем автономно, необходимо собрать и подготовить датасет. Для данной работы мы собрали несколько видеорядов с камеры, установленной по центру лобового стекла, разбитых на отдельные кадры и размеченных вручную. Сбор данных преимущественно проводился в условиях дневного освещения на дороге. Учитывая разнообразие дорожных знаков, мы сосредоточились на их геометрической форме для первичного распознавания, после чего

производится более детальный анализ уже выделенной формы для классификации самого знака.

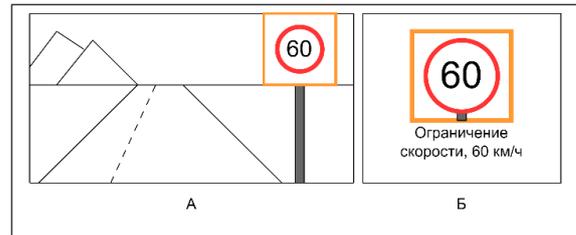


Рис. 3. Схема алгоритма распознавания дорожного знака: А – режим распознавания объекта, Б – режим классификации меньшего изображения с целью извлечения содержания информации знака

Этот подход, схематично представленный на рис. 3, подразумевает двухэтапную обработку изображения. На первом этапе алгоритм фокусируется на определении внешнего вида объекта, акцентируя внимание на его геометрической форме. На втором этапе производится анализ выделенного изображения для извлечения содержимого знака, что включает классификацию изображения и, при необходимости, распознавание текста. Для дорожных знаков выделено десять основных категорий в зависимости от их формы и функционала, что облегчает процесс классификации:

Согласно форме дорожных знаков можно подразделить искомый объект на 10 классов простых геометрических фигур, включая классы для светофора и класс, содержащий только текстовое табло.

При формировании набора данных было собрано 6000 уникальных изображений, часть из которых была размечена вручную, а часть размечена в автоматическом режиме. Отдельно создана тестовая выборка.

B. Результат обучения на реальных данных

В процессе обучения были получены метрики для оценки качества работы нейронной сети. Среди них особое внимание заслуживает метрика средней точности (mAP), определенная в уравнении (1). Данный показатель представлен на графике ниже:

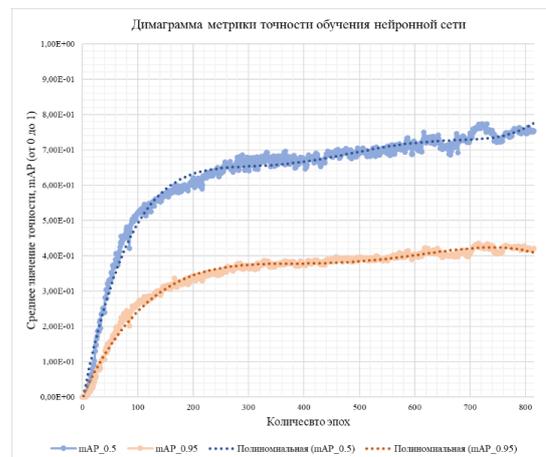


Рис. 4. График метрики точности mAP при обучении на реальных данных

На рис. 4 четко видно, что в течение 815 эпох обучения сети значение mAP для обнаружения объектов

с точностью до 50 % колеблется между 0.7 и 0.8, а для точности от 50 % до 95 % — между 0.4 и 0.5. Одним из наиболее наглядных способов оценить результативность обучения является тестирование алгоритма на данных, с которыми сеть ранее не сталкивалась. Результат обработки такого изображения представлен на рис. 5:



Рис. 5. Результат работы алгоритма распознавания знаков на тестовом изображении

Из рис. 5 видно, что не все знаки были успешно распознаны: знак, расположенный справа, не был идентифицирован, поскольку в наборе данных отсутствовали аналогичные образцы. Тем не менее, другие знаки на изображении были обнаружены корректно и с вероятностью выше 80 %. Повторим, что этого кадра, как и всего маршрута из набора данных, нету в обучающей выборке.

С. Генерация данных

К сожалению, не все дорожные знаки могут быть равномерно представлены в городской среде, а следовательно, и включены в набор данных для обучения. То же касается и разнообразия дорожных ситуаций: не всегда возможно учесть и воссоздать их все в реальных условиях. Однако существует возможность обойти эти ограничения с помощью программ для генерации данных. В нашем случае мы можем создать трехмерные модели дорожных знаков и, используя специализированные программы, сгенерировать фотореалистичные изображения этих знаков.

Процедуру генерации данных проведем на примере конкретного дорожного знака, чтобы затем сравнить эффективность обучения нейронной сети на основе реальных и сгенерированных данных. Знак будет расположен в центре кадра, в то время как перспектива камеры будет варьироваться, создавая различные углы обзора. Кроме того, мы добавим вариации угла падения солнечных лучей, чтобы сделать сцену еще более реалистичной. На рис. 6 представлен один из примеров сгенерированного изображения:



Рис. 6. Пример изображения, созданного с помощью генерации данных

Из рис. 6 видно, что знак находится в центре кадра, но под некоторым углом, а фон имитирует облака и некоторую холмистую местность. Хотя такой визуализации недостает элементов реальной дорожной обстановки, таких как само дорожное полотно, автомобили и прохожие, она содержит все ключевые характеристики, необходимые для обучения и тестирования алгоритмов компьютерного зрения и

нейронных сетей в задачах распознавания дорожных знаков. Рассмотрим результаты, полученные в ходе такого подхода.

Д. Результат обучения на сгенерированных данных

При обучении сети на сгенерированных данных можно рассмотреть график точности, но, несмотря на впечатляющие показатели, такие данные будут хорошо работать только в сгенерированном мире. Рассмотрим диаграмму метрики точности (рис. 7) и работу алгоритма обнаружения знаков на реальных изображениях (рис. 8):



Рис. 7. График метрики точности mAP при обучении на искусственных данных, оранжевым метрика точности для вероятности объекта от 50% до 95%, синим – до 50%

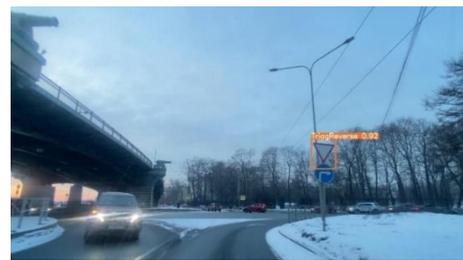


Рис. 8. Результат работы нейронной сети, обученной на синтетических данных, с вероятностью в 92 %, что на изображении присутствует знак «Уступи дорогу».

Из анализа рис. 7 и 8 видно, что использование сети, обученной на искусственном наборе данных, оправдано, однако для достижения удовлетворительных результатов в реальном мире, такой подход недостаточен.

Е. Смешанное обучение

Под смешанным обучением понимается обучение нейронной сети на наборе данных, состоящем как из сгенерированных, так и из реальных данных.

Рассмотрим полученный результат и сравним все три подхода к обучению сети, учитывая показатель mAP, количество изображений, разность точности определения одного знака «Уступи дорогу», так как именно его изображение и было сгенерированным. Показатели представлены в табл. 1:

ТАБЛИЦА 1.

Режим обучения	Выборка	mAP_0.5:0.95	1Т	2Т	3Т	4Т
Реальная выборка	2245	0.45255	0.83	0.90	0.90	-0.49
Искусственная выборка	2592	0.98103	0.95	0.88	0.62	0
Смешанная выборка	4837	0.53255	0.88	0.92	0.94	0.60

Анализ результатов, представленных в таблице, показывает, что внедрение синтетических данных в процессе смешанного обучения приводит к значительному улучшению показателей. Отдельно отметим, что параметр -0.49 и 0 обозначают неверное распознавание и отсутствие распознавания, соответственно. Рассмотрено формирование наборов данных от 2000 изображений согласно [4].

IV. ЛОКАЛИЗАЦИЯ ДОРОЖНЫХ ЗНАКОВ

Помимо распознавания дорожных знаков, также существует задача определения их местоположения, которую успешно решают системы компьютерного зрения.

Особое внимание в данном контексте уделяется использованию одиночной камеры, что исключает применение методов, основанных на создании карт глубины или стереозрения. В условиях городской среды чаще всего встречаются дорожные знаки второго и третьего типоразмеров, что определяет их физический размер. Зная первоначальные размеры объекта, можно вычислить его видимые размеры на изображении в зависимости от расстояния до него, используя (3):

$$S = \frac{FH}{h}, \quad (3)$$

где S – это расстояние до объекта наблюдения, F – фокусное расстояние объектива, H – горизонтальный размер объекта, h – размер матрицы по горизонту.

Таким образом, можно рассчитать расстояние до дорожного знака с точностью до 10 сантиметров. Однако точность такого измерения может снижаться из-за присутствия знаков других размеров на дороге, что может привести к ошибке в оценке расстояния от одного метра и более. Отдельное внимание необходимо уделить расположению знака на самом изображении. Для этого визуально можно разделить обстановку на три секции, схематично указанных на рис. 9: зона непосредственного действия знака (А), зона предстоящего действия знака (Б) и зона, где знак ещё не действует (В):



Рис. 9. Изображение деления картинка на секции, для принятия решения относительно дальности знаков

Рис. 9 показывает сетку проекции перспективы на окружение, которая делит изображение на 7 частей по горизонтали и на 3 части по вертикали. Видно, что уже вступает в действие знак в секции «А» – сужение дороги, а пешеходный переход и все его знаки в зоне «В» – далеко.

Значения «близко» и «далеко» очень расплывчаты и не имеют чётких значений, но принимая во внимание факт наступления действия знака, можно не рассчитывать с точностью до сантиметров, а полагаться на размер опознанного знака и его положение на изображении.

V. ЗАКЛЮЧЕНИЕ

В результате работы продемонстрировано повышение точности определения знаков при добавлении сгенерированных данных, написана программа для генерации изображений дорожных знаков. На тестовой выборке получено значительное улучшение алгоритма распознавания относительно одного знака «Уступи дорогу», также полученные выводы относительного одного знака применимы и к остальным. Также установлено, что для получения значений успешного распознавания знаков на изображении, при их непосредственной близости, достаточно от 2000 изображений.

СПИСОК ЛИТЕРАТУРЫ

- [1] ADAS-системы. Что это и для чего они нужны? [Электронный ресурс]. – Режим доступа: <https://advantum.ru/news-old/tpost/kudnyngi51-adas-sistemi-cto-eto-i-dlya-chego-oni-n> (дата обращения: 14.03.2024).
- [2] Енокаев Р.М. Алгоритм распознавания дорожных знаков // Прикладная математика: современные проблемы математики, информатики и моделирования, 2020. Кубанский гос. университет.
- [3] Mean Average Precision (mAP) in Object Detection [Электронный ресурс]. – Режим доступа: <https://learnopencv.com/mean-average-precision-map-object-detection-model-evaluation-metric/> (дата обращения: 14.03.2024).
- [4] Wenkao Yang, Wei Zhang. Real-time Traffic Signs Detection Based on YOLO Network Model // 2020 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). С 357.