

# Российские учёные научили нейросеть «читать» рукописные буквы русского алфавита

С развитием it-технологий растёт важность быстрого и качественного преобразования рукописного текста в цифровую печатную версию, чтобы было удобнее копировать, редактировать или извлекать из него данные. Очевидно, что первой ступенью в этом процессе станет распознавание букв русского алфавита, написанных от руки. Учёные СФУ разработали новую свёрточную нейросеть (CNN), способную с высокой точностью распознавать изображения рукописных букв. Полученный алгоритм преобразует изображение и «узнаёт» букву, зашифрованную в нём. По словам учёных, точность классификации составляет 99 %.



На сегодняшний день 2,4 % населения в мире говорят по-русски. Сложность распознавания текста, написанного кириллицей от руки, достаточно велика — в особенности, для людей, не знакомых с русским алфавитом. В Интернете распространены сервисы, которые можно использовать для распознавания и преобразования текста любого типа, как цифрового, так и рукописного. Однако использование таких сервисов чревато утечками информации и ненадёжно с точки зрения конфиденциальности и безопасности пользователя. Приложение, которое сможет легко и быстро распознать текст на кириллице, работающее на стороне клиента и не требующее подключения к Интернету, может быть востребовано как у отдельных юзеров, так и у организаций.

*«Самой, пожалуй, интересной особенностью рукописного русского текста является индивидуальный стиль написания букв — то, что мы называем почерком. Стили написания имеют свойство меняться со временем, достаточно сравнить каллиграфические строчки в тетрадях-прописях поколения 70—80-х годов и то, как пишут современные школьники. Даже у одного человека почерк меняется в течение жизни. Целью нашего исследования было распознавание рукописного текста на русском языке нейросетью с использованием моделей глубокого обучения (DL). Насколько нам известно, это первая в мире работа такого рода», — отметил соавтор исследования, студент Института информационных и космических технологий СФУ*  
**Андрей Левков.**



Для достижения цели учёные предприняли ряд шагов. Построили новый набор данных с помеченным изображением в разрешении 32×32 пикселя для 33 букв русского алфавита. Разработали новую архитектуру CNN для проблемы обнаружения рукописных букв русского алфавита и сравнили её с уже существующими мощными моделями CNN. Кроме того, красноярские и петербургские эксперты представили полное описание используемой свёрточной нейросети и исходного кода, чтобы другие исследователи могли воспроизвести эти данные для обнаружения рукописных букв русского алфавита. Для программирования был выбран язык Python и интерактивная среда разработки Jupyter.

Обучение нейросети проводилось с помощью предварительно обработанных данных хранилища CoMNIST — это известная база данных, содержащая образцы рукописного написания букв на латинице и кириллице. Набор данных в базе состоит из 4-полосных изображений с разрешением 278×278 пикселей в формате .png.

«Набор данных содержит 13 299 фотографий, каждая из которых находится в отдельной папке. Папки, в свою очередь, относятся к определенному классу. Таких классов в наборе 33, и каждый соответствует букве русского алфавита. Для каждого класса существует от 300 до 500 изображений. На этих изображениях зафиксированы прописные, печатные и написанные курсивом буквы. И вот приблизительно на 85% этих снимков нейронная сеть (CNN) училась распознавать буквы русского алфавита, а ещё на 15% шла проверка усвоенных «знаний», — рассказала руководитель исследования, доцент кафедры систем искусственного интеллекта СФУ **Анастасия Сафонова**.



Созданный учёными новый уникальный набор данных (изображений) нужен был, чтобы провести независимую проверку разработанной модели. На каждой фотографии была представлена только одна буква в печатном или письменном виде. Набор содержит от 5 до 10 изображений для каждого класса. Чтобы увеличить набор данных и их вариативность, учёные применяли разные функции преобразования изображений — вращали изображения вправо и влево, применяли распределение Гаусса и т. д. В результате эксперты получили 79 794 изображений, на 67 825 из которых нейросеть могла учиться, а 13 084 служили для проверки.

*«Мы сравнили разработанную нашим коллективом модель с наиболее мощными моделями CNN, например, с VGG-16, VGG-19, Xcept, Resnet-101, Mobilenet-V2 и другими. Оказалось, что точность нашей модели во время обучения составляла до 99 %, всё обучение заняло 3 часа. Точность прогнозирования модели составила до 95,83 %. В целом наша модель уступила лишь одной альтернативной — VGG-16, которая продемонстрировала до 99 % точности, самая низкая точность свойственна моделям Xception и Inception-V3», — подчеркнула Анастасия Сафонова.*

Подводя итоги, учёные сообщили, что созданная ими модель нейросети не является окончательной и может быть улучшена в будущем — вероятно, её архитектура будет меняться, чтобы увеличить точность классификации. Также эксперты планируют обучать свою модель для распознавания русского рукописного текста на новом наборе данных и знакомить её с различными стилями письма.

Сообщается, что на основании проделанной работы была зарегистрирована уникальная программа ЭВМ, правообладателем которой выступил Сибирский федеральный университет.

[Пресс-служба СФУ](#), 22 июля 2022 г.

© Сибирский федеральный университет. Редакция сайта: +7 (391) 246-98-60, info@sfu-kras.ru.

Адрес страницы: <https://news.sfu-kras.ru/node/26580>