



added services to meet consumer lifestyle needs. The mobile industry has an important role to play in several respects.

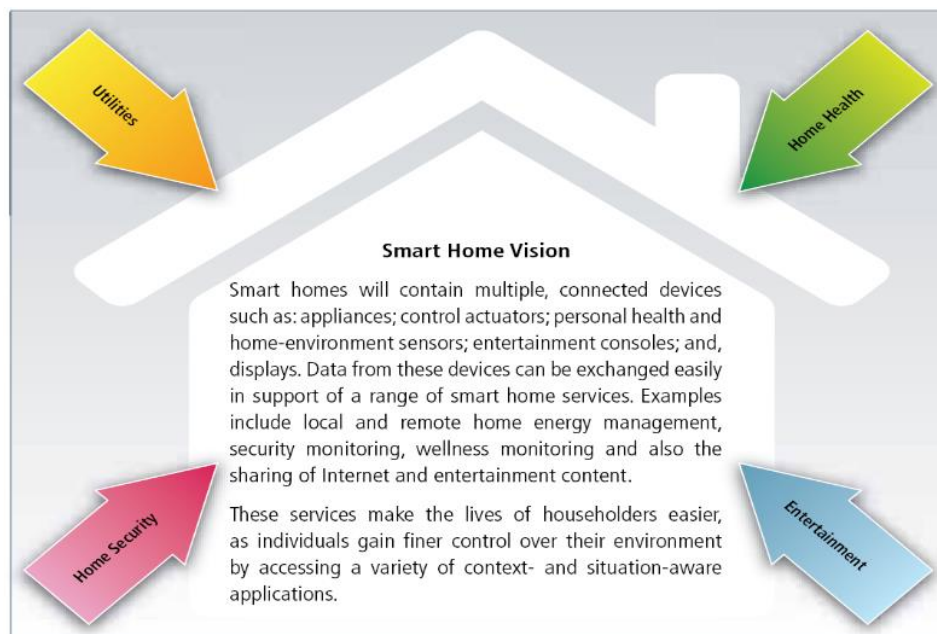


Fig. 1. Smart home vision driven by four key industry segments

Firstly, mobile devices are the most readily-accepted connected consumer devices in the world; most consumers are familiar with mobile devices and their sophisticated, yet intuitive, user interface functions.

References

1. Joan K. Lippincott, (2010) “A mobile future for academic libraries”, Reference Services Review, Vol. 38 Issue: 2, pp.205-213
2. “Best Home Automation System – Consumer Reports”. Retrieved 2016-11-22.

В.А. Акимов, В.Г. Литвинов

ИССЛЕДОВАНИЕ КАЧЕСТВА КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТЬЮ

(Самарский университет)

В настоящее время технологии компьютерного зрения широко распространены. Они применяются в задачах распознавания лиц, пешеходов, объектов, для медицинского анализа, а также во многих других задачах. С развитием технологий и увеличением вычислительных мощностей современных компьютеров стало возможным обучение нейронных сетей с большим числом скрытых слоёв (глубоких нейронных сетей). На данный момент в задачах классификации



изображений наилучшие результаты показывают свёрточные нейронные сети (Convolutional Neural Networks), которые каждый год выигрывают соревнования ImageNet Large Scale Visual Classification Challenge [1]. Целью данной работы является обзор свёрточной нейронной сети, построение модели свёрточной сети для решения задачи классификации изображений, а также анализ качества построенной модели.

Задача классификации в машинном обучении — это формализованная задача, в которой имеется множество объектов, разделённых некоторым образом на классы. Задано конечное множество объектов, для которых известно, к каким классам они относятся. Это множество называется обучающей выборкой. Классовая принадлежность остальных объектов неизвестна. Требуется построить алгоритм, способный классифицировать произвольный объект из исходного множества. Формальное описание задачи классификации: Пусть X — множество описаний объектов, Y — конечное множество номеров (имён) классов. Существует неизвестная целевая зависимость — отображение $y^*: X \rightarrow Y$, значения которой известны только на объектах конечной обучающей выборки $X^m = \{(x_1, y_1 \dots x_m, y_m)\}$. Требуется построить алгоритм $\alpha: X \rightarrow Y$, способный классифицировать произвольный объект $x \in X$.

Свёрточная нейронная сеть была впервые предложена Я.Лекуном и Й.Бенгуа [2]. Ученые, исследуя зрительную кору головного мозга кошки, пришли к выводу, что математическую модель поведения зрительных рецепторов кошки можно построить, опираясь на следующие этапы обработки изображений: свёртка исходного изображения при помощи нескольких фильтров; субдискретизация полученных на предыдущем шаге ключевых признаков; повтор предыдущих шагов, пока на выходе не получится достаточное количество признаков исходного изображения; использование модели полносвязного слоя для получения решения конкретной задачи. Свёрточная сеть состоит из 3 основных видов слоёв: слой свёртки, слой подвыборки (субдискретизации), выходной слой. Слой свёртки (Convolutional Layer) — это основной блок свёрточной нейронной сети, он включает в себя для каждого канала свой фильтр, ядро свёртки которого обрабатывает предыдущий слой по фрагментам (суммируя результаты матричного произведения для каждого фрагмента). Весовые коэффициенты ядра свёртки неизвестны и устанавливаются в процессе обучения. Скалярный результат каждой свёртки попадает на функцию активации, которая является нелинейной функцией. Слой активации обычно логически объединяют со слоем свёртки (считают, что функция активации встроена в слой свёртки). Наиболее распространённой функцией активации в свёрточных сетях является функция ReLU. Слой подвыборки (Subsampling), представляет собой нелинейное уплотнение карты признаков, при этом группа пикселей (обычно размера 2×2) уплотняется до одного пикселя, проходя нелинейное преобразование. Чаще всего используется функция максимума. Операция подвыборки позволяет существенно уменьшить пространственный объём изображения. Слой подвыборки обычно идёт сразу после слоя свёртки перед слоем следующей свёртки. После нескольких выполнений операций свёртки и уплотнения изображения с



помощью подвыборки система перестраивается от конкретной сетки пикселей с высоким разрешением к более абстрактным картам признаков, как правило, на каждом следующем слое увеличивается число каналов и уменьшается размерность изображения в каждом канале. В результате остаётся большой набор каналов, хранящих небольшое число данных. Эти данные объединяются и передаются на обычную полносвязную нейронную сеть (персептрон). Этот слой называется полносвязным (Full-Connected Layer). Архитектура сети представлена на рисунке 1.

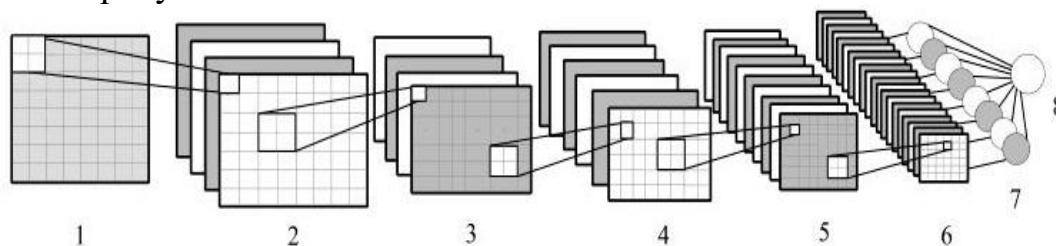


Рисунок 6 - Архитектура сети: 1 вход; 2,4,6 слои свёртки; 3-5 слои подвыборки; 7 полносвязный слой; 8 выход

Упрощенно можно считать, что цель обучения такой нейронной сети состоит в определении значений весов нейронов каждого слоя, чтобы при заданном входном векторе получить на выходе значения сигналов u_s , совпадающие с требуемой точностью с ожидаемыми значениями d_s при $s=1, 2, \dots, M$ [3]. Иными словами, при обучении ставится задача минимизации некоей функции потерь одним из алгоритмов оптимизации нейронной сети. В рамках решения данной задачи была использована функция средней кросс-энтропии по всем учебным примерам. Выбор такой функции потерь связан с использованием в выходном слое функции активации Softmax, которая является обобщением логистической функции для многомерного случая.

В качестве метода обучения был выбран оптимизационный алгоритм Adam (adaptive moment estimation) [4], который является одним из популярных алгоритмов оптимизации нейронных сетей. Данный алгоритм определяет стратегию подбора весов многослойной сети с применением градиентных методов оптимизации, сочетает в себе идею накопления движения по антиградиенту функции потерь и идею более слабого обновления весов для типичных признаков.

Для исследования классифицирующих возможностей свёрточной нейронной сети перейдём к практическому решению задачи классификации. Далее решается задача классификация фруктов, набор данных взят с сайта www.kaggle.com [5]. Набор данных для обучения представляет собой изображения формата JPEG размером 100x100 пикселей. Объём обучающей выборки составляет 15000 изображений, объём тестовой выборки 5000 изображений. Всего в наборе имеется 30 классов различных фруктов. Для построения модели свёрточной нейронной сети использовался язык программирования Python и библиотека TensorFlow [6]. В ходе решения данной задачи классификации была проведена серия экспериментов по выявлению наиболее удачного выбора



структуры нейронной сети, показывающего наилучшее качество классификации изображений. В каждом опыте изменялась структура сети путём добавления новых слоёв свёртки и подвыборки, чередующихся между собой, и отслеживались изменения в качестве классификации. Во всех опытах полносвязный слой для классификации содержал 512 нейронов, выходной полносвязный слой содержал 30 нейронов (по количеству классов предметной области), количество эпох обучения равнялось 20-ти. Результаты экспериментов представлены в таблице 1. Численная оценка качества классификации вычислялась как доля верных ответов нейронной сети на тестовой выборке.

Таблица 1 — Результаты экспериментов

Параметры сети	Качество
1 слой свёртки (32 карты признаков, функция активации ReLU), 1 слой подвыборки (размер 2x2, функция максимума)	95,04%
2 слоя свёртки (32 карты признаков на каждом, функция активации ReLU), 2 слоя подвыборки (размер 2x2, функция максимума)	95,82%
3 слоя свёртки (32 карты признаков на слоях 1-2, 64 карты признаков на слое 3, функция активации ReLU), 3 слоя подвыборки (размер 2x2, функция максимума)	97,80%
4 слоя свёртки (32 карты признаков на слоях 1-2, 64 карты признаков на слоях 3-4, функция активации ReLU), 4 слоя подвыборки (размер 2x2, функция максимума)	96,43%
5 свёрточных слоёв (32 карты признаков на слоях 1-2, 64 карты признаков на слоях 3-5, функция активации ReLU), 5 слоёв подвыборки (размер 2x2, функция максимума)	96,88%

После проведения экспериментов было выяснено, что с наращиванием чередующихся между собой слоёв свёртки и подвыборки возрастает качество классификации, наилучший результат был получен, когда в сети было по 3 слоя свёртки и подвыборки. Затем с дальнейшим наращиванием слоёв качество классификации ухудшалось. Таким образом, в серии экспериментов было найдено оптимальное решение поставленной задачи классификации на заданной предметной области. Полученное решение является локальным максимумом функции качества классификации в зависимости от количества слоёв сети. График данной зависимости показан на рисунке 2.

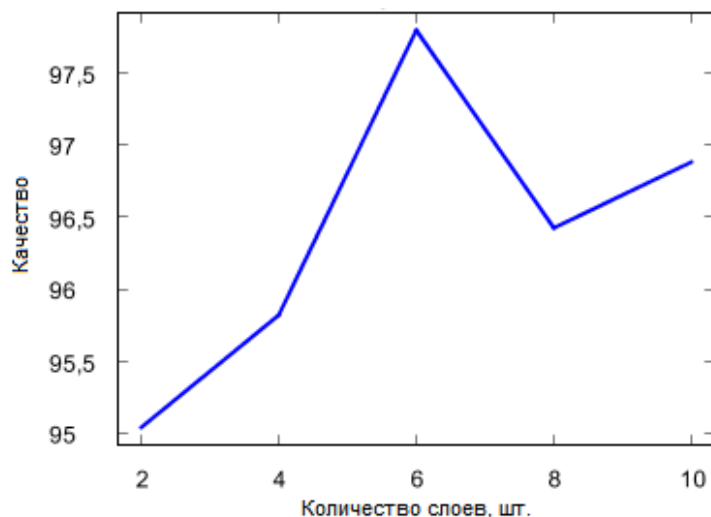


Рисунок 7 – График зависимости качества обучения от кол-ва слоёв

Литература

- 1 Image-net [Электронный ресурс]. – <http://www.image-net.org/challenges/LSVRC/>
- 2 LeCun Y. et al. Convolutional networks for images, speech, and time series //The handbook of brain theory and neural networks. – 1995. – Т. 3361. – №. 10. – С. 1995.
- 3 Осовский С. Нейронные сети для обработки информации / Пер. с пол. И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.: ил.
- 4 Ruder S. An overview of gradient descent optimization algorithms //arXiv preprint arXiv:1609.04747. – 2016.
- 5 Kaggle [Электронный ресурс]. – <http://www.kaggle.com>
- 6 Abadi M. et al. TensorFlow: A System for Large-Scale Machine Learning //OSDI. – 2016. – Т. 16. – С. 265-283.

Г.А. Альгашев, О.П. Солдатова

НЕЙРОПЛАСТИЧНОСТЬ СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

(Самарский университет)

Цель данной научной работы состоит в исследовании эффективности переноса знаний (transfer learning) предварительно обученной свёрточной нейронной сети для решения других задач.

Для классификации и распознавания изображения большую популярность получили свёрточные нейронные сети. Такие сети состоят из нескольких свёрточных слоёв, которые в процессе обучения извлекают признаки подаваемых на вход изображений. Верхние слои сети занимаются извлечением низкоуровневых признаков, а глубокие слои занимаются извлечением высокоуровневых признаков. Чем ближе свёрточный слой находится к выходу сети, тем менее абстрактными становятся извлекаемые признаки (рисунок 1) [1].