

ТЕХНИЧЕСКИЕ НАУКИ

УДК 004.93

АВТОМАТИЗАЦИЯ ПРОЦЕССА РАСПОЗНАВАНИЯ ДОКУМЕНТОВ С ИСПОЛЬЗОВАНИЕМ СВЕРТОЧНЫХ НЕЙРОННЫХ СЕТЕЙ

Вагнер Александр Валерьевич, магистрант, направление подготовки 09.04.04 Программная инженерия, Оренбургский государственный университет, Оренбург
e-mail: alexmednogorsk@mail.ru

Научный руководитель: **Горбачев Дмитрий Владимирович**, кандидат технических наук, доцент, доцент кафедры программного обеспечения вычислительной техники и автоматизированных систем, Оренбургский государственный университет, Оренбург
e-mail: gordi47@mail.ru

***Аннотация.** В данной статье рассматривается тема автоматизации процесса распознавания документов с использованием сверточных нейронных сетей. Это достаточно сложная задача, но ее актуальность заключается в том, что на данный момент обмен документами в основном происходит в электронном виде, вследствие чего возрастает объем документов и возникает потребность в более быстрой и тщательной проверке документов.*

Целью данной работы является разработка сверточной нейронной сети, которая будет распознавать, какое фото было загружено: паспорт или другой документ. Имея на руках такую систему, уже можно автоматизировать небольшую часть работы с документами.

Представленная работа реализовывалась в несколько этапов:

- 1) анализ необходимой научной литературы;*
- 2) проектирование и обучение сверточной нейронной сети на наборе, состоящем из 2000 изображений;*
- 3) апробирование сверточной нейронной сети на тестовом наборе данных из 200 изображений.*

В результате апробации было показано, что данная нейронная сеть способна распознавать изображение паспорта среди любых случайных изображений с точностью в 95%.

Спроектированная нейронная сеть имеет практическую значимость, так как такая нейронная сеть способна оптимизировать деятельность абсолютно любой организации. Достаточно лишь обучить нейросеть на том наборе данных, который используется на предприятии.

Кроме того, данная нейронная сеть может использоваться для распознавания не только изображений документов, но и любых других изображений, например, 3D-моделей, различных объектов и приборов, используемых на предприятии.

В дальнейшем планируется доработать модель сверточной нейронной сети, чтобы такая сеть могла распознавать текст внутри документа, например, срок действия документа, и определять – не просрочен ли документ.

***Ключевые слова:** автоматизация, сверточные нейронные сети, документы, обучение, распознавание, Python, TensorFlow, Keras, OpenCV.*

***Для цитирования:** Вагнер А. В. Автоматизация процесса распознавания документов с использованием сверточных нейронных сетей // Шаг в науку. – 2021. – № 4. – С. 34–39.*

AUTOMATION OF THE DOCUMENT RECOGNITION PROCESS USING CONVOLUTIONAL NEURAL NETWORKS

Vagner Alexander Valerievich, postgraduate student, training program 09.04.04 Software Engineering, Orenburg State University, Orenburg
e-mail: alexmednogorsk@mail.ru

Research advisor: **Gorbachev Dmitry Vladimirovich**, Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Computer Engineering and Automated Systems Software, Orenburg State University, Orenburg
e-mail: gordi47@mail.ru

Abstract. This article discusses the topic of automating the document recognition process using convolutional neural networks. This is a rather difficult task, but its relevance lies in the fact that at the moment the exchange of documents mainly takes place in electronic form, as a result of which the volume of documents increases and there is a need for a faster and more thorough verification of documents.

The goal of this work is to develop a convolutional neural network that will recognize which photo has been uploaded: a passport or another document. Having such a system in hand, it is already possible to automate a small part of work with documents.

The presented work was carried out in several stages:

- 1) analysis of the required scientific literature;
- 2) designing and training a convolutional neural network on a set of 2000 images;
- 3) testing a convolutional neural network on a test dataset of 200 images.

As a result of approbation, it was shown that this neural network is able to recognize a passport image among any random images with an accuracy of 95%.

The designed neural network is of practical importance, since such a neural network is able to optimize the activities of absolutely any organization. It is enough just to train the neural network on the dataset that is used in the enterprise.

In addition, this neural network can be used to recognize not only document images, but also any other images, for example, 3D models, various objects and devices used in the enterprise.

In the future, it is planned to refine the model of a convolutional neural network so that such a network can recognize text within a document, for example, the validity period of a document, and determine if the document has expired.

Key words: automation, convolutional neural networks, documents, training, recognition, Python, TensorFlow, Keras, OpenCV.

Cite as: Vagner, A. V. (2021) [Automation of the document recognition process using convolutional neural networks]. *Shag v nauku* [Step into science]. Vol. 4, pp. 34–39.

На сегодняшний день нейронные сети имеют очень большую популярность в различных сферах деятельности. Например, в медицине с помощью нейронных сетей реализовывают распознавание различных болезней на снимках¹. В экономике нейронные сети позволяют максимально быстро и качественно обрабатывать огромные массивы данных и выдавать различные прогнозы². Также нейронные сети можно встретить в современных автомобилях, в которых есть система автопилота³. Еще нейронные сети могут распознавать объекты, людей, текст. В робототехнике нейронные сети применяются для управления роботами⁴.

На самом деле, нейронные сети можно применить где угодно. Главное, правильно подобрать алгоритм решения задачи, правильно спроектировать нейронную сеть, должным образом обучить, и тогда нейронная сеть будет выдавать требуемый результат [1].

В данной статье рассматривается такая тема, как автоматизация процесса распознавания документов с использованием сверточных нейронных сетей. Иначе говоря, такой процесс называется «распознавание образов». Это достаточно сложная задача, но

в то же время очень актуальная, так как технологии развиваются очень быстро, и поэтому решение определенных задач и принятие определенных решений требуют особой точности, эффективности и достоверности. Поэтому здесь и применяются нейронные сети, в том числе задача распознавания образов.

В настоящее время есть очень важная задача – идентификация документов при оформлении пропусков на охраняемые объекты. В охранной организации есть веб-сервис, который позволяет оформлять заявки на получение пропусков на охраняемые объекты. Со временем число таких объектов стало расти, и соответственно, пропусков стало больше. Работа у специалиста, который проверяет заявки, существенно возросла. Поэтому он может не увидеть, например, просрочен ли документ. И сделать ошибку, пропустив такую заявку. А внедрение интеллектуальной системы, которая бы сама проверяла все документы – существенно упростило бы работу специалиста и ускорило бы процесс выдачи пропусков.

Решением проблемы распознавания формы и содержания документа является проектирование сверточной нейронной сети [2]. Обычные класси-

¹ Нейросети в медицине: что это и как работает? Искусственный интеллект для анализа медицинских изображений [Электронный ресурс]. – Режим доступа: <https://celsus.ai/blog/kak-rabotaet-medicinskaya-nejroset/>. (дата обращения: 25.05.2021).

² Использование нейронных сетей для прогнозирования инфляции: новые возможности [Электронный ресурс]. – Режим доступа: https://elar.urfu.ru/bitstream/10995/66452/1/vesnik_2018_5_007_.pdf (дата обращения: 25.05.2021).

³ Как Tesla обучает автопилот [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/450796/>. (дата обращения: 25.05.2021).

⁴ Растущие нейронные сети в робототехнике [Электронный ресурс]. – Режим доступа: https://www.researchgate.net/publication/330753974_RASTUSIE_NEJRONNYE_SETI_V_ROBOTOTEHNIKE. (дата обращения: 25.05.2021).

ческие нейронные сети здесь не подходят, так как они малоэффективны при работе с изображениями.

Сверточная нейронная сеть – это специальная архитектура искусственных нейронных сетей, нацеленная на эффективное распознавание образов [3]. Идея сверточных нейронных сетей заключается в чередовании сверточных слоёв и субдискретизирующих слоёв. Структура сети – однонаправленная, то есть без обратных связей. В основном сверточные нейронные сети являются многослойными. Для обучения используются стандартные методы, чаще всего метод обратного распространения ошибки. Функция активации нейронов может быть любая, по выбору исследователя.

Рассмотрим типовую структуру сверточной нейронной сети более подробно. Сеть состоит из большого количества слоёв [4]. После начального слоя (входного изображения) сигнал проходит серию сверточных слоёв, в которых чередуется собственно свёртка и субдискретизация (пулинг). Чередование слоёв позволяет составлять «карты признаков» из карт признаков, на каждом последующем слое карта уменьшается в размере, но увеличивается количество каналов. На практике это означает способность распознавания сложных иерархий признаков. Обычно после прохождения нескольких слоёв карта признаков вырождается в вектор или даже скаляр, но таких карт признаков становятся сотни. На выходе сверточных слоёв сети дополнительно устанавливаются несколько слоёв полносвязной нейронной сети (перцептрон), на вход которому подаются оконечные карты признаков [5].

Для реализации сверточной нейронной сети по распознаванию изображений был выбран язык программирования Python версии 3.8.6 64bit. Данная версия языка выбрана потому, что на последней версии Python 3.9.0 нет поддержки библиотеки tensorflow. А сам язык Python был выбран потому, что как раз именно этот язык и предназначен для реализации различных нейронных сетей. Python уже содержит большое количество библиотек по проектированию нейронных сетей, по работе с изображениями, массивами данных и графиками.

Для проектирования нейронной сети была выбрана библиотека 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. Также для работы с нейронной сетью была выбрана библиотека Keras – открытая нейросетевая библиотека, написанная на языке Python, которая в данном случае используется как надстройка над библиотекой TensorFlow.

Для работы с изображениями была выбрана би-

блиотека OpenCV, которая позволяет максимально быстро и просто работать с изображениями, используя всего лишь несколько строк кода на языке Python.

Для работы с вычислениями используется библиотека NumPy – это библиотека с открытым исходным кодом для языка программирования Python, которая предоставляет поддержку многомерных массивов и поддержку высокоуровневых математических функций.

Для работы с графиками была выбрана библиотека Matplotlib – это библиотека на языке программирования Python для визуализации данных двумерной графикой.

Цель – необходимо спроектировать сверточную нейронную сеть, которая будет распознавать, какое фото было загружено: паспорт или другой документ. Поэтому имеем два класса: паспорт и не паспорт. Эти два класса будут использоваться далее при проектировании сверточной нейронной сети.

Как известно, чем больше исходных данных, тем лучше нейронная сеть будет обучена⁵ [6]. Поэтому, в качестве обучающих данных было подготовлено 2000 изображений, среди которых 1000 изображений – фотографии паспортов, а остальные 1000 изображений – случайные изображения.

Обучив такую нейронную сеть на двух классах, уже можно автоматизировать небольшую часть процесса распознавания документов. Если продолжить обучение нейронной сети на остальных видах документов, например, справка об электробезопасности, или справка о работах на высоте, то такая нейросеть будет автоматически распознавать тип документа и предупреждать пользователя и проверяющего о том что был загружен неправильный тип документа.

Для решения поставленной задачи были пройдены следующие этапы:

- были изучены теоретические сведения о сверточных нейронных сетях;
- рассмотрены преимущества и недостатки таких сетей;
- были выбраны технологии для реализации сверточной нейронной сети;
- был собран датасет из 2000 обучающих изображений и из 200 тестовых изображений;
- была спроектирована архитектура сверточной нейронной сети;
- далее данная нейронная сеть была обучена методом с учителем (на маркированных данных);
- было проведено тестирование сверточной нейронной сети на тестовом наборе данных из 200 изображений;
- был получен результат работы сверточной нейронной сети по распознаванию изображения паспорта.

⁵ Подборка датасетов для машинного обучения [Электронный ресурс]. – Режим доступа: <https://vc.ru/dev/68180-podborka-datasetov-dlya-mashinnogo-obucheniya>. (дата обращения: 25.05.2021).

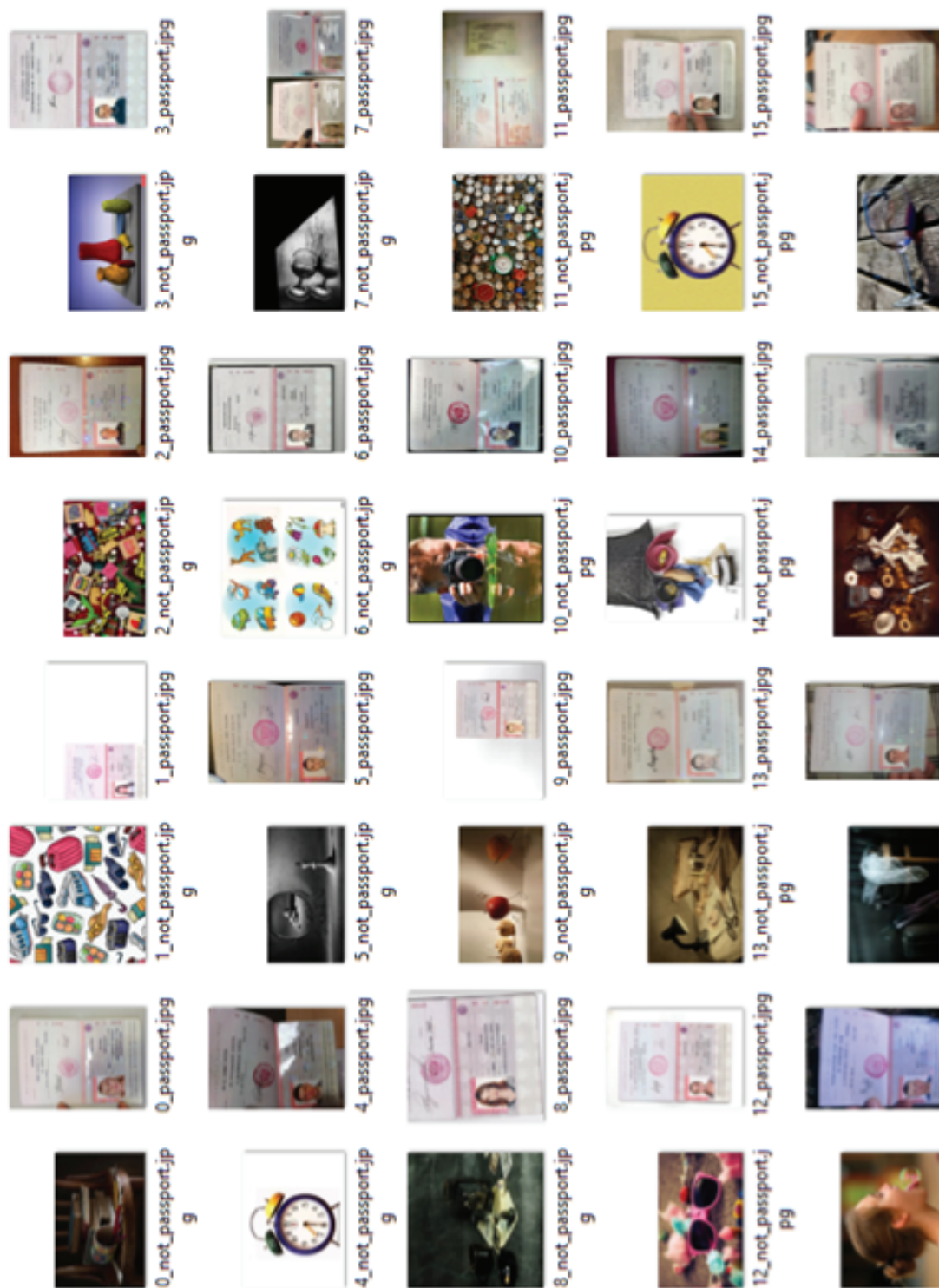


Рисунок 1. Набор данных для обучения сверточной нейронной сети
Источник: разработано автором



Рисунок 2. Результат работы сверточной нейронной сети по распознаванию вида загруженного документа

Источник: разработано автором

Точность обучения сверточной нейронной сети составила 0,9458 или 95%, что является отличным результатом [7]. Данная нейронная сеть способна отделять изображение паспорта среди любых случайных изображений.

Помимо определения вида документа, нейронные сети способны ещё и распознавать текст на документах. Это можно применить, например, при распознавании срока действия документа и определяя – не просрочен ли документ.

Литература

1. Гудфеллоу Я., Бенджио И., Курвилль А. Г93 Глубокое обучение / пер. с англ. А. А. Слинкина. – 2-е изд., испр. – М.: ДМК Пресс, 2018. – 652 с.
2. Рашид, Тарик. С58 Создаем нейронную сеть. : пер. с англ. — СПб.: ООО «Альфа-книга», 2017. – 272 с.
3. Сикорский О. С. Обзор свёрточных нейронных сетей для задачи классификации изображений // Новые информационные технологии в автоматизированных системах. 2017. № 20. [Электронный ре-

сурс]. – URL: <https://cyberleninka.ru/article/n/obzor-svyortochnyh-neyronnyh-setey-dlya-zadachi-klassifikatsii-izobrazheniy> (дата обращения: 24.05.2021).

4. Суровцев И. С., Клюкин В. И., Пивоварова Р. П. Нейронные сети. – Воронеж: ВГУ, 1994. – 224 с.
5. Федотов Д. В., Попов Е. А., Охорзин В. А. Оптимизация структуры сверточной нейронной сети с помощью самоконфигурируемого эволюционного алгоритма в одной задаче идентификации // Вестник СибГАУ. 2015. Т. 16, № 4. С. 857–863.
6. LeCun Y., Bengio Y. Convolutional Networks for Images, Speech, and Time-Series // The Handbook of Brain Theory and Neural Networks. MIT Press, 1995. P.
7. Rishal Hurbans. Grokking Artificial Intelligence Algorithms // Manning Publications Co. 2020.

Статья поступила в редакцию: 27.05.2021; принята в печать: 08.11.2021.

Автор прочитал и одобрил окончательный вариант рукописи.