

Система распознавания дорожных знаков на основе методов машинного обучения

Д. М. Филатов, А. В. Девяткин, Е. В. Серых
СПбГЭТУ «ЛЭТИ»

J. Heikkinen
Independend researcher

А. С. Шабуров

Пермский национальный исследовательский политехнический университет

Abstract. In this paper, traffic sign detector was realized with utilization of methods of machine learning field. Out of great variety of methods in machine learning the most known and reliable were selected. The utilization of deep learning neural network of python library “Keras” was approved by requirement to work with camera frames and use the encoder-decoder structure because of its simplicity in design in comparison with hard-coded methods using frame transformations. The classifier part of the system utilizes the method of Linear Support Vector Machine of the library “scikit-learn” that is lightweight method of distributing the space of features into classes and preprocessing the resulting picture of localized sign to HOG feature space helps to create feature vector and realize classification. The hardware for realization was NVIDIA Jetson TX2 that is GPU for embedded systems. The system was supposed to use in the autonomous small size car 1:8 to real one.

Keywords: machine learning; neural network; deep learning; support vector machine; traffic signs

I. ВВЕДЕНИЕ. ПОСТАНОВКА ПРОБЛЕМЫ.

Создание автономного транспорта, способного перемещаться по дорогам общего пользования наравне с другими участниками движения, является одной из наиболее перспективных задач современного автомобилестроения. Решение данной задачи в будущем позволит существенно повысить безопасность на дорогах, полностью исключив негативное влияние человеческого фактора. Промежуточным шагом в решении данной задачи является внедрение систем активной помощи водителю (ADAS). Однако даже на этом этапе, когда окончательное решение по управлению транспортным средством в большинстве случаев принимает водитель, требования к корректности работы ADAS очень высоки, поскольку неправильное распознавание дорожной ситуации и подача искаженной информации водителю может спровоцировать аварийную ситуацию.

Разработки и исследования в области автономных транспортных средств ведутся в настоящее время очень активно. В качестве примера можно привести исследования построения маршрута движения транспортного средства в работе [1], разработка модуля логики принятия решений в поведении автомобиля [2], использование технологий автономных транспортных средств для помощи навигации слепым людям [3], разработка многофункциональной системы для автономной машины на основе разнородных датчиков измерения расстояния [4].

Одной из ключевых подсистем в системе активной помощи водителю является система распознавания дорожных знаков. В качестве основного источника данных для распознавания знаков дорожного движения наиболее часто используется изображения с камер, так как основная информация знака закодирована с помощью формы и цвета. Расшифровка данной информации осуществляется на основе методов компьютерного зрения. Данный подход имеет высокий потенциал к развитию, но используемые в настоящее время алгоритмы недостаточно точны, безошибочны и устойчивы к изменению внешних условий [5–9], чтобы полностью полагаться на результаты их работы при управлении транспортным средством.

В настоящей работе рассматривается подход с использованием методов машинного обучения для решения задачи распознавания знаков дорожного движения с учетом ограничений в вычислительных ресурсах, а также на основе уменьшенного набора требуемых знаков, которые используются для задания возможных направлений движения транспортного средства.

II. ОПИСАНИЕ СИСТЕМЫ

Разрабатываемая система предполагается к использованию на миниатюрном колесном мобильном роботе с рулевой системой, основанной на механизме Аккермана, представляющем собой уменьшенную копию шасси автомобиля в масштабе 1:8. Ввиду небольших размеров робота в качестве вычислителей были использованы два микрокомпьютера: Intel NUC и NVIDIA Jetson TX2. Первый представляет из себя полноценный CPU на архитектуре x86. Вторым микрокомпьютер представляет собой GPU на архитектуре Pascal и вычислительной мощностью 3 TFLOPS. Последний используется для организации параллельных вычислений при работе нейронной сети глубокого обучения.

Сама система состоит из двух частей, первая производит бинаризацию изображения, создавая карту вероятностей нахождения знака в областях изображения. Дальнейшее маскирование позволяет выделить части изображения, в которых с высокой вероятностью находится знак дорожного движения. Полученные участки представляют собой масштабированные секции, которые могут либо содержать один из интересующих знаков, либо являться ошибкой локализации и не содержать ни одного знака вовсе. Такого рода ошибки предполагаются возможными в данной си-

стеме, ввиду требования к низкой сложности сети для получения высокой скорости работы.

III. ДЕТЕКТИРОВАНИЕ ОБЛАСТЕЙ

В разрабатываемой системе распознавания знаков подсистема построения карты вероятности расположения знака выполняет задачу детектирования знаков на изображении. Решение данной задачи можно найти в различных источниках с использованием таких методов как использование нейронной сети глубокого обучения для получения описывающего прямоугольника знака [10], подход с оценкой частей знака и взаимоотношений между ними [11–13].

Разрабатываемая система использует сеть глубокого обучения, построенную по принципу «энкодер-декодер». Тип структуры нейронных сетей глубокого обучения под названием «Fully convolutional Neural Network» (FCNN) активно применяется для решения задач семантической сегментации, что в свою очередь является более сложной задачей бинаризации изображения с выходом по количеству сегментируемых классов. Пример структуры сети представлен на рис. 1.

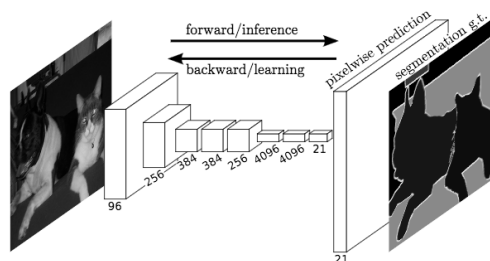


Рис. 1. Пример структуры сети FCNN

В основе работы такого рода сетей лежит принцип выделения специальных признаков (Features) на изображении и дальнейшее их преобразование в матричное представление, которое часто является масштабированием исходного изображения.

За основу сети глубокого обучения была взята распространенная структура сети U-net, что в результате позволило построить сеть уменьшенных размеров, но подобной структуры.

Обучение проводилось на изображениях со знаками шести видов: только прямо, только налево, только направо, только прямо и налево, только прямо и направо, движение прямо запрещено. Результат работы сети с выделением найденной области при наличии засветки солнечным светом представлен на рис. 2.



Рис. 2. Пример работы подсистемы локализации знака

Для обучения сети в качестве функций оценки были использованы коэффициент Сёренсена (1) и коэффициент Жаккара (2).

$$S(A, B) = \frac{|A \cap B|}{|A| + |B|} \quad (1)$$

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

В результате обучения сети коэффициент Сёренсена показал более высокую скорость сходимости и большую устойчивость при большом коэффициенте обучения Стохастического градиентного спуска.

IV. КЛАССИФИКАЦИЯ ЗНАКА

Дальнейшая работа системы после построения карты вероятностей и выделения вероятных областей знаков заключается в классификации сегмента изображения одним из семи классов. Первые шесть классов представляют указанные ранее знаки, которые использовались для обучения сети локализации. Седьмой класс представляет собой описание ошибочного представления, часто его принято называть «фон».

За основу метода классификации был взят алгоритм Linear Support Vector Machine (Линейная Машина Опорных Векторов). Особенностью данного метода является разделение пространства признаков на основе принципа максимума границ (рис. 3).

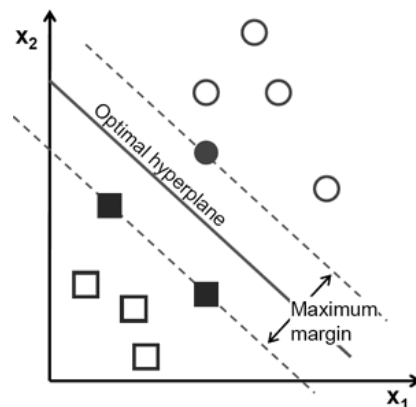


Рис. 3. Пример работы метода LSVM

Для представления выделенного сегмента изображения в виде набора признаков был использован метод Гистограммы Направленных Градиентов (Histogram of Oriented Gradient). Принцип работы данного метода заключается в представлении изображения в виде вектора гистограмм, построенных по градиенту изображения. Такой метод поддерживает инвариантность перемещений объекта на изображении.

На тестовой выборке, состоящей из ста изображений, классификатор показал 87% по оценке F-мерой.

V. ЗАКЛЮЧЕНИЕ

Была разработана система распознавания знаков, позволяющая оценивать и классифицировать установленные дорожные знаки, находящиеся на пути движения автомобиля для принятия решений относительно дальнейшего движения. Использование LSVM и метода HOG для классификации не требует высокой вычислительной мощности, чего нельзя сказать о глубокой нейронной сети, которая при встраивании показывает скорость 14 FPS на Intel Core i5 7th gen и 54 FPS на NVIDIA Jetson TX2. Такое различие в скоростях подтверждает целесообразность использования параллельных вычислителей при использовании нейронных сетей.

СПИСОК ЛИТЕРАТУРЫ

- [1] Gaku Takano, Makoto Obayashi, Keisuke Uto, «Posture-to-posture trajectory planning for an autonomous car using triple clothoid segments», 2017 IEEE 56th Annual Conference on Decision and Control (CDC), Melbourne, VIC, Australia, DOI: 10.1109/CDC.2017.8263682
- [2] Bingqing Xu, Qin Li, «A bounded multi-dimensional modal logic for autonomous cars based on local traffic and estimation», 2017 International Symposium on Theoretical Aspects of Software Engineering (TASE), Sophia Antipolis, France, France, DOI: 10.1109/TASE.2017.8285637
- [3] Manuel Martinez, Alina Roitberg, Daniel Koester, Rainer Stiefelbogen, Boris Schauerte, «Using Technology Developed for Autonomous Cars to Help Navigate Blind People», 2017 IEEE International Conference on Computer Vision Workshop (ICCVW), Venice, Italy, DOI: 10.1109/ICCVW.2017.169
- [4] Ayesha Iqbal, Syed Shaheryar Ahmed, Muhammad Danish Tauqeer, Ali Sultan, Syed Yasir Abbas, «Design of multifunctional autonomous car using ultrasonic and infrared sensors», 2017 International Symposium on Wireless Systems and Networks (ISWSN), Lahore, Pakistan, DOI: 10.1109/ISWSN.2017.8250023
- [5] M. Shneier, «Road sign detection and recognition», Proc. IEEE Computer Society Int. Conf. on Computer Vision and Pattern Recognition, pp. 215–222, 2005.
- [6] A. Ruta, F. Porikli, Y. Li, S. Watanabe, H. Kage, K. Sumi, «A New Approach for In-Vehicle Camera Traffic Sign Detection and Recognition», IAPR Conference on Machine Vision Applications (MVA), Session 15: Machine Vision for Transportation, 2005.
- [7] R. Belaroussi, P. Foucher, J.P. Tarel, B. Soheilian, P. Charbonnier, N. Paparoditis, «Road Sign Detection in Images», A Case Study, 20th International Conference on Pattern Recognition (ICPR), pp. 484–488, 2010.
- [8] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, C. Igel, «Detection of Traffic Signs in Real-World Images: The German Traffic Sign Detection Benchmark», International Joint Conference on Neural Networks, 2013.
- [9] J. Stallkamp, M. Schlipsing, J. Salmen, C. Igel, «Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition», Neural Networks, Vol. 32, pp. 323–332, 2012
- [10] Erhan D., Szegedy C., Toshev A., Anguelov D. Scalable object detection using deep neural networks. arXiv:1312.2249v1 [cs.CV] 8 Dec 2013, URL: <https://arxiv.org/pdf/1312.2249.pdf>.
- [11] Fergus, R., Zisserman, P.P.A. Weakly Supervised Scale-Invariant Learning of Models for Visual Recognition. IJCV 71 (2007), p. 273–303.
- [12] Bouchard, G., Triggs, B. Hierarchical part-based visual object categorization. In: CVPR, Washington, DC, USA, IEEE Computer Society (2005), p. 710-715.
- [13] Felzenszwalb, P., McAllester, D., Ramanan, D. A Discriminatively Trained, Multiscale, Deformable Part Model. In: CVPR. (2008).