

Эксперименты с обучением методов распознавания дорожных знаков на синтетических данных

08, август 2013

DOI: 10.7463/0813.0603378

Чигорин А. А., Конушин А. С.

УДК 004.93

Россия, Московский государственный университет имени М.В. Ломоносова

aachigorin@graphics.cs.msu.ru

ktosh@graphics.cs.msu.ru

Введение.

Задача распознавания дорожных знаков активно исследуется уже не одно десятилетие. Если будет создан метод, работающий с достаточной точностью, его можно использовать в системах помощи водителю, инвентаризации объектов придорожной инфраструктуры или для автоматизированного создания навигационных карт.

Дорожные знаки сделаны, чтобы быть заметными и легко читаемыми и это делает их хорошим объектом для автоматического распознавания. В то же время автоматические алгоритмы должны распознавать знаки в условиях внутриклассовой изменчивости, изменения освещенности, положения точки наблюдения, размытия и перекрытий. Лучшие на сегодняшний день алгоритмы опираются на машинное обучение и требуют наличия большой и репрезентативной обучающей коллекции, чтобы обойти все вышеперечисленные проблемы. Для получения такой коллекции необходимо просмотреть и разметить тысячи километров проезда. Задача становится особенно сложной в случае редко встречающихся классов знаков. Учитывая тот факт, что знаки в разных странах отличаются друг от друга – необходимо собирать подобную коллекцию для каждой новой страны. Чтобы обойти вышеописанную проблему разметки мы предлагаем использовать для обучения синтетически созданную обучающую выборку. Также мы предлагаем алгоритм, предполагающий, что известны параметры трансформаций, которым были подвержены изображения из обучающей выборки. Данное предположение является верным в случае синтетической обучающей коллекции. Наш алгоритм способен не только распознавать класс знака, но и сегментировать его от фона.

1. Обзор литературы.

Обобщенный алгоритм обнаружения и распознавания объектов состоит из трех этапов. На первом этапе ограничивающий прямоугольник объекта выделяется на изображении. Далее происходит уточнение положения объекта внутри полученного ограничивающего прямоугольника, и объект сегментируется от фона. На третьем этапе происходит распознавание класса знака. В данной работе мы рассматриваем задачи, решаемые на первом и втором этапе алгоритма.

Согласно классификации, предложенной в [1], методы распознавания автодорожных знаков можно разделить на две группы – основанные на похожести знаков и основанные на признаках. В методах, основанных на похожести, распознаваемому изображению присваивается метка класса ближайшего к нему прототипа из обучающей выборки. В [2] предложен новый вариант AdaBoost названный SimBoost, спроектированный для распознавания двух классов – “похожие” и “разные”. Класс входного изображения определяется с помощью поиска ближайшего соседа с инвертированной уверенностью SimBoost в качестве метрики расстояния. В [3] предложен гибридный метод на основе классификаторов ближайшего соседа и случайного леса на HOG дескрипторах [4]. Описанный в [3] подход на основе ближайшего соседа похож на предложенный нами, но мы добиваемся лучших результатов за счет использования двухэтапной схемы и синтетических данных.

В подходе, основанном на признаках, входное изображение описывается с помощью какого-либо дескриптора и классификатор, предсказывающий класс знака, обучается на извлеченных признаках. В [5] предложена двухэтапная схема классификации. На первом этапе знаку присваивается метка одной из 11 групп, на основе признаков цвета и формы. Затем знак отделяется (сегментируется) от фона с помощью маски знака, заранее определенной для каждой группы. RGB значения пикселей, лежащих внутри маски знака, подаются на вход SVM-классификатора с гауссовым ядром. В предложенном нами методе мы можем точнее определить маску знака за счет использования большего количества вариаций. В [6] результаты работы многослойной нейронной сети (MLP) и свёрточной нейронной сети (CNN) объединяются в комитет. MLP обучается на HOG дескрипторах, в то время как CNN – на случайным образом отмасштабированных, сдвинутых и повернутых цветных изображениях. Авторы [7] также использовали CNN с обучением на изображениях, преобразованных в цветовое пространство YUV, с Y каналом, обработанным с помощью локального и глобального выравнивания контраста.

2. Создание синтетической обучающей выборки.

Для создания синтетических изображений знаков нами был использован алгоритм, описанный в статье [8]. Визуализация результатов работы основных шагов алгоритма

приведена на рисунке 1. На вход алгоритма генерации подается пиктограмма знака и маска фона, которую легко выделить по пиктограмме. Далее к пиктограмме знака применяется набор трансформаций, состоящий из вариации интенсивности и насыщенности, поворотов вокруг трех осей, перспективных преобразований, добавления шумов и смещений. На последнем этапе к полученному изображению знака с помощью алгоритма смешивания добавляется фон, вырезанный из реальных изображений. В результате получается выборка, состоящая из синтетических изображений знаков и масок фона. Для каждого примера из выборки известно преобразование, которому он был подвержен. Визуализация результатов работы алгоритма приведена на рисунке 2.

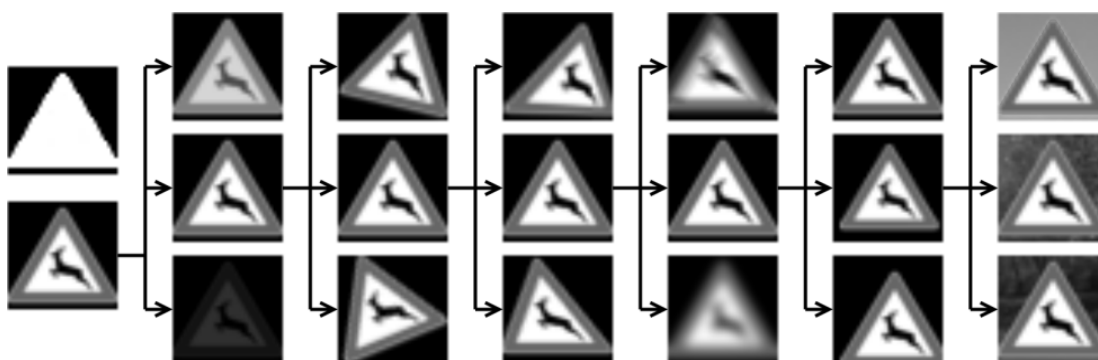


Рисунок 1. Алгоритм генерации синтетических данных.



Рисунок 2. Пример получаемых синтетических данных.

3. Описание предложенного алгоритма.

После получения синтетической обучающей выборки мы извлекаем HOG дескрипторы из всех примеров и используем классификатор на основе поиска ближайшего соседа для присвоения входному изображению метки класса знака.

Метод на основе поиска ближайшего соседа обладает простым свойством – чем больше данных в обучающей выборке, тем лучше результаты распознавания. Наш

алгоритм генерации синтетических данных содержит 16 параметров трансформаций. Если каждый параметр при генерации будет представлен хотя бы тремя значениями, то мы получим больше 50 миллионов синтетических примеров только на один класс. Чтобы решить данную проблему мы разбиваем поиск ближайшего соседа на два этапа. На первом этапе мы определяем приближенные параметры трансформации за счет поиска по выборке, сгенерированной с параметрами, варьируемыми в широком диапазоне значений с большим шагом. На втором этапе мы ищем ближайшего соседа в окрестности трансформации, определенной на первом этапе.

В случае дорожных знаков на первом этапе мы определяем приближенные параметры трансформации и маску знака. Далее мы отделяем знак от фона и удаляем смещения по краям, используя найденную маску. И производим поиск ближайшего соседа в окрестности найденной трансформации, чтобы получить окончательную метку класса. Визуализация работы алгоритма приведена на рисунке 3. Примеры четырех ближайших соседей, получаемых после первого этапа алгоритма, приведены на рисунке 4. Видно, что удается точно определить параметры трансформации, что позволяет отделить знак от фона.

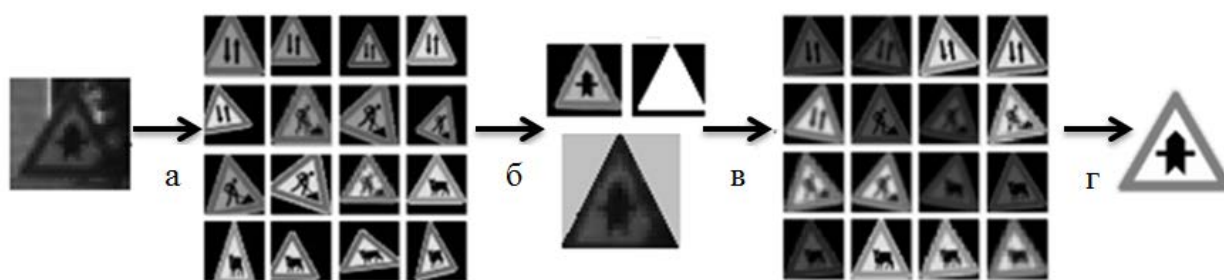


Рисунок 3. Визуализация работы предложенного алгоритма распознавания: а) Поиск входного изображения среди трансформаций в широком диапазоне б) найдена приближенная трансформация и маска знака. Знак отделен от фона и выровнен в) поиск отделенного от фона знака в окрестности найденной трансформации г) присвоение класса знака.

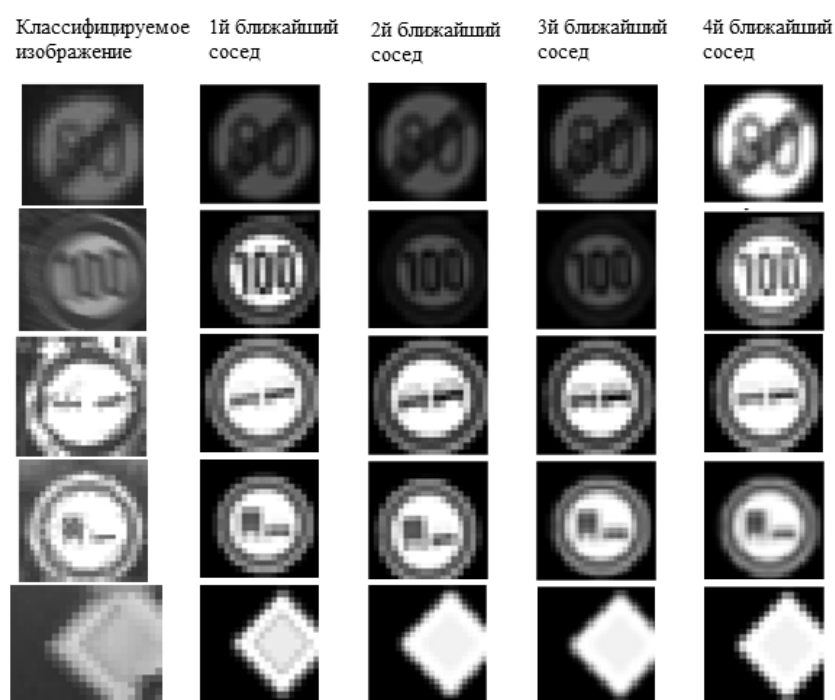


Рисунок 4. Примеры четырех ближайших соседей, получаемых после первого этапа предложенного алгоритма.

4. Эксперименты и результаты.

Чтобы подтвердить эффективность классификатора на основе ближайшего соседа мы обучили на нашей синтетической выборке, состоящей из 100 000 изображений знаков, следующие типы классификаторов:

- 1) Linear Discriminant Analysis (LDA)
- 2) Linear SVM (LSVM)
- 3) Классификатор на основе К ближайших соседей (k-NN)

Мы провели сравнение получаемых точностей классификации на базе GTSTB [9]. Она состоит из 50000 изображений 43 классов знаков. Размеры встречаемых в коллекции знаков варьируются от 15x15 до 222x199 пикселей. Обучающая и тестовая выборка состоят из 39209 и 12630 изображений, соответственно. Мы тестировали предложенный алгоритм на тестовой части выборки и обучались на синтетических данных. Параметры дескриптора HOG были подобраны путем сравнения нескольких наборов параметров в связке с k-NN классификатором. Результаты сравнения приведены в таблице 1. Для дальнейших экспериментов был использован набор параметров “HOG 3”.

Таблица 1. Подбор параметров дескриптора HOG.

Name	Cell	Block	Stride	Num. of bins	Semicircle	Acc. [%]
HOG 1	5x5	10x10	5x5	8	true	92.35
HOG 2	4x4	8x8	2x2	8	true	94.06
HOG 3	3x3	6x6	3x3	9	true	95.03

Все гиперпараметры классификаторов были подобраны с помощью метода скользящего контроля. Результаты классификации приведены в таблице 2. Тренировочные наборы № 1 и № 2 содержат синтетические и реальные данные соответственно. Из сравнения результатов для наборов № 1 и № 2 видно, что в случае обучения на синтетических данных LDA и LSVM работают намного хуже, чем в случае обучения на реальных данных. Мы объясняем эту разницу тем, что распределение, лежащее в основе генератора синтетических данных, намного сложнее и линейные классификаторы, такие как LDA и LSVM, не могут разделить получаемую выборку. В то же время нелинейный метод на основе поиска ближайшего соседа (k-NN) лучше справляется со сложной выборкой. Чтобы подтвердить наши выводы мы провели простой эксперимент, создав набор № 3, содержащий синтетические данные после этапа сегментации и выравнивания предложенного нами алгоритма. Результирующая выборка получилась проще, за счет отсутствия в генераторе синтетических данных ряда трансформаций. Видно, что ошибка классификации у LDA и LSVM упала в два раза. Лучшие на сегодняшний день методы классификации знаков на данной базе достигают точности 99.46% и основаны на объединении результатов работы нескольких глубоких нейронных сетей. Глубокими нейронными сетями называют такие нейронные сети, у которых количество слоев с настраиваемыми во время обучения весами больше двух. Хотя точность предложенного нами метода ниже, он лучше подходит для использования модуля распознавания в связке с модулем обнаружения, так как позволяет выровнять знак и отделить его от фона. Типичные ошибки предложенного алгоритма приведены на рисунке 5. Они вызваны отсутствием соответствующей трансформации в алгоритме генерации синтетических данных. К таким трансформациям относятся: перекрытия, блики, сильное размытие за счет движения камеры.

Таблица 2. Сравнение различных классификаторов при обучении на реальных и на синтетических данных.

Номер тренировочного набора	Тип данных	LDA [%]	Linear SVM [%]	k-NN [%]
1	Синтетические	43.6	79.01	93.15
2	Реальные	93.28	95.7	72.81
3	Синтетические после сегментации	83.22	91	96.91 (предложенный метод)



Рисунок 5. Типичные ошибки предложенного алгоритма.

Мы также протестировали предложенный нами метод на двух других базах знаков. На базе шведских знаков [10] на “visible” подмножестве изображений наш метод показал точность 97.47% на первой части и 99.13% на второй части. Сравнение с результатами из статьи создателей базы [11] провести невозможно, так как в статье заявлены результаты для одновременного обнаружения и распознавания знаков. На базе бельгийских знаков [12] нами достигнута точность в 95.57%. В [13] была получена точность в 97%, но распознаванию подвергались только обнаруженные, а значит самые простые, знаки.

5. Технические детали.

Модель предложенного двухуровневого классификатора занимает 10 Гб в оперативной памяти и состоит из 650000 синтетических изображений знаков с масками фона. Время классификации равно 0.3 секунды на одном современном процессоре. Процедура генерации синтетической выборки указанного размера занимает 15 минут и может быть существенно ускорена, так как большую часть времени занимают операции дискового ввода-вывода.

Заключение.

Предложен метод распознавания знаков дорожного движения, спроектированный для обучения на синтетически созданных данных. В методе предполагается, что для каждого обучающего примера известны преобразования, которым он был подвергнут. Описан алгоритм создания синтетической обучающей выборки. Метод распознавания работает на основе поиска ближайшего соседа в пространстве дескрипторов, полученных за счет извлечения градиентных признаков. Предложенный метод также позволяет отделить знак от фона и выровнять его границы, что в дальнейшем помогает на этапе классификации. Проведено тестирование метода на общедоступных базах знаков из трех стран. Получены результаты, сравнимые по точности с описанными в других статьях.

Работа поддержана РФФИ; грант №12-01-33085 по теме "Методы настройки вероятностных моделей в задачах анализа изображений, данных лазерного сканирования, компьютерного зрения".

Список литературы

1. Dalal N., Triggs W. Histogram of oriented gradients for human detection // Proc. IEEE Comput. Soc. Conf. Comput. Vis. and Pattern Recog. 2005. P. 886-893. DOI: [10.1109/CVPR.2005.177](https://doi.org/10.1109/CVPR.2005.177)
2. Paclik P., Novovicova J., Duin R. Building Road-Sign Classifiers Using a Trainable Similarity Measure // IEEE Trans. Intell. Transp. Syst. 2006. Vol. 7, no. 3. P. 309-321. DOI: [10.1109/TITS.2006.880627](https://doi.org/10.1109/TITS.2006.880627)
3. Ruta A., Porikli F., Watanabe S., Li Y. In-vehicle Camera Traffic Sign Detection and Recognition // Machine Vision and Applications. 2011. Vol. 22, no. 2. P. 359-375.
4. Zaklouta F., Stanculescu B., Hamdoun O. Traffic sign classification using K-d trees and Random Forests // IEEE International Joint Conference on Neural Networks. 2011. P. 2151-2155. DOI: [10.1109/IJCNN.2011.6033494](https://doi.org/10.1109/IJCNN.2011.6033494)
5. Maldonado-Bascon S., Lafuente-Arroyo S., Gil-Jimenez P., Gomez-Moreno H., Lopez-Ferreras F. Road-Sign Detection and Recognition Based on Support Vector // IEEE Trans. Intell. Transp. Syst. 2007. Vol. 8, № 2. P. 264-278. DOI: [10.1109/TITS.2007.895311](https://doi.org/10.1109/TITS.2007.895311)
6. Ciresan D., Meier U., Masci J., Schmidhuber J. A Committee of Neural Networks for Traffic Sign Classification // IEEE International Joint Conference on Neural Networks. 2011. P. 1918-1921. DOI: [10.1109/IJCNN.2011.6033458](https://doi.org/10.1109/IJCNN.2011.6033458)
7. Sermanet P., Lecun Y. Traffic Sign Recognition with Multi-Scale Convolutional Networks // The 2011 International Joint Conference on Neural Networks. IEEE Conference Publications, 2011. P. 2809-2813. DOI: [10.1109/IJCNN.2011.6033589](https://doi.org/10.1109/IJCNN.2011.6033589)
8. Моисеев Б., Чигорин А. Классификация автодорожных знаков с помощью свёрточной нейросети, обученной на синтетических данных // 22-я Международная Конференция по Компьютерной Графике и Машинному Зрению (GraphiCon'2012) : труды. (Москва, 1-5 Октября 2012 г.). Москва, МГУ, 2012. P. 284-287.
9. Stallkamp J., Schlipsing M., Salmen J., Igel C. Man vs. Computer: Benchmarking Machine Learning Algorithms for Traffic Sign Recognition // Neural Networks. 2012. Vol. 32. P. 323-332. <http://dx.doi.org/10.1016/j.neunet.2012.02.016>
10. Larsson F. Traffic signs dataset. Available at: <http://www.cvl.isy.liu.se/research/traffic-signs-dataset> , accessed 01.07.2013.
11. Larsson F., Felsberg M. Using Fourier Descriptors and Spatial Models for Traffic Sign Recognition // Image Analysis : Proceedings of the 17th Scandinavian Conference on Image Analysis, SCIA 2011 (Ystad, Sweden, May 2011). Springer Berlin Heidelberg, 2011. P. 238-249. DOI: 10.1007/978-3-642-21227-7_23 (Ser. Lecture Notes in Computer Science; vol. 6688).
12. Timofte R. KUL Belgium Traffic Sign Classification Benchmark. Available at: <http://homes.esat.kuleuven.be/~rtimofte/> , accessed 01.07.2013
13. Timofte R., Zimmermann K., Gool Van L. Multi-view traffic sign detection, recognition, and 3D localization // IEEE Workshop on Applications of Computer Vision (WACV 2009). December 2009. P. 1-8. DOI: [10.1109/WACV.2009.5403121](https://doi.org/10.1109/WACV.2009.5403121).

Experiments with teaching of recognition methods for road signs with the use of synthetic data

08, August 2013

DOI: 10.7463/0813.0603378

Chigorin A.A., Konushin A.S.

Lomonosov Moscow State University, Russian Federation

aachigorin@graphics.cs.msu.ruktosh@graphics.cs.msu.ru

This article considers the task of traffic sign recognition with learning from synthetic data. The method for synthetic dataset generation from one traffic signs pictogram for each class was described. The authors proposed a method for traffic signs segmentation and classification; the method is based on the nearest neighbor search in the space of gradient features derived from synthetic data. The proposed method was evaluated on databases from three different countries – Germany, Belgium and Sweden. A comparative analysis of classification accuracy of the proposed method with training from synthetic and real data was carried out.

Publications with keywords: [pattern recognition](#), [classification algorithm](#), [synthetic data](#), [traffic signs](#)

Publications with words: [pattern recognition](#), [classification algorithm](#), [synthetic data](#), [traffic signs](#)

References

1. Dalal N., Triggs W. Histogram of oriented gradients for human detection. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. and Pattern Recog*, 2005, pp. 886-893. DOI: [10.1109/CVPR.2005.177](https://doi.org/10.1109/CVPR.2005.177)
2. Paclik P., Novovicova J., Duin R. Building Road-Sign Classifiers Using a Trainable Similarity Measure. *IEEE Trans. Intell. Transp. Syst.*, 2006, vol. 7, no. 3. pp. 309-321. DOI: [10.1109/TITS.2006.880627](https://doi.org/10.1109/TITS.2006.880627)
3. Ruta A., Porikli F., Watanabe S., Li Y. In-vehicle Camera Traffic Sign Detection and Recognition. *Machine Vision and Applications*, 2011, Vol. 22, no. 2, pp. 359-375.
4. Zaklouta F., Stanculescu B., Hamdoun O. Traffic sign classification using K-d trees and Random Forests. *IEEE International Joint Conference on Neural Networks*, 2011, pp. 2151-2155. DOI: [10.1109/IJCNN.2011.6033494](https://doi.org/10.1109/IJCNN.2011.6033494).

5. Maldonado-Bascon S., Lafuente-Arroyo S., Gil-Jimenez P., Gomez-Moreno H., Lopez-Ferreras F. Road-Sign Detection and Recognition Based on Support Vector. *IEEE Trans. Intell. Transp. Syst.*, 2007, vol. 8, no. 2, pp. 264-278. DOI: [10.1109/TITS.2007.895311](https://doi.org/10.1109/TITS.2007.895311).
6. Ciresan D., Meier U., Masci J., Schmidhuber J. A Committee of Neural Networks for Traffic Sign Classification. *IEEE International Joint Conference on Neural Networks*, 2011, pp. 1918-1921. DOI: [10.1109/IJCNN.2011.6033458](https://doi.org/10.1109/IJCNN.2011.6033458)
7. Sermanet P., Lecun Y. Traffic Sign Recognition with Multi-Scale Convolutional Networks. *The 2011 International Joint Conference on Neural Networks*, IEEE Conference Publications, 2011. P. 2809-2813. DOI: [10.1109/IJCNN.2011.6033589](https://doi.org/10.1109/IJCNN.2011.6033589)
8. Moiseev B., Chigorin A. Klassifikatsiya avtodorozhnykh znakov s pomoshch'yu svertochnoy neyroseti, obuchennoy na sinteticheskikh dannyykh [Traffic signs classification with convolutional neural network trained on synthetically generated data]. *Proc. of the 22nd International Conference on Computer Graphics and Vision (GraphiCon'2012)*, Moscow, 1-8 October 2012, Voscow, MSU, 2012, pp. 284-287.
9. Stallkamp J., Schlipsing M., Salmen J., Igel C. Man vs. Computer: Benchmarking Machine Learning Algorithms for Traffic Sign Recognition. *Neural Networks*, 2012, vol. 32, pp. 323-332. <http://dx.doi.org/10.1016/j.neunet.2012.02.016>
10. Larsson F. *Traffic signs dataset*. Available at: <http://www.cvl.isy.liu.se/research/traffic-signs-dataset> , accessed 01/07/2013.
11. Larsson F., Felsberg M. Using Fourier Descriptors and Spatial Models for Traffic Sign Recognition. *Image Analysis : Proceedings of the 17th Scandinavian Conference on Image Analysis, SCIA 2011*, Ystad, Sweden, May 2011. Springer Berlin Heidelberg, 2011, pp. 238-249. DOI: [10.1007/978-3-642-21227-7_23](https://doi.org/10.1007/978-3-642-21227-7_23) (Ser. *Lecture Notes in Computer Science*; vol. 6688).
12. Timofte R. *KUL Belgium Traffic Sign Classification Benchmark*. Available at: <http://homes.esat.kuleuven.be/~rtimofte/> , accessed 01.07.2013.
13. Timofte R., Zimmermann K., Gool Van L. Multi-view traffic sign detection, recognition, and 3D localization. *IEEE Workshop on Applications of Computer Vision (WACV 2009)*, December 2009, pp. 1-8. DOI: [10.1109/WACV.2009.5403121](https://doi.org/10.1109/WACV.2009.5403121).