

ТЕХНИЧЕСКИЕ НАУКИ

РАЗРАБОТКА СИСТЕМЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Алексеев Юлиа Вячеславовна

*бакалавр по направлению подготовки «Информационные системы и технологии»
Институт сферы обслуживания и предпринимательства (филиал) ДГТУ, г. Шахты*

Аннотация

В статье описывается проблема распознавания образов, а также одна из самых сложных задач в этой области – распознавание изображений. В статье уделено особое внимание свёрточным нейронным сетям, как наиболее эффективному методу ее решения, а также описаны подходы к их обучению и результаты проведенных экспериментов. По результатам исследований была реализована система распознавания изображений, принципы работы которой подробно рассмотрены в статье.

Abstract

The article describes a pattern recognition problem as well as one of the most difficult tasks in this field, which is image recognition. The article focuses on convolutional neural networks as the most effective way to solve this, describes training techniques for neural networks, and the results of the experiments. Based on the research results, the image recognition system was implemented, the main ideas of which are discussed in detail in the article.

Ключевые слова: распознавание изображений, нейронные сети, сверточные нейронные сети, глубокое обучение, обучение нейронных сетей.

Keywords: image recognition, neural networks, convolutional neural networks, deep learning, training neural networks.

Задачу распознавания образов можно описать как отнесение предоставляемых данных к определенному классу при помощи выявления важных характеристик этих входных данных из всего набора менее важных. Проблема распознавания образов является одной из главных в теории интеллектуальных систем. Одной из самых сложных задач, решаемой в этой области, является распознавание изображений [3].

В настоящее время задача по распознаванию изображений часто решается на практике: идентификация человека в системах безопасности, анализ медицинских снимков пациентов, построение машинного зрения и т.д. Довольно часто изображения для анализа могут быть искажены шумом или просто повреждены. В сравнении со многими существующими методами распознавания изображений, искусственные нейронные сети способны решать эту задачу даже на подобных изображениях.

Искусственные нейронные сети обладают эффективными механизмами обучения, которые помогают избежать необходимости отбора важных для распознавания признаков из всего их набора. Необходимо также отметить, что параллельная работа нейронов способствует более эффективному и быстрому решению задачи распознавания.

Несмотря на то, что для успешной реализации нейросетевого подхода требуется большая выборка данных для обучения сети и специальная архитектура, которая бы эффективно решала поставленную задачу, для системы распознавания изображений был использован именно этот подход. Основными причинами этого выбора стали:

- возможность решения задач с плохо выделяемыми зависимостями между входными данными, что позволяет работать с неполными данными;
- возможность отсева неинформативных для анализа характеристик данных, таким образом исчезает необходимость дополнительного анализа;
- возможность переобучения нейронных сетей в новых условиях окружающей среды.

В процессе практической реализации нейронных сетей зачастую возникает ряд проблем, которые необходимо принимать во внимание: высокие вычислительные ресурсы, затрачиваемые на обучение нейронных структур, проблема переобучения и т.д.

Deep learning и сверточные НС

Первые архитектуры глубоких нейронных сетей появились еще в 80-е годы прошлого века, но сама идея применения таких сетей начала развиваться сравнительно недавно. Такой перерыв в развитии данной области связан с трудностями, возникшими в процессе обучения [2].

В настоящее время алгоритмы Deep learning широко используются для решения многих практических задач. Нейронные сети, обучаемые при помощи алгоритмов Deep learning, не только показывают более высокую точность в сравнении с альтернативными методами обучения сетей, но и в ряде задач проявляют понимание семантики информации, которая им подается.

Теория Deep learning дополняет обычные методы машинного обучения специальными алгоритмами для анализа входной информации на нескольких уровнях представления.

В результате, нейронная сеть, обученная с помощью алгоритмов Deep learning, тратит меньше входной информации для обучения, но в то же время она способна анализировать информацию с гораздо более высокой точностью [5].

Свёрточные нейронные сети – еще одна популярная тема в области искусственного интеллекта.

В основном такая архитектура применяется для работы с изображениями. Основная причина этого заключается в наличии специальных слоев, которые помогают получить лучшие результаты по сравнению с другими топологиями сетей. Структура сверточной нейронной сети изображена на рисунке 1.

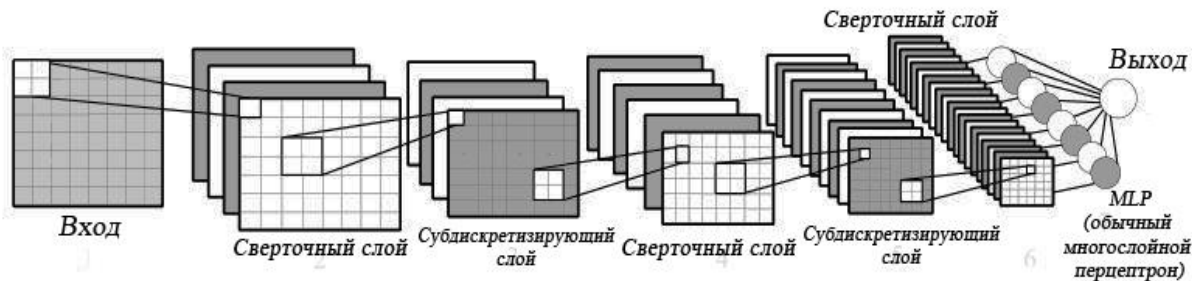


Рисунок 1. Структура сверточной нейронной сети

Обучение сверточных нейронных сетей на малом объеме данных

При построении нейронной сети для распознавания изображений на малом объеме обучающих данных может быть использован один из трех подходов [1]:

1. полное обучение нейронной сети;
2. использование bottleneck-признаков уже обученной сети;
3. надстройка последних слоев обученной сети.

Полное обучение нейронной сети

Одним из подходов обучения сверточных нейронных сетей является обучение с нуля. В данном случае необходимо помнить, что обучение на малом объеме данных может привести к проблеме переобучения.

Один из вариантов решения проблемы переобучения – это увеличение количества данных для обучения. Но во многих случаях этого недостаточно, так как данные могут быть сильно коррелированы. Дропаут (Dropout) также помогает бороться с переобучением, не позволяя слою увидеть дважды один и тот же образ, таким образом, действуя аналогично аугментации (Augmentation). Таким образом, можно сделать вывод, что аугментация и дропаут помогают нарушить случайные корреляции, возникающие в данных.

На практике метод полного обучения сети реализуется крайне редко, так как проблематично найти набор данных достаточного размера для обучения сети. Вместо этого обычно принято брать уже обученную сверточную нейронную сеть и использовать ее для решения поставленной задачи.

Использование bottleneck-признаков уже обученной сети

Второй подход состоит в том, чтобы использовать нейронную сеть, которая была предварительно обучена на большом наборе данных, например, ImageNet (1,2 миллионов изображений). Такая сеть уже обладает признаками, которые полезны для большинства проблем компьютерного зрения.

Основная идея этого подхода состоит в следующем: использовать сверточную нейронную сеть,

предварительно обученную на наборе данных, затем убрать последний полносвязный слой. После чего запустить один раз полученную модель на имеющемся обучающем наборе данных для тестирования, записывая при этом bottleneck-признаки в массив. Bottleneck-признаками называются те, которые извлекаются из малоразмерного скрытого слоя, расположенного уже ближе к последним скрытым слоям нейронной сети. В завершении, необходимо обучить полносвязный слой с учетом этих сохраненных признаков.

Подстройка последних слоев обученной сети

Идея третьего подхода заключается в подстройке последних слоев, например, сверточных и полносвязных. Работая уже с обученной нейронной сетью, необходимо лишь переучить ее на новом наборе данных, используя при этом очень небольшие обновления весов. Другими вариантами в рамках этого подхода являются также: подстройка всех слоев обученной сети или использование предыдущих слоев и подстройка слоев, находящихся ближе к выходу сети.

Система для распознавания изображений на базе SOA

Реализация алгоритмов распознавания изображений на основе сверточных нейронных сетей требует большой вычислительной мощности, и такие расчеты можно осуществить с помощью параллельных или распределенных систем. Основная тенденция развития современных высокопроизводительных распределенных вычислительных систем – это гибридные системы, состоящие из CPU и GPU, а также аппаратных ускорителей на FPGA. Главная проблема создания таких гибридных систем – это сложность объединения разнородных вычислительных ресурсов в единую распределенную систему.

В настоящее время при организации распределенных вычислений используются Grid-системы. При организации Grid-вычислений необходимо решить проблему объединения разнородных вычис-

лительных ресурсов в единую распределенную систему. В настоящее время в ИТ индустрии при создании информационных систем получила большое распространение сервис-ориентированная архитектура, благодаря своему главному преимуществу – возможности эффективной интеграции информационных ресурсов [4, с. 1463].

Архитектура построенной гибридной реконфигурируемой системы для высокопроизводительных вычислений изображена на рисунке 2. Как представлено на рисунке ниже, решаемая вычислительная задача состоит из клиентской и серверной частей.

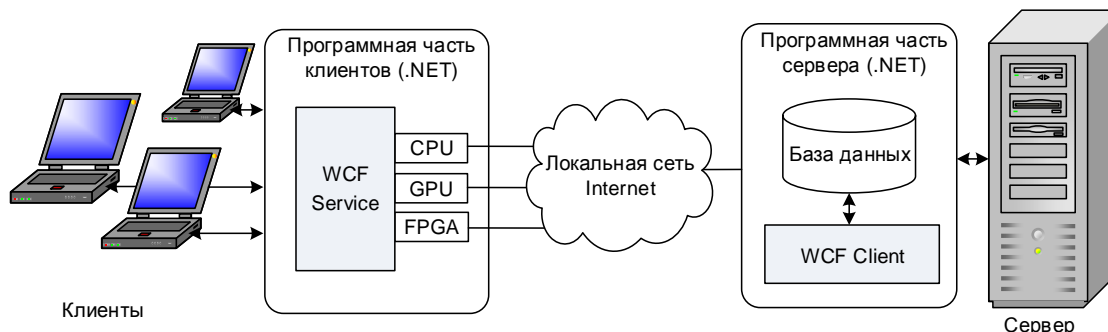


Рисунок 2. Архитектура предлагаемой системы для высокопроизводительных вычислений

Экспериментальные исследования

Для проведения экспериментальных исследований был использован набор изображений CIFAR-10, который часто применяют для сравнения алгоритмов распознавания изображений на основе сверточных нейронных сетей.

CIFAR-10 содержит 60 тысяч цветных изображений с размерностью 32x32 в 10 классах, в каждом классе находится 6 тысяч изображений. Для проведения обучения имеется 50 тысяч учебных изображений, для проведения тестов 10000 тысяч.

На данный момент многими разработчиками и исследователями в области Deep learning используется библиотека TensorFlow. Это новая библиотека

для создания систем машинного обучения с открытым исходным кодом от компании Google.

Для распределенного обучения нейронных сетей был применен асинхронный стохастический градиентный спуск. Главная особенность этого алгоритма заключается в том, что каждый клиент имеет собственный экземпляр данных и обучаемую модель. Такой подход позволяет значительно снизить нагрузку на сеть.

При проведении экспериментальных исследований на созданной системе для распознавания изображений были использованы три нейронные сети, построенные на основе рассмотренных подходов. Результаты проведенных экспериментов показаны на рисунке 3.

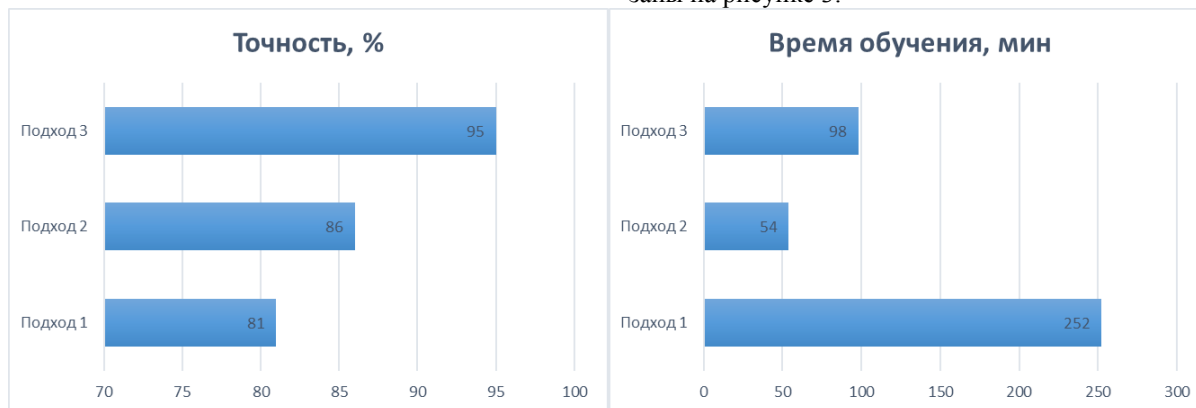


Рисунок 3. Результаты проведенных экспериментов

На основе полученных результатов экспериментальных исследований, можно сделать вывод, что для задачи распознавания изображений на основе сверточных нейронных сетей наилучший результат из соотношения время обучения / точность показывает подход №3, основанный на подстройке последних слоев обученной сети.

Вывод

В статье описаны достоинства применения алгоритмов Deep learning и сверточных нейронных

сетей для решения задачи распознавания изображений. Автором рассмотрены и реализованы три различных подхода обучения нейросетевых структур на малом объеме данных. Для сравнения этих подходов были проведены экспериментальные исследования, результаты которых были проанализированы и приняты во внимания при построении системы для распознавания изображений.

Список литературы:

1. Building powerful image classification models using very little data. — 2016 [Электронный ресурс].

Дата обновления: 05.06.2016. — URL: <https://blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html> (дата обращения: 26.06.2017).

2. Глубинное обучение: возможности, перспективы и немного истории. — 2016 [Электронный ресурс]. Дата обновления: 6.09.2016. — URL: <https://habrahabr.ru/company/it-grad/blog/309024/> (дата обращения: 20.06.2017).

3. Задача распознавания образов: Языки программирования. — 2011 [Электронный ресурс]. Дата обновления: 23.11.2011. — URL: <http://life->

prog.ru/view_zam2.php?cat=4&id=269&page=9 (дата обращения: 20.06.2017).

4. Shrihari, S.V., Pawan, K., Vishnu, M., Mahajan, S. Technological aspects of grid computing (2015) 2nd International Conference on Electronics and Communication Systems, ICECS 2015, art. no. 7124829, pp. 1463-1468.

5. Третье поколение нейросетей: «глубокие нейросети». — 2014 [Электронный ресурс]. Дата обновления: 27.11.2014. — URL: http://scorcher.ru/neuro/neuro_nets/neuro_nets1.php (дата обращения: 20.06.2017).