

# Подготовка системы распознавания объектов на базе TensorFlow и Keras

А. И. Тур<sup>1</sup>, А. Н. Кокоулин<sup>2</sup>, А. А. Южаков<sup>3</sup>

Пермский национальный исследовательский  
политехнический университет (ПНИПУ)

<sup>1</sup>tur.aleksandr93@mail.ru, <sup>2</sup>liga\_asu@mail.ru,  
<sup>3</sup>edu@at.pstu.ac.ru

А. Н. Лукичев

Санкт-Петербургский государственный  
электротехнический университет  
«ЛЭТИ» им. В.И. Ульянова (Ленина)

**Аннотация.** В данной статье рассмотрен пример повышения эффективности распознавания объекта по его изображению искусственной нейронной сетью на базе TensorFlow и Keras. Представлены результаты экспериментов по выбору количества распознаваемых классов, изменения параметров распознавания, а так же модификации обучающей выборки. На основании полученных данных сделаны выводы и предложен вариант реализации системы для заданной задачи.

**Ключевые слова:** искусственная нейронная сеть; глубокое обучение; распознавание; TensorFlow; Keras

## I. ВВЕДЕНИЕ

В последнее время стала популярна тема «цифровой утилизации и переработки» отходов. В группу этих проектов входят интеллектуальные конвейеры по сортировке мусора и различные средства приёма у населения отходов. Из последних наиболее популярным являются так называемые Reverse vending machine (дословно «обратный торговый автомат», автомат по приёму тары, RVM). Это автомат, позволяющий людям возвращать пустые контейнеры для напитков, такие как бутылки и банки, для переработки. Обычно автомат устроен таким образом, чтобы возвращать конечному пользователю сумму эквивалентную стоимости сданной тары. Это и делает его «реверсивным» торговым автоматом – вместо того, чтобы пользователь вкладывал деньги и получал продукт, пользователь ставит продукт в автомат и получает за это денежное вознаграждение.

Для реализации подобного проекта необходимо «обучить» автомат различать контейнеры. Это может быть сделано по средствам спектрометров, анализаторов, считывателей штрих-кодов. Однако для увеличения прибыли от такого автомата необходимо сделать его максимально дешёвым в плане производства и обслуживания. Поэтому было принято решение попробовать отказаться от уже перечисленных методов и сконцентрироваться на распознавании внешнего вида объекта. С этой целью было решено протестировать ряд открытых нейронных сетей для заданной задачи. В этой статье будут рассмотрены тесты системы распознавания на базе TensorFlow и Keras.

TensorFlow – открытая программная библиотека для машинного обучения, разработанная компанией Google для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и классификации образов, достигая качества человеческого восприятия.

TensorFlow является системой машинного обучения Google Brain второго поколения. В то время как эталонная реализация работает на единичных устройствах, TensorFlow может работать на многих параллельных процессорах (как CPU, так и GPU, опираясь на архитектуру CUDA для поддержки вычислений общего назначения на графических процессорах).

Keras – открытая нейросетевая библиотека, написанная на языке Python. Она представляет собой надстройку над фреймворками DeepLearning4j, TensorFlow и Theano. Прежде всего она нацелена на оперативную работу с сетями глубинного обучения, при этом спроектирована так, чтобы быть компактной, модульной и расширяемой. Она была создана как часть исследовательских усилий проекта ONEIROS (англ. Open-ended Neuro-Electronic Intelligent Robot Operating System). Эта библиотека содержит многочисленные реализации широко применяемых строительных блоков нейронных сетей, таких как слои, целевые и передаточные функции, оптимизаторы, и множество инструментов для упрощения работы с изображениями и текстом.

Основные задачи для разрабатываемой системы: распознавать заданные типы объектов (банки, бутылки и другие предметы) в нормальном и незначительно деформированном состоянии, минимизировать неверное распознавание (что повлечёт за собой дополнительные траты при последующей переработке).

## II. СОЗДАНИЕ СИСТЕМЫ РАСПОЗНАВАНИЯ

Для создания сети были сделаны обучающие выборки банок и бутылок (как целых, так и деформированных), а также предметов отличных от них. Было принято решение создать 2 отличные по реализации сети: 6-ти классовую (банки, бутылки, стеклянные бутылки, огонь, рука человека, различный мусор), способную распознавать не только предметы, но и возможные негативные происшествия, и две 2-ух классовых сети (банки/не банки и бутылки/не бутылки), которые должны работать совместно. Результаты первого варианта обучения на одинаковых выборках представлены в табл. 1.

Для получения данных и последующих результатов использовалась тестовая выборка изображений:

- 5 фотографий бутылок (образцы №1–5);
- 5 фотографий банок (образцы №6–10);

- 5 фотографий иных объектов (образцы №11–15) которые могут быть помещены в отсек распознавания.

ТАБЛИЦА I ПЕРВЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ

Образцы	6cl		can		pet	
	+/-	%	+/-	%	+/-	%
№ 1	1	98	1	93	1	84
№ 2	1	98	1	80	1	92
№ 3	1	93	1	85	1	90
№ 4	1	95	1	77	1	83
№ 5	1	98	1	86	1	87
№ 6	0	93	0	89	0	57
№ 7	0	16	0	94	1	76
№ 8	1	46	0	65	1	60
№ 9	0	65	1	61	1	81
№ 10	0	73	0	81	0	68
№ 11	1	96	1	82	1	98
№ 12	1	99	1	98	1	99
№ 13	1	98	1	89	1	98
№ 14	1	94	1	92	1	98
№ 15	1	99	1	99	1	98
Правильность	73%		73%		87%	
Ложноположительных	4		0		2	
Ложноотрицательных	0		4		0	

В столбцах «+/-» указано правильно (1) или нет (0) распознан объект, а в «%» – процент уверенности сети в своём ответе.

Таким образом, видно, что подготовленные сети будут справлять в среднем одинаково (4 ошибки у сети из 6 классов, 2–4 ошибки у системы из 2-ух классовых сетей).

### III. МОДИФИКАЦИЯ СИСТЕМЫ РАСПОЗНАВАНИЯ

Для улучшения показателей было решено изменить размер скользящего сканирующего окна с 23 пикселей, до 56. Результаты представлены в табл. 2 и 3.

Для 6-ти классовой системы результаты немного улучшились (с 4 до 3 ошибок), однако уверенность сети в своих ответах снизилась.

ТАБЛИЦА II РЕЗУЛЬТАТЫ ПРИ УВЕЛИЧЕНИИ СКОльзяЩЕГО ОКНА для 6-ти классовой сети

Образцы	6cl		6cl_56	
	+/-	%	+/-	%
№ 1	1	98	1	91
№ 2	1	98	1	92
№ 3	1	93	1	85
№ 4	1	95	1	86
№ 5	1	98	1	91
№ 6	0	93	0	82
№ 7	0	16	0	67
№ 8	1	46	1	63
№ 9	0	65	1	69
№ 10	0	73	0	53
№ 11	1	96	1	55
№ 12	1	99	1	83
№ 13	1	98	1	98
№ 14	1	94	1	93
№ 15	1	99	1	98
Правильность	73%		80%	
Ложноположительных	4		3	
Ложноотрицательных	0		0	

ТАБЛИЦА III РЕЗУЛЬТАТЫ ПРИ УВЕЛИЧЕНИИ СКОльзяЩЕГО ОКНА для 2-ух классовых сетей

Образцы	can		can_56		pet		pet_56	
	+/-	%	+/-	%	+/-	%	+/-	%
№ 1	1	93	1	90	1	84	1	91
№ 2	1	80	1	94	1	92	1	76
№ 3	1	85	1	85	1	90	1	69
№ 4	1	77	1	96	1	83	1	67
№ 5	1	86	1	77	1	87	1	99
№ 6	0	89	0	90	0	57	0	70
№ 7	0	94	0	95	1	76	1	92
№ 8	0	65	0	80	1	60	1	64
№ 9	1	61	0	90	1	81	1	88
№ 10	0	81	1	55	0	68	1	58
№ 11	1	82	1	91	1	98	1	98
№ 12	1	98	1	90	1	99	1	99
№ 13	1	89	1	90	1	98	1	98
№ 14	1	92	1	97	1	98	1	98
№ 15	1	99	1	92	1	98	1	98
Правильность	73%		73%		87%		93%	
Ложноположительных	0		0		2		1	
Ложноотрицательных	4		4		0		0	

Для системы из 2-ух классовых сетей результат почти не изменился – от 4 до 1 ошибки (нижний порог уменьшился из-за улучшения качества распознавания бутылок).

Следующим шагом были предприняты попытки повысить результативность, изменяя визуальные качества обучающей выборки: повышение контрастности (can\_c, pet\_c), повышение освещённости (can\_l, pet\_l), повышение чёткости (can\_s, pet\_s). Изображения тестовой выборки изменениям не подвергались. Результаты обучения представлены в табл. 4 и 5.

Воздействия на обучающую выборку, в сети, распознающей банки, показали, что при изменении чёткости изображения результат значительно улучшается. Поскольку основные проблемы на предшествующих этапах были связаны с распознаванием банок, полученная сеть может значительно повысить общую эффективность распознавания системы из двух сетей.

ТАБЛИЦА IV РЕЗУЛЬТАТЫ ПОСЛЕ МОДИФИКАЦИИ ОБУЧАЮЩЕЙ ВЫБОРКИ для сети, РАСПОЗНАЮЩЕЙ БАНКИ

Образцы	can		can_c		can_l		can_s	
	+/-	%	+/-	%	+/-	%	+/-	%
№ 1	1	93	1	92	1	99	1	98
№ 2	1	80	1	99	1	99	1	96
№ 3	1	85	1	97	1	99	1	96
№ 4	1	77	1	98	1	99	1	91
№ 5	1	86	1	61	1	55	1	58
№ 6	0	89	0	92	0	94	0	84
№ 7	0	94	0	99	0	73	0	99
№ 8	0	65	0	59	1	87	1	99
№ 9	1	61	1	86	0	91	1	88
№ 10	0	81	0	83	1	52	1	74
№ 11	1	82	1	94	1	92	1	72
№ 12	1	98	1	98	1	99	1	99
№ 13	1	89	1	97	1	99	1	96
№ 14	1	92	1	99	1	99	1	99
№ 15	1	99	1	99	1	99	1	99
Правильность	73%		73%		80%		87%	
Ложноположительных	0		0		0		0	
Ложноотрицательных	4		4		3		2	

ТАБЛИЦА V РЕЗУЛЬТАТЫ ПОСЛЕ МОДИФИКАЦИИ ОБУЧАЮЩЕЙ ВЫБОРКИ ДЛЯ СЕТИ, РАСПОЗНАЮЩЕЙ БУТЫЛКИ

Образцы	pet		pet_c		pet_l		pet_s	
	+/-	%	+/-	%	+/-	%	+/-	%
№ 1	1	84	0	61	0	94	0	74
№ 2	1	92	0	57	1	54	0	60
№ 3	1	90	0	77	0	97	0	51
№ 4	1	83	0	75	0	92	0	58
№ 5	1	87	1	99	1	99	1	87
№ 6	0	57	1	79	1	94	0	56
№ 7	1	76	1	96	1	86	1	96
№ 8	1	60	1	64	0	90	0	52
№ 9	1	81	1	87	1	98	1	52
№ 10	0	68	1	63	1	84	0	60
№ 11	1	98	1	95	1	99	1	96
№ 12	1	99	1	99	1	99	1	99
№ 13	1	98	1	97	1	99	1	99
№ 14	1	98	1	98	1	99	1	98
№ 15	1	98	1	90	1	99	1	98
Правильность	87%		73%		73%		53%	
Ложноположительных	2		0		1		3	
Ложноотрицательных	0		4		3		4	

Воздействия на обучающую выборку, в сети, распознающей бутылки, ухудшили результаты. Аналогичный тест с 6-ти классовой сетью тоже показал снижение общего процента правильности распознавания.

Таким образом, лучшей 6-ти классовой сетью является bcl\_56 (3 ошибки). Для определения лучшей системы из 2-ух классовых сетей необходимо учитывать комбинации ошибок, так как данные сети будут работать одновременно. Поэтому лучше всего использовать сети can\_s и pet\_56.

#### IV. ВЫВОДЫ

В ходе экспериментов было выявлено, что для выбранной задачи система из 2-ух классовых сетей суммарно будет ошибаться реже (ошибка в 6, возможные ошибки в 7 и 10 образце), чем из шести (ошибки в 6, 7 и 10 образцах).

Подробные результаты представлены в табл. 6.

ТАБЛИЦА VI РЕЗУЛЬТАТЫ ВЫБРАННЫХ ДЛЯ РЕАЛИЗАЦИИ СЕТЕЙ

Образцы	can_s		pet_56	
	+/-	%	+/-	%
№ 1	1	98	1	91
№ 2	1	96	1	76
№ 3	1	96	1	69
№ 4	1	91	1	67
№ 5	1	58	1	99
№ 6	0	84	0	70
№ 7	0	99	1	92
№ 8	1	99	1	64
№ 9	1	88	1	88
№ 10	1	74	1	58
№ 11	1	72	1	98
№ 12	1	99	1	99
№ 13	1	96	1	98
№ 14	1	99	1	98
№ 15	1	99	1	98
Правильность	87%		93%	
Ложноположительных	0		1	
Ложноотрицательных	2		0	

При реализации системы из 6-ти классовой сети, автомат будет достаточно плохо собирать бутылки, периодически добавляя в них банки. Это негативно скажется на последствиях переработки данных ёмкостей.

Если сомнительные объекты (одна из сетей проголосовала за, а другая против) будут автоматически считаться посторонними (не бутылка и не банка), то это позволит автомату повысить единообразность собранных отходов, не допуская посторонних включений, что является одной из основных преследуемых целей.

В целом, на данный момент, сети лучше работают с бутылками, чем с банками. Этот факт вызван, прежде всего, особенностями внешнего вида изучаемых объектов. Изображения банок достаточно часто совпадают с изображениями коробок и прямоугольных упаковок. Отказаться от таких образов в отрицательной обучающей выборке не является хорошей идеей, так как тогда в собранный объём банок будут попадать посторонний мусор. В дальнейшем мы собираемся решить данный вопрос увеличением обучающей выборки за счёт увеличения числа изображений банок.

#### СПИСОК ЛИТЕРАТУРЫ

- [1] Reverse vending 101: a beginner's guide [Электронный ресурс]. – 2011. – Режим доступа: <https://www.tomra.com/en/collection/reverse-vending/reverse-vending-news/2017/how-does-a-reverse-vending-machine-work/>. – Дата обращения 30 декабря 2017.
- [2] Cost of a Deposit Return Reverse Vending Machine. [Электронный ресурс]. – 2013. – Режим доступа: <http://www.zerowastescotland.org.uk/sites/default/files/reverse%20v%20-%20CFE%20response.pdf>. – Дата обращения 30 декабря 2017.
- [3] Keras: The Python Deep Learning library [Электронный ресурс]. – Keras documentation. – Режим доступа: <https://keras.io/#why-this-name-keras>. – Дата обращения 10 января 2018).
- [4] Хижняков. Ю.Н. Современные проблемы теории управления: учеб. пособие. Изд-во Перм. нац. исслед. политехн. ун-та. Пермь, 2015. 237 с. ISBN 978-5-398-01500-3.
- [5] Слива М.В. Использование миникомпьютера Raspberry PI для преподавания основ робототехники // Культура, наука, образование: проблемы и перспективы: материалы III Всероссийской научно-практической конференции / Нижневартковский государственный университет – Нижневартковск, 2014. С. 326-328.
- [6] Линев Ф.А. Интернет вещей на основе Raspberry PI / Линев Ф.А., Киселева С.Д. // Молодежный научно-технический вестник. 2015. №5. С. 22.