

УДК 004.032.26:581.522-047.44

ИСПОЛЬЗОВАНИЕ СВЁРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ АНАЛИЗА СОСТОЯНИЯ РАСТЕНИЙ



А.В. Барковский
Магистрант кафедры информатики БГУИР, инженер-программист



А.В. Жвакина
Доцент кафедры информатики БГУИР, кандидат технических наук, доцент

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь
COOO HiQo Solutions, Республика Беларусь
E-mail: anton@swarmer.me, zhvakina@bsuir.by

Аннотация. На протяжении всего времени существования компьютеров одними из самых важных задач были задачи, связанные с восприятием изображений и видео, также в совокупности называемые задачами компьютерного зрения. Наибольшая точность при решении таких задач на сегодняшний день достигается при использовании свёрточных нейронных сетей. В данном докладе рассматривается возможность использования свёрточных нейронных сетей для классификации растений в биологических исследованиях.

Ключевые слова: свёртка, нейронная сеть, классификация изображений, компьютерное зрение.

Последние достижения в разработке свёрточных нейронных сетей открывают широкий простор для их применения во множестве различных отраслей и наук, одной из которых является биология.

Автоматическая обработка изображений при помощи нейронных сетей может использоваться для классификации растений по виду или состоянию здоровья, обнаружения болезней, паразитов или сорных растений, автоматического отслеживания прогресса роста растения под воздействием различных факторов и многих других задач.

Задача является актуальной, так как сейчас на биологическом факультете Белорусского Государственного Университета данные мероприятия проводятся вручную, требуют больших затрат времени и человеческих ресурсов. Использование нейронной сети, способной самостоятельно отличать здоровые растения от умирающих, позволит автоматизировать процесс проведения биологических исследований. Это увеличит объём исследований, проводимых биологами, упростит наблюдение за экземплярами растений и снизит нагрузку на работников лабораторий.

Сравнительный анализ возможностей таких методов анализа изображений, как вручную подобранные свёртки, гистограммы цветов или ориентированных градиентов показал, что нейронные сети наиболее эффективны для решения задачи классификации растений. Они позволяют достичь необходимого уровня точности и обеспечивают скорость распознавания, возрастающую при аппаратной реализации.

Свёрточные сети могут состоять из множества видов свёрточных операций[1]. Схема стандартной свёрточной нейронной сети представлена на рисунке 1.

Стоит отметить, что хотя нейронные сети – это относительно ресурсоёмкий метод обработки изображений, с развитием вычислительных компонентов (в частности, графических

ускорителей) в последние годы они зачастую пригодны для использования в условиях ограниченных вычислительных ресурсов. К примеру, простые свёрточные нейронные сети могут обрабатывать данные со скоростью сотен кадров в секунду. Также существует нейронная сеть, способная с высокой точностью разделять изображение на объекты, работающая со скоростью около 100 кадров в секунду на мобильном телефоне[2].

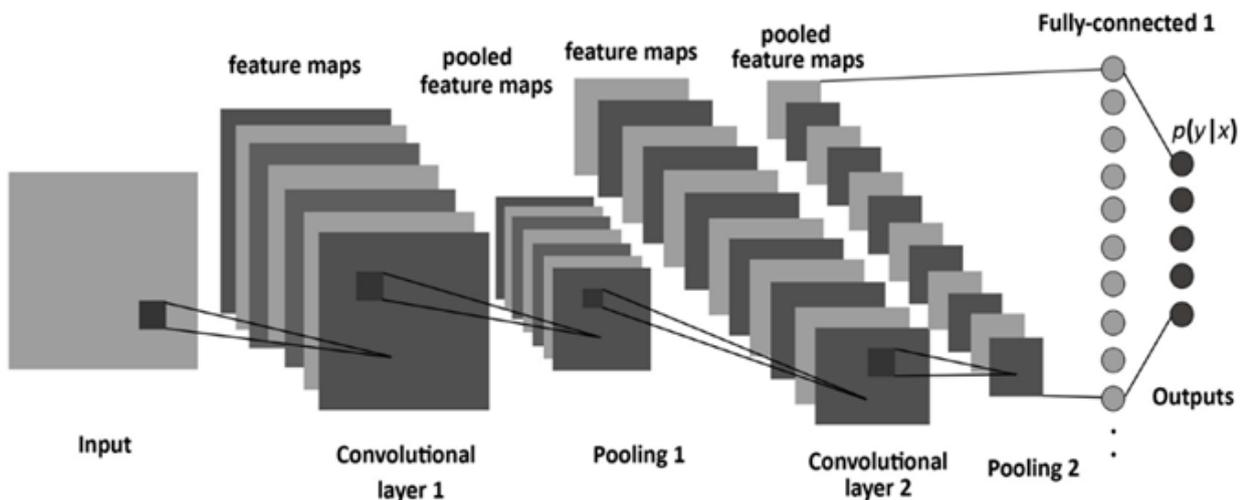


Рисунок 1. Свёрточная нейронная сеть

В работе исследовался вид свёрточных сетей, особенно хорошо адаптированный под мобильные устройства. Так как доступные в телефонах и стационарных компьютерах ресурсы продолжают расти, ресурсоёмкость нейронных сетей будет отходить на второй план, и в ближайшем будущем применимость и актуальность свёрточных сетей возрастет, особенно в рамках решаемой задачи.

Для разработки нейронной сети использована архитектура MobileNet. Среди её преимуществ[3]:

- 1 Низкая требовательность к ресурсам, что допускает возможность её использования в мобильных и других устройствах с низкой вычислительной мощностью.
- 2 Уменьшенное время, необходимое на обучение сети.
- 3 Адаптируемость под различные требования при помощи двух простых и хорошо изученных гиперпараметров.
- 4 Относительная (в сравнении с некоторыми другими нейронными сетями) простота и небольшое количество параметров снижают тенденцию сети к переобучению, когда точность, достигнутая на тренировочных данных, не обобщается на другие входные данные.
- 5 Точность, очень приближенная к другим намного более сложным видам свёрточных нейронных сетей.

Недостатком является то, что свёртки, разделяемые по глубине, несмотря на их эффективность, несколько снижают её точность. Также в дальнейшей работе планируется исследовать возможность квантизации MobileNet, в частности её производительность и точность после этой процедуры.

В ходе разработки нейронной сети исследовались различные вспомогательные методы, позволяющие улучшить точность результатов нейронных сетей или достичь более быстрой сходимости при обучении. Среди этих методов:

- ReLU (rectified linear unit) – наиболее широко используемая в данный момент функция активации, определяемая как $f(x) = \max(x, 0)$. Её преимущество над используемой ранее сигмоидальной функцией в том, что она избегает угасания градиентов при высоких входных

значениях, что облегчает обучение глубоких нейронных сетей. Одной возможной проблемой является застревание нейрона в деактивированном нулевом состоянии при установке в процессе обучения очень низких входных весов. Для решения этой проблемы можно использовать ELU — exponential linear unit, которые могут принимать значения от -1 и обладают ненулевой производной на всей области определения.

– Инициализация весов должна проводиться определённым образом, чтобы избежать неконтролируемого угасания или возрастания активаций. Широко используется Xavier-инициализация. Эта процедура инициализирует веса таким образом, чтобы средние значения активаций всех слоёв сети были схожи.

– Нормализация входных и выходных данных необходима из-за формы используемых функций активации. В случае с сигмоидальными функциями критично нормализовать данные так, чтобы их значения были строго больше -1 и строго меньше 1, так как сигмоидальная функция неспособна достичь других значений. Даже при использовании других функций активации, таких как ReLU, целесообразно пронормировать данные, чтобы значения не находились далеко от типичных значений, от -1 до 1.

– Батч-нормализация означает нормирование данных в пределах батча не только на входе и выходе, но и между слоями, что уменьшает сдвиг, который сети нужно изучить при переходе между слоями, облегчая весь процесс обучения[4].

– Обрезание градиента – ограничение градиента, используемого при градиентном спуске, некоторым значением. Это позволяет избежать слишком больших изменений весов при обучении, что, как было эмпирически установлено в ходе исследований, особенно важно при обучении рекуррентных нейронных сетей.

– Аугментация тренировочных данных дополнительными примерами помогает нейронной сети обобщаться на новые входные данные. Примером аугментации могут служить повороты картинок, добавление шума, сдвиг цветов, и удаление некоторых частей изображения.

– Выброс (dropout) означает деактивацию некоторого количества случайных нейронов в процессе обучения таким образом, что их значения делаются равными 0. Это также помогает нейронной сети надёжнее обобщаться на новые данные благодаря тому, что сеть учится проводить одни вычисления несколькими нейронами.

– Регуляризация (угасание весов) – добавление в loss-функцию выражения, зависящего от весов нейронов (квадрат весов в l2-регуляризации), с некоторым коэффициентом. Ещё один из способов, который не позволяет нейронной сети переобучиться на тренировочных данных, с помощью ограничения их возможности расти до больших значений.

– Балансирование loss-функции требуется, если в тренировочных данных некоторые классы встречаются намного чаще других. В этом случае сеть учится хорошо работать с часто встречающимися данными, но может игнорировать редко встречающиеся, так как их суммарное влияние на функцию потерь невелико.

– Оптимизаторы обучения ускоряют процесс схождения модели к оптимальным параметрам путём адаптации скорости обучения под каждый параметр в отдельности. Некоторые из известных вариантов оптимизации – добавление момента в стандартный стохастический градиентный спуск, RMSprop, Adagrad, AdaDelta, Adam. Хотя оптимизаторы могут ускорять или даже делать возможным схождение модели к оптимальной, результаты проведенных исследований показали неочевидность выгоды от их использования. В частности, результаты исследований показали ухудшение обобщения модели на новые данные при использовании оптимизаторов по сравнению с обычным стохастическим градиентным спуском [5].

– Оптимальное формирование батчей для обучения влияет на сходимость модели. Использование слишком больших батчей ухудшает обобщение модели на новые данные.

Также данные для обучения могут изначально быть в некотором порядке, который помешает процессу обучения (к примеру, сначала только один класс, затем только второй класс). Решению проблемы способствует перемешивание тренировочных данных перед обучением.

При реализации модели на основе MobileNet использованы библиотеки TensorFlow и Keras, облегчающие описание, обучение и эксплуатацию нейронных сетей. Эти библиотеки появились относительно недавно и в них присутствуют пробелы в обычно требуемом функционале. Поэтому были разработаны собственные компоненты для реализации необходимых возможностей.

Таким образом, нейронная сверточная сеть, разработанная с использованием архитектуры MobileNet, облегчает задачу исследования состояния растений по внешним признакам, позволит выполнять широкомасштабные эксперименты и экономить при этом человеческие ресурсы и время.

Список литературы

- [1]. An Introduction to different Types of Convolutions in Deep Learning. [Электронный ресурс] – Режим доступа: <https://towardsdatascience.com/types-of-convolutions-in-deep-learning-717013397f4d>. – Дата доступа : 12.03.2017.
- [2]. Mobile Real-time Video Segmentation. [Электронный ресурс] – Режим доступа: <https://research.googleblog.com/2018/03/mobile-real-time-video-segmentation.html>. – Дата доступа : 12.03.2017.
- [3]. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. [Электронный ресурс] – Режим доступа: <https://arxiv.org/abs/1704.04861>. – Дата доступа : 09.03.2017.
- [4]. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. [Электронный ресурс] – Режим доступа: <https://arxiv.org/abs/1502.03167>. – Дата доступа : 12.03.2017.
- [5]. The Marginal Value of Adaptive Gradient Methods in Machine Learning. [Электронный ресурс] – Режим доступа: <https://arxiv.org/abs/1705.08292>. – Дата доступа : 09.03.2017.

APPLICATION OF CONVOLUTIONAL NEURAL NETWORKS FOR PLANT CONDITION ANALYSIS

A. BARKOVSKY

Master student, Department of Computer Science BSUIR, software engineer

A.V. ZHVAKINA, PhD

Associate Professor, Department of Computer Science BSUIR

*Belarusian State University of Informatics and Radioelectronics, Republic of Belarus
HiQo Solutions, Republic of Belarus
E-mail: anton@swarmer.me, zhvakina@bsuir.by*

Abstract. Throughout the whole existence of computers, among the most important tasks were tasks related to the perception of images and video, also referred to as computer vision problems. The greatest accuracy in solving such problems today is achieved with the use of convolutional neural networks. This paper discusses the possibility of using convolutional neural networks for classifying plants in biological studies.

Key words: B convolution, neural network, image classification, computer vision.