

А. Л. Приоров, М. А. Кулагин
ФГБОУ ВО Ярославский государственный университет им. П.Г. Демидова

Применение нейронной сети в задаче детектирования железнодорожных светофоров

Представлен алгоритм детектирования железнодорожных светофоров с использованием нейронных сетей. Для решения поставленной задачи собрана обучающая и тестовая база изображений, спроектирована нейронная сеть. Тестирование проходило с использованием различных модификаций сети, а именно – использование разных целевых функций, необходимых для обучения по методу градиентного спуска. Так же было проведено сравнение работы нейронной сети с алгоритмом Виола-Джонса на той же базе изображений. В будущем алгоритм, построенный на основе нейронных сетей, может составить часть системы детектирования и распознавания объектов железнодорожной инфраструктуры на изображении, а также будет полезен инженерам в проектировании систем видеонаблюдения.

Нейронные сети, алгоритм Виола-Джонса, детектирование, целевая функция.

Введение

В последнее время большое внимание уделяется созданию интеллектуальных транспортных средств. Уже сейчас существуют автомобили, движущиеся без помощи человека, беспилотные самолеты и другие транспортные средства. Для того чтобы автомобиль мог двигаться в пространстве по всем правилам дорожного движения и принимать решения за человека, требуется видеть и анализировать окружающее пространство. Поэтому сейчас одним из самых популярных приложений машинного обучения является техническое зрение [1, 2].

В работе затронута тема детектирования железнодорожных светофоров с применением нейронной сети [3] и алгоритма Виола-Джонса. Задача является актуальной, так как алгоритм поиска области интереса на изображении в будущем будет лежать в основе системы, позволяющей локомотиву двигаться без участия машиниста. Создание цифровой железной дороги является на данный момент одним из приоритетных направлений развития данной области, поэтому решение задачи детектирования объектов железнодорожной инфраструктуры позволит ускорить процесс развития интеллектуального транспорта.

Применение алгоритмов машинного обучения и технического зрения для решения научно-технических задач требуют создание большой обучающей и тестовой базы данных. На данный момент уже существуют готовые базы изображений различных объектов, но задача детектирования железнодорожных светофоров требует создание собственной базы, необходимой для получения высоких результатов.

Формирование базы изображений

Для формирования соответствующей базы использовался набор из 4305 изображений размером 64x64 пикселя (рис. 1).

Алгоритмы, спроектированные с применением нейронных сетей, требуют гораздо большего объема данных, поэтому первоначальное количество изображений в базе было увеличено с применением: аффинных преобразований (рис. 2а), зеркального отражения (рис. 2б), выравнивания гистограммы (рис. 2в).



Рис. 1. Примеры тестовых изображений из базы: а) изображения с объектом интереса (положительная выборка); б) изображения без объекта интереса (отрицательная выборка)



Рис. 2. Примеры полученных изображений: а) аффинные преобразования – растяжение изображения; б) зеркальное отражение изображения; в) выравнивание гистограммы

Первым вариантом расширения базы данных изображений являются аффинные преобразования, а именно – растяжение. Изображения были изменены с применением функции `warpAffine()` из библиотеки `OpenCV`. Второй вариант увеличения количества изображений является зеркальное отражение. Третий вариант – выравнивание гистограммы, которое осуществлялось с использованием стандартной функции `equalizeHist` из библиотеки `OpenCV`.

В итоге, после расширения базы изображений, их конечное количество для обучения и тестирования составило 10000 изображений. Вся база изображений была разбита на три части: обучающая база – 7500 изображений (3007 положительная выборка и 4493 отрицательная выборка); база для валидации – 1500 изображения (750 положительная выборка и 750 отрицательная выборка); база для тестирования – 1000 изображений (500 положительная выборка и 500 отрицательная выборка). Все изображения, используемые для обучения, были размечены. Метка “1” означает присутствие на изображении железнодорожного светофора, метка “0” – его отсутствие.

В работе сравниваются алгоритмы обучения, используемые для детектирования области интереса на изображении, с применением нейронной сети и алгоритма Виола-Джонса.

Алгоритм Виола-Джонса

Тестирование алгоритма Виола-Джонса проходило с использованием готовых функций из библиотеки `OpenCV`, предназначенной для компьютерного зрения. Более подробное описание алгоритма Виола-Джонса представлено в [4]. В этой работе, благодаря существующим программным решениям, то есть готовым библиотекам и наборам функций в них, обучение проходило в два этапа. На первом этапе применялась

функция `opencv_createsamples()`, которая приводила весь положительный набор изображений, то есть изображения с железнодорожным светофором, в общий формат и единый файл `sample.vec`. Второй этап заключался в создании итогового каскада классификаторов с использованием функции `opencv_traincascade()`. В итоге, результаты работы алгоритма показали в среднем 91% верного детектирования объекта на изображении. Работа алгоритма, протестированного в задаче детектирования железнодорожных светофоров, представлена на блок схеме (рис. 3б).



Рис. 3. Детектирование с применением: а) нейронной сети; б) алгоритма Виола-Джонса

Нейронная сеть

Так же в данной работе, помимо алгоритма Виола-Джонса, проектируется и используется нейронная сеть с одним скрытым слоем (рис. 4). Ее обучение проводится с помощью метода градиентного спуска. На основании алгоритмов машинного обучения и компьютерного зрения, построенная нейронная сеть обучается, затем тестируется (рис. 3а). Реализовался алгоритм на языке программирования Python с использованием базовых библиотек OpenCV и NumPy. Примеры реализации других алгоритмов детектирования, но только автомобильных светофоров представлены в [5–6].

Самый левый слой нейронной сети является входным слоем, состоящим из 4096 нейронов. Крайний правый – является выходным слоем и состоит из 2 нейронов, средний слой называют скрытым и состоит из 50 нейронов (рис. 4).

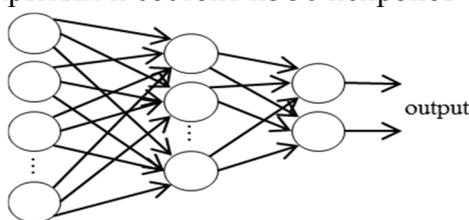


Рис. 4. Схематичная структура нейронной сети

Каждое изображение преобразуется в размер 4096x1, для того чтобы на каждый нейрон подать значение только одного пикселя для последующего обучения сети. На выходе нейронной сети активируется один из двух нейронов, демонстрирующий присутствие или отсутствие на изображении заданного объекта.

Обучения основывается на методе градиентного спуска [7], для подбора весов и смещений нейронной сети. Для количественной оценки того, насколько хорошо подобраны веса используется целевая функция (cost function). Существуют различные целевые функции. В данной работе приводится сравнение двух из них: квадратичной и функции кросс-энтропии:

$$C(w, b) = \frac{1}{2n} \sum_x (y(x) - a)^2$$

$$C(w, b) = -\frac{1}{n} \sum_x [y(x) \ln a + (1 - y(x)) \ln(1 - a)],$$

где w – совокупность всех весов нейронной сети; b – все смещения; n – совокупность всех учебных примеров; a – вектор выходов сети, когда x действует на входе сети. Задача состоит в подборе таких весов и смещений, чтобы функция затрат была минимальной. Метод градиентного спуска чаще всего используется, когда параметры (веса и смещения) функции не могут быть рассчитаны и должны быть найдены с помощью алгоритма оптимизации. Изначально подбираются произвольные значения параметров, после чего с каждым шагом коэффициенты меняются таким образом, чтобы значение функции уменьшалось. Изменение функции затрат относительно предыдущего положения вычисляется с помощью градиента функции.

В итоге нейронная сеть была обучена с использованием двух целевых функций, по очереди. В табл. 1 представлены результаты обучения на различных эпохах (эпоха – это целый цикл обучения и тестирования сети на всей обучающей и тестовой базе изображений). Видно, что функция кросс-энтропии увеличивает скорость обучения нейронной сети и результаты точности работы алгоритма на тестовой базе.

Обучение нейронной сети с использованием двух целевых функций

Таблица 1.

Целевая функция	Эпоха обучения	Точность работы алгоритма на тестовой базе
Квадратичная функция	10	54
	50	78
	80	93
Кросс-энтропия	10	77
	50	85
	80	94

Выводы

С помощью алгоритмов Виола-Джонса и нейронной сети, реализованы и протестированы алгоритмы поиска объекта на изображении. Выяснено, что использование кросс-энтропии как целевой функции для обучения нейронной сети значительно сокращает время на ее обучение и увеличивает точность работы алгоритма. Алгоритм Виола-Джонса показал несколько худшие результаты, чем нейронная сеть.

Для дальнейшего улучшения результатов требуется увеличение базы данных изображений, а также применение глубоких нейронных сетей. Описанные алгоритмы в

будущем могут составить часть системы детектирования и распознавания объектов железнодорожной инфраструктуры на изображении.

Библиографический список

1. Шапиро Л., Стокман Дж. Компьютерное зрение. М. БИНОМ, Лаборатория знаний, 2006.
2. Форсайт Д.А., Понс Д. Компьютерное зрение. Современный подход. М.: «Вильямс», 2004.
3. Осовский С. Нейронные сети для обработки информации. М. Финансы и статистика, 2002.
4. Viola P., Jones M. Rapid Object Detection using a Boosted Cascade of Simple Feature. Mitsubishi Electric Research Labs, Cambridge, 2001.
5. Zhou X., Yuan J., Liu H. A traffic light recognition algorithm based on compressive tracking. Beijing Key Laboratory of Information Service Engineering, Beijing Union University, 2015.
6. Hwang T.H., Joo I.H., Cho S.I. Detection of Traffic Lights for Vision-Based Car Navigation System. Lecture Notes in Computer Science (PSIVT '06), Vol.4319, Dec. 2006. pp. 682–691.
7. Michael N. Neural Networks and Deep Learning . 2015. Chapter № 1.

APPLICATION OF A NEURAL NETWORK IN THE TASK OF DETECTING RAILWAY TRAFFIC LIGHTS

Priorov A. L., Kulagin M. A.

P.G. Demidov Yaroslavl State University

This article presents an algorithm for detecting railway traffic light. This is algorithm based on using neural network. In this task was assembled a training and test image database, the neural network is designed. The testing was carried out using various cost functions required for gradient descent training. Also, a comparison was made of the work of a neural network with the algorithm Viola-Jones.

In the future, the algorithm will form part of the system for detecting and recognizing railway infrastructure objects on the image, and will also be useful to engineers in the design of CCTV.

Neural network, algorithm Viola-Jones, detection, cost function.