

УДК 004.89

ПРИМЕНЕНИЕ СВЕРТОЧНОЙ НЕЙРОННОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ ФРУКТОВ

¹Данг Т.Ф.Т., ¹Юрченко А.В., ²Динь В.Т., ¹Ляшенко Д.А., ¹Нгуен Т.К.

¹Национальный исследовательский Томский государственный университет, Томск,
e-mail: dangthiphuongchung2018tsu@gmail.com, lyashenkodmtriy@gmail.com, cuongntit@vamaru.edu.vn;

²Национальный исследовательский Томский политехнический университет,
Томск, e-mail: dinh88@mail.ru

В этой статье обсуждается использование глубокого обучения для распознавания фруктов и его применение. В начале описываются особенности глубинного обучения и подробно характеризуется структура популярных архитектур глубоких нейронных сетей, а также подходы к обучению нейросетевой модели. Далее анализируется потенциал использования глубокого обучения в качестве продвинутого инструмента интеллектуального анализа данных. В ходе работы были рассмотрены некоторые современные подходы к решению задачи распознавания фруктов по их изображению и было дано подробное описание работы сверточной нейронной сети и архитектурного метода EfficientNet, используемого для распознавания фруктов с использованием набора данных Fruit360. Результаты показали, что предложенная модель достигает точности 95%. Конечной целью проведенного исследования была разработка автоматизированных технических решений для повышения точности распознавания фруктов по их изображениям. Представленная в данной работе система также направлена на повышение эффективности и простоты процесса расчета цен на товары в торговом центре. Ключевым нововведением в проделанном исследовании являются алгоритмы распознавания на изображениях фруктов на основе глубокого обучения. Насколько нам известно, результаты данного исследования превосходят результаты, представленные в других работах, также основанных на использовании описанного выше набора данных. Проведенные в этом исследовании эксперименты показали, что предложенный подход является перспективным решением проблемы распознавания пищевых изображений. Описанная в данной статье работа включает в себя дальнейшее улучшение производительности алгоритмов и интеграцию их в систему реального мира.

Ключевые слова: компьютерное зрение, машинное обучение, нейронные сети, сверточные нейронные сети, распознавание фруктов, набор данных Fruits 360

A FRUIT RECOGNITION SYSTEM USING A CONVOLUTIONAL NEURAL NETWORK

¹Dang T.F.T., ¹Yurchenko A.V., ²Din V.T., ¹Lyashenko D.A., ¹Nguyen T.K.

¹National Research Tomsk State University, Tomsk, e-mail: dangthiphuongchung2018tsu@gmail.com,
lyashenkodmtriy@gmail.com, cuongntit@vamaru.edu.vn;

²National Research Tomsk Polytechnic University, Tomsk, e-mail: dinh88@mail.ru

This paper discusses the use of deep learning (DL) for fruit recognition and its applications. We first briefly introduced deep learning and described in detail the structure of some popular deep neural network architectures and approaches for model training. We also analyzed the potential of using deep learning as an advanced data mining tool. In the course of the paper, we reviewed some state-of-the-art approaches to the task of fruit recognition from images. Further, the convolutional neural network (CNN) and EfficientNet for fruit recognition using the Fruit360 dataset will be explained in more detail. The results show that the proposed model achieves an accuracy of up to 95%. The ultimate goal of our research is to develop automated technical solutions to improve the accuracy of fruit recognition from their images. The system we propose in this paper is aimed at increasing the efficiency and simplicity of the process of calculating the price of goods in a retail center. The key novelty in this work is the algorithms for fruit image recognition based on deep learning. To the best of our knowledge, the results of this research have outperformed the results of other submitted work using this dataset. Our experiments have shown that the proposed approach is a promising solution to the food recognition problem. This article's work includes further improvements in algorithm performance and integration into a real-world system.

Keywords: computer vision, machine learning, neural networks, convolutional neural networks, fruit recognition, dataset Fruits 360

В последние годы для распознавания на изображениях объектов, в том числе фруктов, широко используются различные техники с применением технологии компьютерного зрения, такие как машинное обучение и особенно глубокое обучение [1].

Благодаря возможностям современных информационных технологий все чаще применяются методы компьютер-

ного зрения, машинного обучения и особенно глубокого обучения для распознавания на изображениях объектов, в том числе для распознавания товаров на полках в торговых точках. Важно отметить, что использование глубоких нейронных сетей для идентификации, классификации и различения товаров по их изображениям показывает большую эффективность

по сравнению с использованием других алгоритмов. Несмотря на это, существующие в настоящее время подходы еще не достигают такого уровня точности, при котором она могла бы быть использована в практическом применении.

Цель данной работы заключается в разработке, тестировании и реализации алгоритма для распознавания различных видов фруктов на изображениях, полученных с цифровой камеры, с помощью глубокого обучения.

Материалы и методы исследования

Глубокое обучение – это подраздел машинного обучения, который в свою очередь является подразделом искусственного интеллекта. Оно представляет собой набор методов, которые позволяют моделировать высокоуровневые абстракции данных. В рамках глубокого обучения компьютерная модель анализирует и извлекает полезную информацию из изображений, звуковых или текстовых материалов для последующего ее применения. Эти модели могут достигать очень высокой точности, превышающей возможности человека. Модели обычно обучаются с использованием большого набора маркированных данных и сложных нейронных сетевых архитектур с множеством слоев, что позволяет достигать высокой точности.

Концепция глубокого обучения была впервые представлена еще в 1980-х гг., однако эта идея не стала сразу популярной благодаря двум причинам: требование огромного количества помеченных данных и значительной вычислительной мощности [2]. В последнее десятилетие наблю-

дается рост числа прикладных программ с реализацией в них глубокого обучения, включая обработку естественного языка, классификацию изображений, поисковую систему и др. Обучение заключается в необходимости вывести полезную информацию из набора данных и создать внутреннюю картину того, что может использовать для действий эксперт.

Искусственная нейронная сеть (ИНС, Artificial Neural Networks, ANN), математическая модель, построенная по принципу организации и функционирования биологических нейронных сетей нервных клеток живого организма и являющаяся наиболее распространенным алгоритмом в области машинного обучения [3]. Она состоит из интегрированных вычислительных блоков, называемых нейронами. ИНС также состоит из входного, скрытого и выходного слоев. Входной слой берет на вход, например, изображение и передает его скрытому слою, а затем выходной слой выдает результат – максимальную вероятность того, какой объект имеется на изображении. Возможно иметь несколько скрытых слоев для работы с более сложными функциями.

Сверточная нейронная сеть (англ. Convolutional Neural Network, CNN)

Принципиальная структура сверточных нейронных сетей (СНС) похожа на структуру обычных нейронных сетей и состоит из нейронов, обучаемых весов и смещений [4]. На вход каждого нейрона поступает набор данных, с помощью функции активации выполняется определенная математическая операция, после которой полученное значение передается на следующий слой.

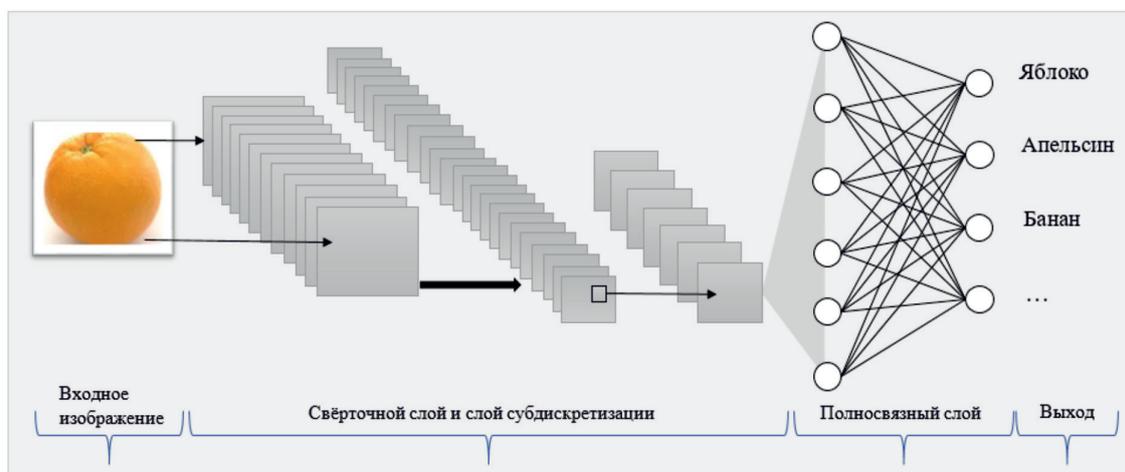


Рис. 1. Принципиальная структура сверточных нейронных сетей

EfficientNet

В 2019 г. компания Google представила модель EfficientNet, которая в настоящее время является одной из самых современных моделей сверточных нейронных сетей [5]. В этой статье показано, что точность модели CNN возрастает с увеличением количества фильтров в каждом слое, глубины (количества слоев в модели) и разрешения (размера входного изображения). Однако при увеличении этих величин стоимость вычислений экспоненциально увеличивается. Поэтому было создано несколько моделей семейства архитектур EfficientNet, которые отличаются количеством используемых параметров.

Существует 8 реализаций EfficientNet, считая от B0 до B7 по мере роста сложности сетевой архитектуры. В большинстве ситуаций результаты первоначального теста показывают более высокие точность и скорость. Такие модели могут использоваться для создания более точных и эффективных моделей, а также для идентификации и распознавания изображений, что дает преимущества в ситуациях интенсивного режима работы. В табл. 1 приве-

дена краткая информация об архитектуре EfficientNet-B0.

Набор данных

Для обучения и тестирования были выбраны изображения из набора данных fruits 360, который находится в открытом доступе на портале Kaggle. Этот набор данных содержит 77917 изображений различных фруктов по 103 категориям [6, 7]. Картинки фруктов были получены с помощью неподвижной камеры во время вращения фрукта с помощью электродвигателя. В качестве фона была использована белая бумага. В связи с неравномерностью освещения для извлечения фруктов с фона был применен алгоритм рекурсивной заливки. После удаления фона все картины были сжаты до размеров 100×100 пикселей по стандартным форматам RGB. Из набора данных fruits 360 мы взяли 17624 изображения из 25 категорий. Мы использовали 13218 изображений (75%) в качестве обучающего множества, а остальные 4406 изображений (25%) для тестирования модели. 25 категорий фруктов, которые мы использовали для проведения эксперимента, показаны в табл. 1 и на рис. 2.

Таблица 1

Архитектура EfficientNet-B0

Этап i	Оператор F_i	Разрешение $H_i \times W_i$	Каналы C_i	Слои L_i
1	Conv3x3	224 x 224	32	1
2	MBCConv1, k3x3	112 x 112	16	1
3	MBCConv6, k3x3	112 x 112	24	2
4	MBCConv6, k3x3	56 x 56	40	2
5	MBCConv6, k3x3	28 x 28	80	3
6	MBCConv6, k3x3	14 x 14	112	3
7	MBCConv6, k3x3	14 x 14	192	4
8	MBCConv6, k3x3	7 x 7	320	1
9	Conv1x1 & Pooling & FC	7 x 7	1280	1

Таблица 2

Пример фруктов в базе fruits 360

Название фруктов	Количество изображений для обучения	Количество изображений для тестирования
Желтое яблоко	492	164
Яблоко Гренни Смит	492	164
Яблоко Ред Делишес	492	164
Абрикос	492	164
Авокадо	426	142
Банан	489	163
Вишня	492	164
Кокос	489	163
Синий виноград	984	328
Белый виноград	489	163
Грейпфрут	489	163

Окончание табл. 2		
Название фруктов	Количество изображений для обучения	Количество изображений для тестирования
Кумкват	489	166
Лимон	492	164
Лайм	489	163
Мандарин	489	163
Манго	489	163
Апельсин	480	160
Груша	492	164
Зеленый перец	444	148
Желтый перец	666	222
Красный перец	666	222
Клубника	492	164
Помидор	738	246
Ананас	489	163
Киви	468	156

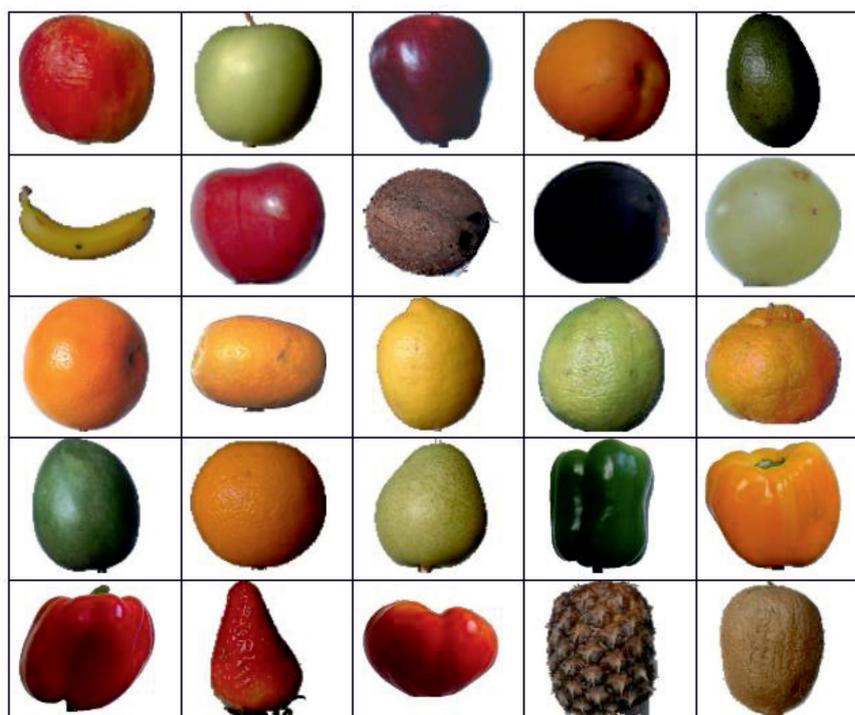


Рис. 2. Образцы фруктов в базе fruits 360

Оценки классификаторов

Как правило, результаты решения проблемы двоичной классификации обычно представлены в матрице ошибок (Confusion Matrix) (рис. 3), которая содержит 4 ячейки:

– верноположительные (TP) объекты, которые были классифицированы как положительные и действительно являются положительными (принадлежащими к данному классу);

– верноотрицательные (TN) объекты, которые были классифицированы как отрицательные и действительно отрицательные (не принадлежат к данному классу);

– ложноположительные (FP) объекты, которые были классифицированы как положительные, но фактически отрицательные;

– ложноотрицательные (FN) объекты, которые были классифицированы как отрицательные, но фактически положительные.

		Реальность	
		+	-
Прогноз	+	True Positive (истинно-положительное решение): прогноз совпал с реальностью, результат положительный произошел, как и было предсказано ML-моделью	False Positive (ложноположительное решение): ошибка 1-го рода, ML-модель предсказала положительный результат, а на самом деле он отрицательный
	-	False Negative (ложноотрицательное решение): ошибка 2-го рода – ML-модель предсказала отрицательный результат, но на самом деле он положительный	True Negative (истинно-отрицательное решение): результат отрицательный, ML-прогноз совпал с реальностью

Рис. 3. Матрица ошибок

Для оценки качества моделей используются три основных метрики: доля правильно классифицированных объектов (Accuracy), Точность (Precision) и Полнота (Recall) [8].

Accuracy – широко используемая и легкая для понимания метрика. Это отношение всех правильных прогнозов к общему числу всех предсказанных образцов.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Точность (precision) – это доля прогнозируемых положительных результатов, которые являются действительно верно-положительными результатами для всех положительно предсказанных объектов. Другими словами, точность дает нам ответ на вопрос «Из всех объектов, которые классифицированы как принадлежащие классу, сколько на самом деле принадлежит ему?»

$$\text{Точность} = \frac{TP}{TP + FP} \quad (2)$$

Полнота (recall) – пропорция всех верно-положительно предсказанных объектов к общему количеству действительно положительных. То есть полнота показывает, сколько образцов из всех положительных примеров были классифицированы правильно. Чем выше значение полноты, тем меньше положительных примеров пропущено в классификации.

$$\text{Полнота} = \frac{TP}{TP + FN} \quad (3)$$

Настройка компьютера

Для обучения глубоких нейронных сетей, таких как EfficientNet, необходимо иметь сервер с мощными вычислительными характеристиками. В табл. 3 указаны конфигурации аппаратного и программного обеспечения системы, использованной для проведения нашего обучения.

Таблица 3
Настройка компьютера

Название	Описание
Оперативная память	16 Гб
Процессор	Intel Core i7-4770 Haswell 4 x 3400-3900 МГц
Видеокарта	GeForce GTX 1060 Ti
Операционная система	Ubuntu 19.04
Python	3.7
Numpy	1.15

Результаты исследования и их обсуждение

В этой работе показано применение EfficientNet-b0 на набор данных Fruit 360 для определения улучшенной производительности системы классификации. Из набора данных Fruits 360 мы взяли 17624 изображения из 25 различных категорий: 75% из них используются для обучения, а 25% – для тестирования модели. Обучение сети проводится в 35 эпохах с размером партии 20. Сравнение предложенной модели с существующими моделями пока-

зывает, что результаты нашей модели являются положительными и многообещающими для реального применения. Благодаря такой повышенной точности и аккуратности будет более целесообразно повысить общую эффективность машины в распознавании фруктов. В качестве демонстрации была разработана программа на Python с использованием библиотеки PyQt. Главное окно программы показано на рис. 4.



Рис. 4. Главное окно программы

Заключение

В данной статье рассмотрен механизм распознавания фруктов, основанный на алгоритме EfficientNet. Скорость распознавания значительно улучшилась за время проведения эксперимента. Среди всех рассмотренных случаев модель достигла

лучшей точности при тестировании 98% в случае 4 от 11 до 15 эпох и лучшей точности при обучении 96,79% в 13 эпоху. Этот результат будет служить основой для разработки весов самообслуживания с автоматическим распознаванием товаров.

Список литературы

1. Chen, J.J., Ngo C.W. Deep-based ingredient recognition for cooking recipe retrieval. Proceedings of the 24th ACM International Conference on Multimedia. 2016. P. 32–41.
2. Rocha A., Hauage D.C., Wainer J., Goldenstein S. Automatic fruit and vegetable classification from images. Comput. Electron. Agric. 2010. Vol. 70. P. 96–104.
3. Deng L., Hinton G., Kingsbury B. New types of deep neural network learning for speech recognition and related applications: An overview. IEEE International Conference on Acoustics, Speech and Signal Processing. 2013. P. 8599–8603.
4. Zhang Y., Wu L. Classification of fruits using computer vision and a multiclass support vector machine. Sensors. 2012. Vol. 12. No. 9. P. 12489–12505.
5. Tan M., Le Q.V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. Proceedings of the 36th International Conference on Machine Learning. 2019. Vol. 5. arXiv: 1905.11946.
6. Kaggle Inc. Fruits 360 dataset. [Электронный ресурс]. URL: <https://kaggle.com/moltean/fruits> (дата обращения: 03.07.2021).
7. Muresan H., Oltean M. Fruit recognition from images using deep learning. Proceeding of the Acta Univ. Sapientiae, Informatica. 2018. Vol. 10. No. 1. P. 26–42.
8. Дудченко П.В. Метрики оценки классификаторов в задачах медицинской диагностики // Молодежь и современные информационные технологии: сборник трудов XVI Международной научно-практической конференции студентов, аспирантов и молодых учёных, г. Томск. Томск: Изд-во ТПУ, 2019. С. 164–165.