

А.А. Золкин, В.С. Тормозов, Ю.В. Гуменникова, И.В. Степина

ТЕЗИСЫ К ВОПРОСУ МЕТОДОЛОГИИ ПРОГРАММНОГО
ОБЕСПЕЧЕНИЯ ПРИМЕНЕНИЯ ТЕХНОЛОГИИ НЕЙРОСЕТЕВОГО
АППАРАТА РАСПОЗНАВАНИЯ

Аннотация. Рассматривается общая проблема задачи локализации и классификации объектов в растровых изображениях. На данный момент не существует универсальных методов и технологий, позволяющих идеально решить задачу локализации и классификации объектов на изображении. Проанализированы наиболее популярные и актуальные методы для локализации и классификации объектов на изображении. Разработан и приведен алгоритм технологии распознавания. Продемонстрированы результаты работы программы. Обобщена целевая карта использования алгоритмического метода. Для решения задач, поставленных в рамках данного исследования, было принято решение использовать дескрипторы HOG&SVM в модуле локализации, а локализованное изображение передавать на CNN, реализующие модуль классификации. Рассматривается и обосновывается алгоритм выработки собственной модели проектирования технологии распознавания, базирующейся на методологии применения нейросетевого аппарата в программном обеспечении.

Ключевые слова: методы распознавания, вычислительное моделирование, инжиниринг, сверточные нейронные сети, задачи классификации и локализации.

A.L. Zolkin, V.S. Tormozov, Yu.V. Gumennikova, I.V. Stepina

THESES TO THE QUESTION OF THE SOFTWARE METHODOLOGY
FOR THE APPLICATION OF THE TECHNOLOGY OF THE RECOGNITION
NEURAL NETWORK DEVICE

Abstract. The article deals with the general problem of the problem of localization and classification of objects in raster images. The task of localization as a technology of iterative recognition of road signs as a typical object for demonstrating recognition methodology. At the moment, there are no universal methods and technologies that can ideally solve the problem of localizing and classifying objects in an image. The most popular and relevant methods for localization and classification of objects in the image were analyzed. An algorithm for recognition technology has been developed and presented. The results of the program work are demonstrated. The target map of using the algorithmic method is generalized. To solve the problems posed in this study, it was decided to use HOG & SVM descriptors in the localization module, and transfer the localized image to the CNN that implements the classification module. An algorithm for developing an own model for designing recognition technology based on the methodology for using a neural network apparatus in software is considered and justified.

Keywords: recognition methods, computational modeling, engineering, convolutional neural networks, classification and localization problems.

Введение

Первым этапом в задаче распознавания дорожных знаков как общебазовой задачи является локализация их расположения на изображении. Для данной задачи были рассмотрены наиболее известные методы и сделан выбор в пользу классификаторов HOG&SVM, основанных на гистограммах направленных градиентов с использованием метода опор-

Золкин Александр Леонидович

доцент кафедры информатики и вычислительной техники. Поволжский государственный университет телекоммуникаций и информатики, город Самара. Сфера научных интересов: автоматика и автоматизация; информатика и вычислительная техника; прикладная информатика; программирование; транспорт. Автор 300 опубликованных научных работ.

Электронный адрес: alzolkin@list.ru

Тормозов Владимир Сергеевич

старший преподаватель кафедры программного обеспечения вычислительной техники и автоматизированных систем. Тихоокеанский государственный университет, город Хабаровск. Сфера научных интересов: искусственный интеллект; распознавание образов; цифровая обработка изображений. Автор более 20 опубликованных научных работ.

Электронный адрес: 007465@pnu.edu.ru

Гуменникова Юлия Валериевна

кандидат физико-математических наук, доцент, доцент кафедры высшей математики. Самарский государственный университет путей сообщения, город Самара. Сфера научных интересов: интернет вещей; прикладная информатика; программирование. Автор более 50 опубликованных научных работ.

Электронный адрес: gumennikov@yandex.ru

Степина Ирина Васильевна

кандидат технических наук, доцент, доцент кафедры строительного материаловедения. Национальный исследовательский университет Московский государственный строительный университет, Москва. Сфера научных интересов: автоматика и автоматизация; химия привитых поверхностных соединений на основе элементоорганических веществ и растительного сырья; управление составом, структурой и свойствами гидроксिलированных подложек путем химического модифицирования; биостойкие композитные материалы на основе модифицированного растительного сырья. Автор более 80 опубликованных научных работ.

Электронный адрес: sudeykina@mail.ru

ных векторов, так как такой подход позволяет получать большую точность в сравнении с методом Виолы – Джонса и сравнительно меньшие затраты ресурсов и времени относительно нейросетевых подходов.

Вторым этапом в задаче распознавания дорожных знаков является классификация локализованных объектов.

Методы исследования

Для данной задачи рассмотрены наиболее известные методы и сделан выбор в пользу сверточной нейронной сети, так как такой подход позволит получить большую точность в сравнении с другими подходами, основанными на машинном обучении или цветовых особенностях изображений, и сравнительно меньшие затраты ресурсов и времени относительно обычного многослойного перцептрона [10].

Архитектура сверточной нейронной сети

Сверточные нейронные сети (далее – CNN) – это глубокие искусственные нейронные сети, которые используются главным образом для классификации изображений.

CNN также могут кластеризовать изображения по сходству и выполнять распознавание (локализацию) объектов в сценах [8].

CNN могут идентифицировать лица, уличные знаки, опухоли и многие другие аспекты визуальных данных в зависимости от обучающей выборки, которая была задана. Можно выделить следующие преимущества сверточной нейронной сети: меньшее количество параметров в сравнении с полносвязанной сетью; устойчивость к перемене расположения и поворотам распознаваемого объекта. Однако CNN обладают и рядом недостатков: долгое время обучения; необходимость большой обучающей выборки [9].

Несмотря на недостатки, на сегодняшний день архитектура сверточной нейронной сети является наиболее пригодной для решения задачи классификации изображений.

Сверточная нейронная сеть в общем виде состоит из следующих слоев и их чередования: слои свертки; пулинговые слои (слои подвыборки); полносвязные слои. Сверточные нейронные сети используют различные функции активации.

Пример возможной структуры сверточной нейронной сети представлен на Рисунке 1.

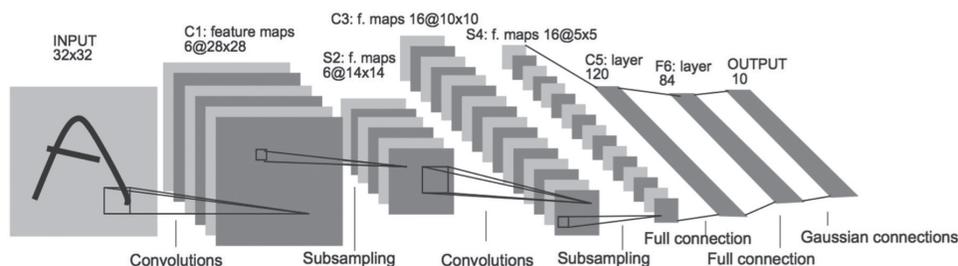


Рисунок 1. Пример структуры сверточной нейронной сети

Результаты

Для создания классифицирующей модели с использованием сверточной нейронной сети можно выделить три основных этапа:

1. Формирование и подготовка обучающей базы данных.
2. Проектирование нейронной сети.
3. Обучение нейронной сети.

Рассмотрим реализацию каждого из этих этапов.

Подготовка обучающей базы данных для разрабатываемой технологии. В обучающем наборе GTSRB изображения не обязательно имеют фиксированные размеры, и знак не обязательно центрируется на каждом изображении. Каждое изображение содержит около 10 % границы вокруг фактического дорожного знака. Как было описано в предыдущих разделах, изображения в обучающем наборе отличаются друг от друга в плане освещенности, контрастности и положении знака, поэтому их необходимо привести к некоторому общему виду. Для решения данной проблемы был использован разработанный модуль `preprocessing.py`, решающий данную проблему. Все изображения были масштабированы до триединого размера 48×48 пикселей [9].

Проектирование и обучение нейронной сети. Для решения задачи классификации дорожных знаков было принято решение воспользоваться библиотечным решением Keras, которое позволяет проектировать нейронные сети различных структур без необходимости их реализации с нуля, что значительно ускоряет разработку. Однако выбор и

проектирование правильной структуры сети – важный вопрос, который в значительной мере влияет на конечный результат.

Keras – открытая библиотека для работы с нейросетевыми алгоритмами, написанная на языке программирования Python. Она является надстройкой над другими открытыми библиотеками и фреймворками: DeepLearning4j, TensorFlow и Theano [2].

В рамках решения описанной ранее задачи была разработана и использована следующая модель сверточной нейронной сети: шесть сверточных слоев, три слоя подвыборки, два полносвязных слоя, а также дополнительные нормализующие слои [4].

Для более подробного представления о структуре разработанной модели приведем часть кода, отвечающего за формирование данной модели (см. Рисунок 2).

```
def cnn_model():
    model = Sequential()

    model.add(Conv2D(32, (3, 3), padding='same',
                    input_shape=(3, IMG_SIZE, IMG_SIZE),
                    activation='relu'))
    model.add(Conv2D(32, (3, 3), activation='relu'))
    model.add(MaxPooling2D(pool_size=(2, 2)))
    model.add(Dropout(0.2))

    model.add(Conv2D(64, (3, 3), padding='same',
                    activation='relu'))
    model.add(Conv2D(64, (3, 3), activation='relu'))
    model.add(MaxPooling2D(pool_size=(2, 2)))
    model.add(Dropout(0.2))

    model.add(Conv2D(128, (3, 3), padding='same',
                    activation='relu'))
    model.add(Conv2D(128, (3, 3), activation='relu'))
    model.add(MaxPooling2D(pool_size=(2, 2)))
    model.add(Dropout(0.2))

    model.add(Flatten())
    model.add(Dense(512, activation='relu'))
    model.add(Dropout(0.5))
    model.add(Dense(NUM_CLASSES, activation='softmax'))
    return model
```

Рисунок 2. Часть листинга программы проектирования модели сети

Разработка и проектирование технологии алгоритма комплексного распознавания дорожных знаков. После разработки отдельных модулей можно переходить к проектированию и разработке комплексного алгоритма локализации и классификации объектов на изображении. Для реализации алгоритма комплексного распознавания на примере объектов высокой контрастности (дорожных знаков) использовался язык программирования Python и сторонние библиотеки с открытым исходным кодом: OpenCV, Dlib, NumPy, Keras.

В общем виде предлагаемый алгоритм можно описать следующим образом:

- исходное изображение подается на вход HOG&SVM дескрипторов, с помощью которых локализуются потенциальные искомые объекты;
- потенциальные локализованные области исходного изображения подаются на вход модуля подготовки изображений;
- изображения обрабатываются (выравнивается яркость, контрастность, приводятся к единому размеру) и подаются на вход CNN;

Тезисы к вопросу методологии программного обеспечения применения технологии ...

- обученная модель возвращает идентификатор для каждой потенциальной области;
- локализованные и классифицированные знаки обводятся рамкой на исходном изображении и подписываются в соответствии с идентификатором, который выдала CNN.

Принцип работы предлагаемого алгоритма показан в виде блок-схемы на Рисунке 3.

Целью данного исследования является разработка универсального алгоритма локализации и классификации объектов на изображении, поэтому блок-схема, представленная на Рисунке 3, построена без привязки к конкретным моделям и дескрипторам, обученным ранее.

Разработанный алгоритм достаточно универсален и может применяться для других практико-ориентированных задач [5].

Алгоритм, представленный на Рисунке 3, был разработан и применен для задачи распознавания дорожных знаков.

Пример успешной работы разработанного решения представлен на Рисунке 4.

Тестирование классифицирующей сверточной нейронной сети. После локализации потенциальный знак подается на вход заранее обученной сети. Точность классификации очень важна для корректной работы всей системы. Если локализованный объект будет распознан неправильно, то разработанная система не только не приносит пользы, но и может создать проблемы, дезинформируя пользователя [6; 7].

Для тестирования работы нейронной сети необходимо подготовить большую выборку разнообразных изображений с пометками о наличии на них объекта, относящегося к тому или иному классу. Данная задача достаточно трудоемкая, так как для корректных результатов необходимо несколько тысяч таких изображений. В пакете данных GTSRB [5] присутствует набор тестовых промаркированных данных, состоящий из 12630 изображений для тестирования, который будет использоваться при тестировании разработанной модели.

Заключение и выводы

На основании разработанного алгоритма и его программной реализации можно сделать следующие выводы.

1. Представленный алгоритм достаточно универсален и может быть применен к любой подобной задаче.

2. Разработанное решение выдает как верные, так и неверные результаты.

Было проведено численное тестирование разработанного решения, результатами которого стали следующие показатели.

1. Модуль локализации детектирует дорожный знак с вероятностью 77,65 %.

2. Модуль классификации верно классифицирует дорожный знак с вероятностью 97 %.

Полученные результаты вполне удовлетворительны и применимы на практике, однако, возможно, в ближайшем будущем появятся еще более точные нейросетевые подходы для локализации объектов.

На основании сделанного выбора разработан универсальный подход и алгоритм для решения задачи локализации и классификации объектов на изображении, который впоследствии может быть применен для других практико-ориентированных задач локализации и распознавания образов.

Разработанное техническое решение может быть расширено и использовано как система помощи водителю (автопилот). Предложенный алгоритм может применяться для совершенно иной задачи со схожей проблематикой [1; 3; 11].

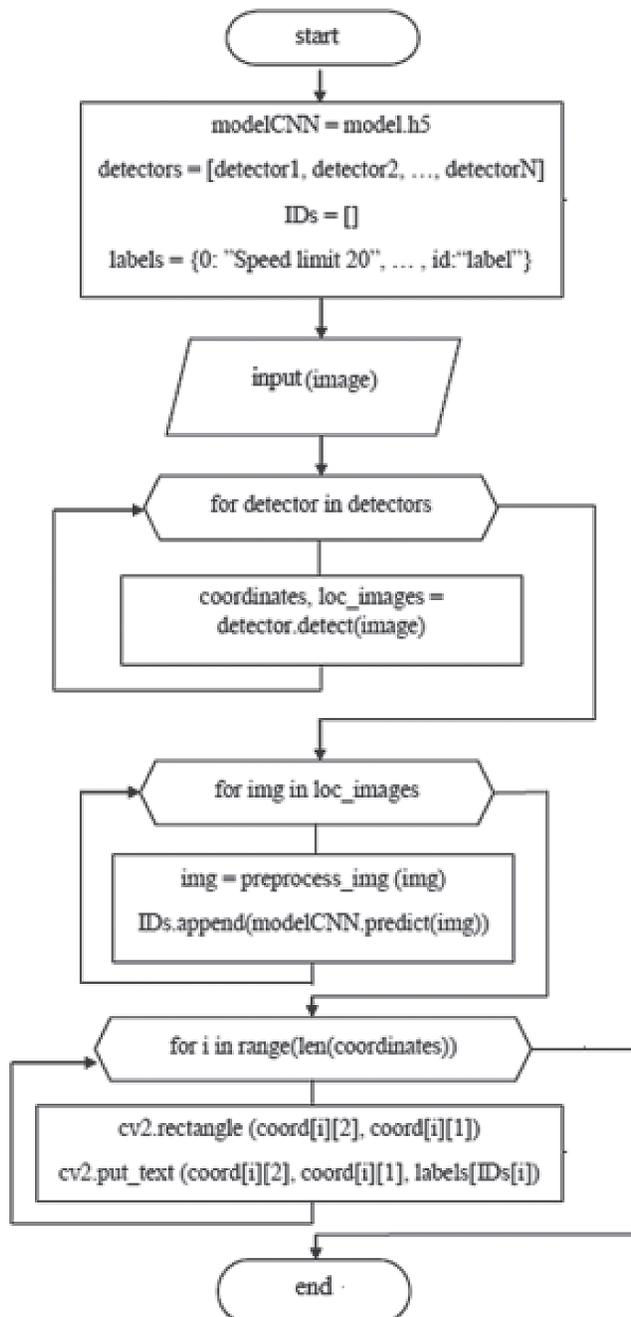


Рисунок 3. Блок-схема алгоритма локализации и классификации объектов на изображении

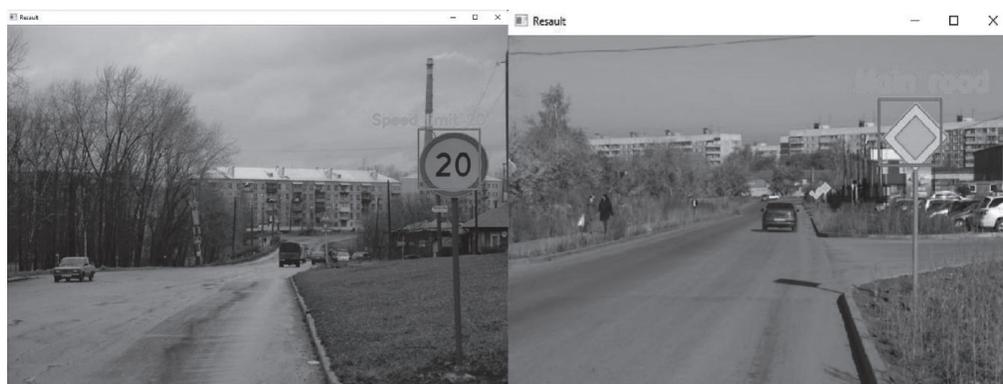


Рисунок 4. Пример успешной работы разработанного решения

Литература

1. Чирков М., Лачинина Т., Чистяков М. Знания и информация как синергия платформенного подхода цифровизации глобального развития // Свободная мысль. 2020. № 5 (1683). С. 37–44. DOI: 10.24411/0869-4435-2020-00003.
2. Ciresan D. C., Meier U., Masci J., Schmidhuber J. (2011) A committee of neural networks for traffic sign classification. In International Joint Conference on Neural Networks, 4 p.
3. Delfani F., Samanipour H., Beiki H., Yumashev A., Akhmetshin E. (2020) A robust fuzzy optimisation for a multi-objective pharmaceutical supply chain network design problem considering reliability and delivery time. International Journal of Systems Science: Operations and Logistics, pp. 1–25. DOI: 10.1080/23302674.2020.1862936.
4. Habibi Aghdam H, Jehani Heravi. E. (2017) Guide to Convolutional Neural Networks a Practical Application to Traffic-Sign Detection and Classification, 282 p.
5. Jain V. and Seung S. H. (2008) Natural image denoising with convolutional networks, 8 p.
6. Liu C.-L., Yin F., Wang Q.-F., Wang D.-H. (2011) ICDAR 2011 chinese handwriting recognition competition. International Conference on Document Analysis and Recognition, pp. 1464–1469.
7. Mark Summerfield (2009) Programming in Python 3: a complete introduction to the Python language, 636 p.
8. Melin P. (2015) Design of Intelligent Systems Based on Fuzzy Logic, Neural Networks and Nature-Inspired Optimization. Patricia Melin, Oscar Castillo, JanuszKacprzyk. Springer International Publishing, 637 p.
9. Ronan C., Jason W. (2008) A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning. Proceedings of the 25th International Conference on Machine Learning, ICML '08, New York, NY, USA, ACM, pp. 160–167.
10. Stallkamp J., Schlipsinga M., Salmena J., Igelb Man C. vs. (2012) Computer: Benchmarking Machine Learning Algorithms for Traffic Sign Recognition, 11 p.
11. Tormozov V.S., Zolkin A.L., Vasilenko K.A. (2020) Optimization of neural network parameters based on a genetic algorithm for prediction of time series. International Multi-Conference on Industrial Engineering and Modern Technologies, FarEastCon, p. 9271536. DOI: 10.1109/FarEastCon50210.2020.9271536.

References

1. Chirkov M., Lachinina T., Chistyakov M. (2020) *Znanija i informacija kak sinergija platformennogo podhoda cifrovizacii global'nogo razvitija* [Knowledge and information as a synergy of the platform approach to digitalization of global development]. *Svobodnaja mysl'*, No. 5 (1683), pp. 37–44 (in Russian). DOI: 10.24411/0869-4435-2020-00003.
2. Cirean D., Meier U., Masci J., Schmidhuber J. (2011) A committee of neural networks for traffic sign classification. In *International Joint Conference on Neural Networks*, 4 p.
3. Delfani F., Samanipour H., Beiki H., Yumashev A., Akhmetshin E. (2020) A robust fuzzy optimisation for a multi-objective pharmaceutical supply chain network design problem considering reliability and delivery time. *International Journal of Systems Science: Operations and Logistics*, pp. 1–25. DOI: 10.1080/23302674.2020.1862936.
4. Habibi Aghdam H., Jehani Heravi E. (2017) *Guide to Convolutional Neural Networks a Practical Application to Traffic-Sign Detection and Classification*, 282 p.
5. Jain V. and Seung S. (2008) Natural image denoising with convolutional networks, 8 p.
6. Liu C.-L., Yin F., Wang Q.-F., Wang D.-H. (2011) ICDAR 2011 chinese handwriting recognition competition. *International Conference on Document Analysis and Recognition*, pp. 1464–1469.
7. Mark Summerfield (2009) *Programming in Python 3: a complete introduction to the Python language*, 636 p.
8. Melin P. (2015) *Design of Intelligent Systems Based on Fuzzy Logic, Neural Networks and Nature-Inspired Optimization*. Patricia Melin, Oscar Castillo, JanuszKacprzyk. Springer International Publishing, 637 p.
9. Ronan C., Jason W. (2008) A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning. *Proceedings of the 25th International Conference on Machine Learning, ICML '08, New York, NY, USA, ACM*, pp. 160–167.
10. Stallkampa J., Schlipfing M., Salmena J., Igelb Man C. vs. (2012) *Computer: Benchmarking Machine Learning Algorithms for Traffic Sign Recognition*, 11 p.
11. Tormozov V.S., Zolkin A.L., Vasilenko K.A. (2020) Optimization of neural network parameters based on a genetic algorithm for prediction of time series. *International Multi-Conference on Industrial Engineering and Modern Technologies, FarEastCon*, p. 9271536. DOI: 10.1109/FarEast-Con50210.2020.9271536.