

УДК 004.81

05.00.00 Технические науки

**ИСПОЛЬЗОВАНИЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ РАСПОЗНАВАНИЯ АВТОМОБИЛЬНЫХ НОМЕРОВ. ПРЕИМУЩЕСТВА И НЕДОСТАТКИ ПО СРАВНЕНИЮ С ШАБЛОННЫМ МЕТОДОМ**

Михалеви́ч Ю́рий Серге́евич  
магистрант  
РИНЦ SPIN-код: 9486-7713  
0@39.yt

Ткаченко Василий Владимирович  
к.э.н., доцент  
РИЦН SPIN-код: 6878-2800  
tkachenkovasso@yandex.ru  
*Кубанский государственный аграрный университет, Россия, 350044, Краснодар, Калинин ул., 13*

Задача распознавания автомобильных номеров является одной из стандартных задач компьютерного зрения. Программное обеспечение, поставляемое вместе с системами видеонаблюдения, в большинстве случаев, предлагает функционал распознавания автомобильных номеров. В тоже время, существует множество подходов к решению данной задачи. Наиболее распространенными являются подходы, основанные на использовании шаблонов. Такие подходы предлагают предсказуемое и небольшое время выполнения и низкий процент ошибок. Однако, данные подходы оказываются намного менее эффективными, когда встает необходимость распознать автомобильный номер, который может находиться в непредсказуемом месте, быть набран неопределенным шрифтом и на нестандартном фоне и не иметь жесткого формата. Таковы, например, номера автомобилей в США. Одним из способов увеличить эффективность и качество распознавания таких номеров, является использование нейронных сетей. Предполагается, что использование нейронных сетей может существенно повысить качество распознавания. Однако, использование нейронных сетей сопряжено со сложностью их тренировки и часто оказывается менее производительным, чем использование шаблонных методов. В данной статье рассматривается возможность использования сверточной нейронной сети, тренированной на базе данных MNIST (Mixed National Institute of Standards and Technology, — Смешанный Национальный Институт Стандартов и Технологий, с англ.). Данная статья носит обзорный характер и рассматривает применение шаблонов и нейронных сетей для распознавания автомобильных номеров с точки зрения качества распознавания, производительности и сложности применения

UDC 004.81

Technical sciences

**USING CONVOLUTIONAL NEURAL NETWORKS FOR LICENSE PLATES RECOGNITION. ADVANTAGES AND DISADVANTAGES IN COMPARISON WITH TEMPLATE-BASED METHOD**

Mikhalevich Yuriy Sergeevich  
master's degree student  
RSCI SPIN-code: 9486-7713  
0@39.yt

Tkachenko Vasily Vladimirovich  
Cand.Econ.Sci., assistant professor  
RSCI SPIN-code: 6878-2800  
tkachenkovasso@yandex.ru  
*Kuban State Agrarian University, Russia, 350044, Krasnodar, Kalinina str., 13*

Car license plates recognition problem is one of the typical tasks of computer vision. Video surveillance software usually provides license plates recognition function. Meanwhile, there are many approaches to solve this problem, where template-based methods are the most common. Such methods providing predictable and short enough execution time, and little percent of mistakes. However, such methods are far less effective in case there is a need to recognize car's license plate, which may be located in unpredictable place, typed in undefined font and on non-standard background, or without strict formatting. For example, USA car license plates. One of the methods to increase effectiveness and quality of such license plates recognition is to use neural networks. It is assumed, that neural networks usage can significantly increase recognition quality. Nevertheless, neural networks usage entails difficulties of it's training, and often becomes less efficient as template-based methods usage. This article discusses probability of usage of convolutional neural network, which was trained using MNIST (Mixed National Institute of Standards and Technology) database. This article is a review of usage of templates and neural networks for car's license plate recognition in terms of quality, performance and complexity of the usage

Ключевые слова: СВЕРТОЧНЫЕ НЕЙРОННЫЕ СЕТИ, НЕЙРОННЫЕ СЕТИ, КОМПЬЮТЕРНОЕ ЗРЕНИЕ, МАШИННОЕ ОБУЧЕНИЕ, АВТОМОБИЛЬНЫЕ НОМЕРА, РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ, ОБЗОР

Keywords: CONVOLUTIONAL NEURAL NETWORKS, NEURAL NETWORKS, COMPUTER VISION, MACHINE LEARNING, LICENSE PLATES, IMAGE RECOGNITION, REVIEW

В настоящее время интеллектуальные системы и решения получают все большее развитие. Одной из сфер применения интеллектуальных систем является видеонаблюдение. Обнаружение движения, дыма, огня, границ, подсчет людей в очереди, наблюдение за кассой, обнаружение людей, распознавание лиц, распознавание автомобильных номеров, – все эти и многие другие задачи решаются с использованием интеллектуальных методов обработки изображений и видео. Однако, до сих пор не существует совершенного и универсального решения многих из перечисленных задач. Применяемые методы постоянно совершенствуются и иногда даже заменяются. В этом контексте представляет интерес задача распознавания автомобильных номеров.

Целью исследования является – проведение сравнительного анализа методов распознавания образов и разработка концепции по использованию сверточных нейронных сетей для задачи распознавания автомобильных номеров.

Знакомясь со статьями на данную тему, можно убедиться (на ряду с авторами этих статей) в том, что оптимальным методом решения этой задачи является применение шаблонов. И, действительно, многие системы распознавания автомобильных номеров используют методы, основанные на применении шаблонов. Суть таких методов заключается в том, что алгоритм располагает информацией о том, как выглядит автомобильный номер, какой у него может быть фон, как выглядят и на каких позициях могут находиться определенные символы (дифференцируются буквы и цифры), учитывает шрифт, используемый для их начертания, сразу рассматривает возможное место обнаружения автомобильного номера, а не все изоб-

ражение. Все это позволяет добиться высокого качества распознавания стандартизированных номерных знаков. Производительность таких методов более чем удовлетворительна. Обыкновенно, уверенность авторов, защищающих методы, основанные на шаблонах, проистекает из того факта, что они рассматривают применение своих методов в рамках одной страны (или территории, на который автомобильные номерные знаки строго стандартизированы). Например, с использованием данного метода можно легко распознавать автомобильные номера, используемые на территории Российской Федерации (Рисунок 1).



Рисунок 1 – Пример номера, для распознавания которого практически применять шаблоны

Однако, какой метод применять в случае, если необходимо распознавать номера на таможенном пункте, где проезжает множество автомобилей с номерами, удовлетворяющими разным шаблонам? Можно использовать множество шаблонов, в ущерб эффективности. Или усложнить один шаблон, сделав его более универсальным, в ущерб качеству. А какой метод использовать, если нам необходимо распознавать автомобильные номера вне формата, находящиеся в неожиданном месте (транзитные номера), содержащие различное количество букв и цифр, размещенные на различном фоне (США) (Рисунок 2)?



Рисунок 2 – Автомобильные номерные знаки США

Возможным ответом на все эти вопросы является использование искусственных нейронных сетей. В машинном обучении и в когнитивистике, под искусственными нейронными сетями понимают семейство моделей, вдохновленных биологическими нейронными сетями (центральной нервной системой животных, в частности – мозгом), и используются для того, чтобы оценить или приблизить функции, которые могут зависеть от огромного количества входных данных и обычно неизвестны. Искусственные нейронные сети обычно представляются как системы соединенных

между собой «нейронов», которые обмениваются сообщениями друг с другом. Соединения имеют множество весов, которые могут быть настроены на основании опыта, делая нейронные сети адаптируемыми к входным данным и способными обучаться.[2][5]

Например, нейронная сеть для распознавания текста определена как множество «входных нейронов», которые могут быть активированы пикселями входного изображения. После того, как они будут взвешены и преобразованы функцией (определенной проектировщиком сети), активность этих нейронов будет передана другим нейронам. Этот процесс будет повторен до тех пор, пока выходной нейрон не будет активирован. Это определит, какой символ был прочитан.[5]

Как и другие методы машинного обучения – системы, обучающиеся из данных – нейронные сети используются для решения множества задач, которые тяжело решить с использованием стандартного логического программирования, включая задачи компьютерного зрения и распознавания речи.[3]

Существует несколько видов искусственных нейронных сетей. Отличают feed-forward (от англ. feed – кормить и forward – вперед) и рекуррентные нейронные сети. В первых – связи между нейронами не могут формировать цикл, во вторых – формируют, с целью предоставления обратной связи более верхним уровням.

Также, нейронные сети классифицируют исходя из их архитектуры – многослойный персептрон (многослойные нейронные сети), глубокие (сложный многослойные нейронные сети, производящие множество нелинейных преобразований), сверточные (содержащие сверточные слои) и другие.[2]

В данной статье мы рассмотрим применение сверточных нейронных сетей. В машинном обучении, сверточные нейронные сети являются видом feed-forward искусственных нейронных сетей, в которой индивидуальные



нейронные уложен таким образом, что они соответствуют перекрывающимся регионам в поле зрения. Сверточные нейронные сети были вдохновлены биологическими процессами и являются вариацией многослойного персептрона, спроектированного для использования минимального объема предобработки. Они находят широкое применение в обработке видео и изображений, советующих системах и обработке естественного языка.[4]

В случае, если такие сети используются для распознавания изображений, они состоят из множества слоев из небольших скоплений нейронов, которые принимают небольшие части входного изображения. Называются эти скопления – рецептивными полями. Результат работы данных скоплений затем размещается так, что они пересекаются для того, чтобы приобрести лучшее представление оригинального изображения; это повторяется для каждого следующего слоя.[6]

Однако, с применением нейронных сетей сопряжены несколько сложностей – сложность обучения сетей и низкая производительность. Для того, чтобы решить вторую проблему, сверточные сети могут содержать локальные или глобальные слои субдискретизации (subsampling, pooling слои), которые комбинируют вывод кластеров нейронов (Рисунок 3), тем самым снижая количество (и детализацию) информации.[1] Если правильно подобрать параметры таких слоев, можно существенно повысить производительность сети, не потеряв, при этом, в качестве.

Для решения задачи распознавания номеров было спроектировано несколько сверточных нейронных сетей с использованием фреймворка для построения нейронных сетей Caffe. Центральные слои одной из них описаны ниже в формате Caffe.

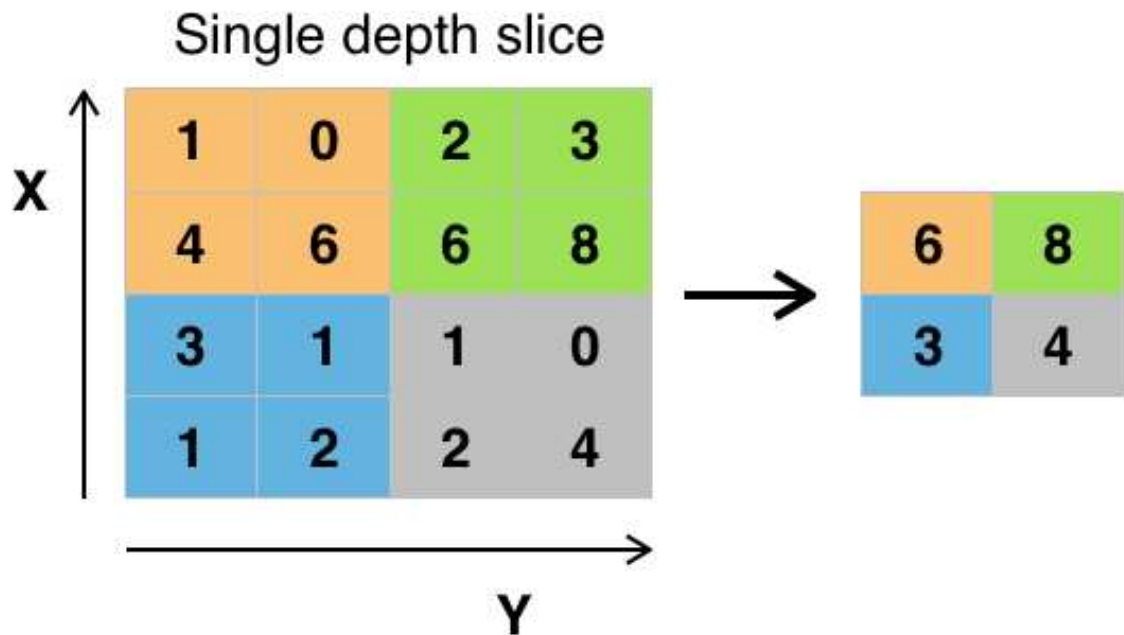


Рисунок 3 – Пример работы слоя субдискретизации. Фильтр max pooling, 2x2, с шагом 2.

```

layer {
  name: "image28x28"; type: "MemoryData"; top: "image28x28"; top: "label"
}
layer {
  name: "conv1"; type: "Convolution"; top: "conv1"; bottom: "image28x28";
  convolution_param { weight_filler { type: "xavier" }; bias_filler { type: "constant" } }
}
layer {
  name: "pool1"; type: "Pooling"; top: "pool1"; bottom: "conv1";
  pooling_param { pool: MAX; kernel_size: 2; stride: 2; }
}
layer {
  name: "conv2"; type: "Convolution"; top: "conv2"; bottom: "pool1";
  convolution_param { weight_filler { type: "xavier" }; bias_filler { type: "constant" } }
}
layer {
  name: "pool2"; type: "Pooling"; top: "pool2"; bottom: "conv2";
  pooling_param { pool: MAX; kernel_size: 2; stride: 2; }
}

```

```

}
layer {
  name: "ip1"; type: "InnerProduct"; top: "ip1"; bottom: "pool2";
  inner_product_param { num_output: 500; weight_filler { type: "xavier" };
bias_filler { type: "constant" } }
}
layer {
  name: "relu1"; type: "ReLU"; top: "ip1"; bottom: "ip1";
}
layer {
  name: "drop1"; type: "Dropout"; top: "drop1"; bottom: "ip1";
}
layer {
  name: "ip2"; type: "InnerProduct"; top: "ip2"; bottom: "drop1";
  inner_product_param { num_output: 500; weight_filler { type: "xavier" };
bias_filler { type: "constant" } }
}

```

Для краткости, были убраны слои с входными и выходными нейронами, некоторые параметры и детали настройки отдельных слоев.

Как вы можете видеть, данная структура является отражает типовую структуру сверточной нейронной сети (Рисунок 4).

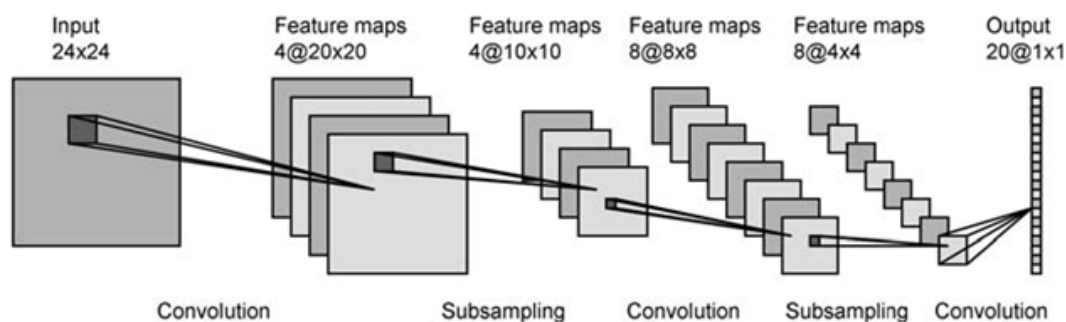


Рисунок 4 – Типовая структура сверточной нейронной сети.

Более важным вопросом является подбор параметров описанных слоев. Данный вопрос слишком комплексный для того, чтобы осветить его в рамках данной обзорной статьи.



**Выводы:**

1. Проведена сравнительная характеристика методов распознавания автомобильных номеров, выявлены их достоинства и недостатки.
2. Для решения задачи распознавания номеров было спроектировано несколько сверточных нейронных сетей с использованием фреймворка для построения нейронных сетей Caffe.
3. В результате проведенных исследований, в ряде тестовых условий (были использованы различные наборы видео) удалось добиться качества распознавания номеров, превосходящее шаблонный метод при времени работы сети лишь незначительно превышающем время работы шаблонного метода. В настоящий момент производятся исследования, направленные на оптимизацию работы сети.

**Литература**

1. Зейлер М.Д., Фергус Р. Stochastic Pooling for Regularization of Deep Convolutional Neural Networks – 2013.
2. Коллоберт Р., Вестон Дж. A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning // Материалы 25-ой международной конференции о вопросах машинного обучения. ICML '08. – Нью-Йорк, США: ACM, 2008. – С. 160–167.
3. Корекадо К., Мур Т., Номура О., Андо Х., Нкано Т., Матсуго М., Ивата А. A Convolutional Neural Network VLSI for Image Recognition Using Merged/Mixed Analog-Digital Architecture // Интеллектуальные информационные и инженерные системы, основанные на знаниях. – 2003 – С. 169–176.
4. Крижевский А. ImageNet Classification with Deep Convolutional Neural Networks – 2013.
5. ЛеКун И. LeNet-5, convolutional neural networks – 2013.
6. «Экономическая характеристика географической информационной системы мониторинга тракторного парка в сельском хозяйстве» / Третьяков И.И., Ткаченко В.В. // В сборнике: «Научное обеспечение агропромышленного комплекса».

**References**

1. Zeiler M.D., Fergus R. Stochastic Pooling for Regularization of Deep Convolutional Neural Networks – 2013.
2. Collobert R., Weston J. 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, 2008. – P. 160–167.
3. Korekado K., Morie T., Nomura O., Ando H., Nakano T., Matsugu M., Iwata A. A Convolutional Neural Network VLSI for Image Recognition Using Merged/Mixed Analog-

Digital Architecture // Knowledge-Based Intelligent Information and Engineering Systems. – 2003 – P. 169–176.

4. Krizhevsky A. ImageNet Classification with Deep Convolutional Neural Networks – 2013.

5. LeCun Y. LeNet-5, convolutional neural networks – 2013.

6. «JeKONOMICHESKAJa HARAkTERISTIKA GEOGRAFIChESKOJ INFORMACIONNOJ SISTEMA MONITORINGA MASHINOTRAKTORNOGO PARKA V SEL"SKOM HOZJaJSTVE» / Tret'jakov I.I., Tkachenko V.V.. // V sbornike: «NAUChNOE OBESPEChENIE AGROPROMYShLENNOGO KOMPLEKSA».