

# ВЛИЯНИЕ ВЫБОРА КЛЮЧЕВЫХ СЛОВ НА КАЧЕСТВО ИХ РАСПОЗНАВАНИЯ

**Аннотация.** В статье рассматривается влияние выбора ключевых слов на качество их распознавания с использованием сверточных нейронных сетей (CNN). Для решения этой задачи были применены три модели CNN: Xception, ResNet50 и MobileNetV2. Основная цель работы заключалась в сравнении эффективности этих моделей в зависимости от выбранного ключевого слова, а также в определении ключевых слов, которые наиболее успешно поддаются распознаванию. В процессе исследования были использованы современные программные инструменты, такие как библиотека Keras, для построения и обучения моделей. Полученные результаты демонстрируют, что выбор ключевых слов и архитектура используемой модели оказывают существенное влияние на точность распознавания.

**Ключевые слова:** выбор ключевых слов, распознавание образов, машинное обучение, сверточные нейронные сети, классификация изображений, библиотека Keras.

## THE IMPACT OF KEYWORD SELECTION ON RECOGNITION QUALITY

MALINSKIY S.V., GUBIN M.V., ALADIN P.A., GALENDA S.V.

**Abstract.** This paper examines the impact of keyword selection on recognition accuracy using convolutional neural networks (CNN). To address this task, three CNN models were applied: Xception, ResNet50, and MobileNetV2. The main objective of the study was to compare the effectiveness of these models depending on the selected keyword, as well as to identify the keywords that are most successfully recognized. Modern software tools, such as the Keras library, were used for model construction and training. The results obtained demonstrate that both the choice of keywords and the architecture of the model have a significant impact on recognition accuracy.

**Keywords:** keyword selection, image recognition, machine learning, convolutional neural networks, image classification, Keras library.

Обучение нейронных сетей для распознавания ключевых слов как альтернативы подписи представляет собой актуальную задачу, особенно в контексте защиты персональных данных. С развитием технологий глубокого обучения и компьютерного зрения появились новые возможности для создания автоматизированных систем, способных с высокой точностью анализировать и различать почерки людей, а также идентифицировать их авторов [1].

Эти системы могут быть полезны при работе с различными документами, где требуется автоматическая идентификация на основе ключевых слов, написанных от руки. Они также могут применяться в пунктах контроля доступа — например, при въезде на территорию предприятия или для доступа в ограниченные зоны [2].

Использование ключевых слов вместо традиционной подписи — это относительно новая практика, и на данный момент аналогичных коммерческих решений нет. Теория показывает, что такой метод может повысить уровень безопасности и удобства аутентификации, сохраняя при этом конфиденциальность персональных данных.

В настоящей работе было исследовано применение нейросетей и замена подписи на ключевые слова. Человек записывает ключевое слово, и на основе анализа почерка и правильности написанного слова происходит его идентификация [3]. Такой подход добавит дополнительный уровень защиты и повысит точность аутентификации.

Для исследования задачи верификации ключевых слов были выбраны три модели нейронных сетей: MobileNetV2, Xception и ResNet50 [4].

Эти модели были выбраны благодаря своими архитектурными особенностями и потенциалу в задачах распознавания и классификации изображений [5].

MobileNetV2 — легковесная и эффективная модель, разработанная для использования на мобильных и встроженных устройствах. Архитектура MobileNetV2 содержит начальный уровень полной свертки с 32 фильтрами, за которым следуют 19 остаточных слоев [6].

Она использует глубинно-разделяемые свертки, что позволяет значительно сократить количество параметров и вычислительных затрат по сравнению с традиционными сверточными сетями.

Модель Xception (англ. Xception – eXtreme Inception) была предложена в 2016 году как развитие модели Inception. Она реализует идеи, подобные Inception, но считается более оптимальной. Xception использует концепцию (глубинно-сепарабельных) сверток, что позволяет значительно улучшить эффективность и производительность сети. Модель включает 36 сверточных слоев, структурированных в 14 модулей, что обеспечивает мощные возможности для извлечения признаков из изображений [7].

Модель ResNet50 состоит из 50 слоев и использует остаточные блоки (residual blocks), которые позволяют градиентам более эффективно распространяться через сеть. Эта архитектура решает проблему исчезающего градиента, что делает обучение глубоких нейронных сетей более стабильным. ResNet50 была предложена в 2015 году в работе "Deep Residual Learning for Image Recognition" (Kaiming He) и с тех пор стала одной из самых популярных моделей для задач компьютерного зрения [8].

Для обучения моделей были собраны 2250 изображений (9 различных ключевых слов) от 5 человек (3 мужчин и 2 женщин). 1575 изображений составили обучающую выборку. Оставшиеся 675 изображений составили проверяющую (тестовую) выборку.

Разнообразие почерков, различный уровень освещенности при сканировании и различное качество изображений обеспечивают множество информативных примеров для обучения нейронной сети и способствуют её адаптации к различным условиям.

Пример выбранных ключевых слов представлен на рисунке 1.

- 1 Анализ 4 Мониторинг 7 Ревизия  
2 Сигнал 5 Книга 8 Аудит  
3 Проверка 6 Контроль 9 Испекция

Рисунок 1 – Пример ключевых слов

Для анализа результатов было построено отображение, позволяющее проследить процесс обучения для лучших по точности распознавания ключевых слов (Аудит, Контроль, Ревизия) и сравнить с худшими по точности (Сигнал, Книга, Мониторинг) на обучающей выборке. Модель – Xception (рис. 2).

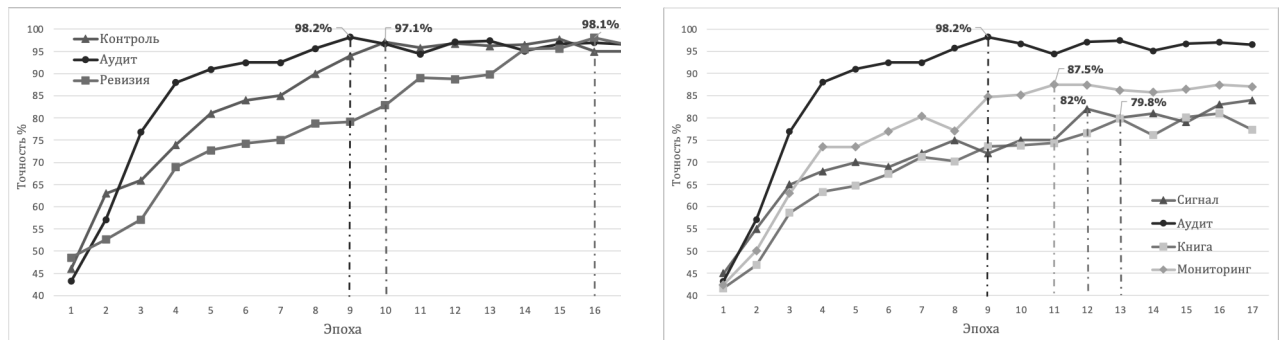


Рисунок 2 – Точность распознавания на обучающей выборке

Результаты обучения для ключевых слов (Аудит - А, контроль - К, ревизия - Р) отображены в таблице, которая демонстрирует точность распознавания для каждой протестированной модели (табл.).

Таблица – Оценки качества распознавания ключевых слов

Модель	Эпох	Время обучения, (с)	Качество распознавания (обучение, %)			Качество распознавания (тестирование, %)			Качество распознавания (среднее значение, %)	Среднее время распознавания (мс)
			А	К	Р	А	К	Р		
Xception	20	92	98,2	97,1	98,1	96	91,4	91,6	95,4	19
ResNet50	33	110	97,1	92,8	97,1	93	90,1	87,4	92,8	17
MobileNetV2	24	74	96,9	89,9	92,5	91	88,3	86,4	91,42	16,8

Полученные результаты наглядно показывают, что лучшим (из трёх выбранных ключевых слов) является – “Аудит”. Это ключевое слово показало высокие результаты обучения и тестирования соответственно: Xception – (98,2%; 96%), ResNet50 – (97,1%; 93%), MobileNetV2 – (97,1%; 91%).

Исследования позволяют считать его оптимальным для задачи идентификации человека на основе рукописного ключевого слова (взамен подписи).

Однако для обеспечения надежности системы распознавания необходимо каждый раз проводить проверку на новом коллективе. Это позволит определить ключевые слова, которые обеспечат наилучшую распознаваемость и точность идентификации в условиях данной группы пользователей.

Остальные ключевые слова (контроль , ревизия) продемонстрировали стабильные результаты распознавания, однако средняя точность “контроль” ниже на 3,5%, “ревизия” – ниже на 4,9% (по сравнению с “Аудит” на тестовой выборке).

Модель Xception показала наивысшие результаты распознавания для всех ключевых слов (в среднем 97,8%). На основании этого выбора и высоких показателей модели Xception было принято решение использовать именно её для создания интеллектуальной системы автоматической идентификации.

На основе ключевого слова – “Аудит” и модели – Xception была разработана интеллектуальная система с использованием языка программирования Python [9], а графический интерфейс был реализован с помощью библиотеки PyQt [10].

Пример распознавания ключевого слова разработанной интеллектуальной системой представлен на рис. 3.

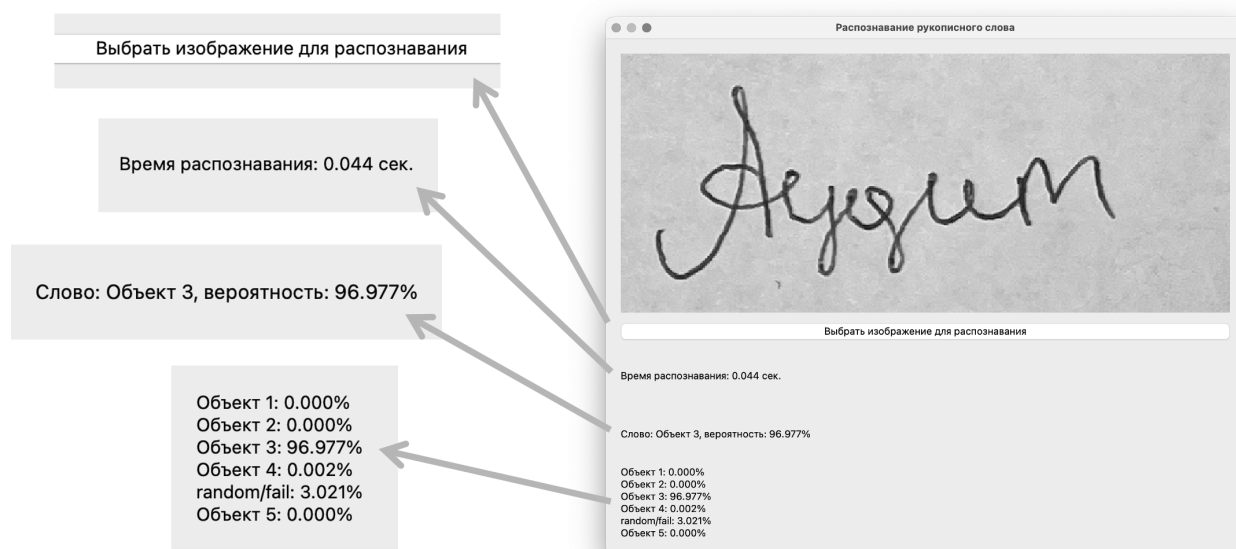


Рисунок 3 – Пример распознавания автора ключевого слова интеллектуальной системой

## Литература

1. Искусственный интеллект, нейронные сети и распознавание рукописного текста [сайт]. URL: <https://www.myscript.com/ru/ai> (дата обращения: 05.10.2024).
2. Вьюгин В. В. Математические основы машинного обучения и прогнозирования. - М.: МЦНМО, 2022 г. - 400 с
3. Скрыпников А.В., Денисенко В.В., Хитров Е.Г., Евтеева К.С., Савченко И.И. Распознавание рукописного текста с использованием нейронных сетей // Современные наукоемкие технологии. – 2021. – № 6–1. – С. 91–95.
4. Жерон О. Прикладное машинное обучение с помощью Scikit-Learn, Keras и TensorFlow. Концепции, инструменты. - М.: Диалектика, 2020 г. - 1040 с.
5. Рамсундар Б., Заде Р. TensorFlow для глубокого обучения– М.: BHV, 2019 г. – 256 с.
6. MobileNet, MobileNetV2, and MobileNetV3 [сайт]. URL: <https://keras.io/api/applications/mobilenet/> (дата обращения: 08.11.2023)
7. Xception: компактная глубокая нейронная сеть [сайт]. URL: [https://ai-news.ru/2018/01/xception\\_kompaktnaya\\_glubokaya\\_nejronnaya\\_set.html](https://ai-news.ru/2018/01/xception_kompaktnaya_glubokaya_nejronnaya_set.html) (дата обращения: 02.10.2024)
8. ResNet 50 – Keras Applications [сайт]. URL: <https://keras.io/api/applications/resnet/> (дата обращения: 02.10.2024).
9. Пол Бэрри. Изучаем программирование на Python. - М.: Эксмо, 2022 г. - 624 с.
10. ШлееМ. Ш68. Qt 5.10. Профессиональное программирование на Python. - СПб.: БХВ-Петербург, 2018. - 1072 с.