

# Методика предобработки данных машинного обучения для решения задач компьютерного зрения

А. Е. Трубин<sup>1\*</sup>, А. А. Морозов<sup>2</sup>, А. Е. Зубанова<sup>2</sup>, В. А. Ожередов<sup>1,3</sup>, В. С. Корепанова<sup>1,4</sup>

<sup>1</sup> Университет «Синергия», Москва, Россия

<sup>2</sup> Орловский государственный университет им. И. С. Тургенева, Орел, Россия

<sup>3</sup> Институт космических исследований Российской академии наук, Москва, Россия

<sup>4</sup> ООО «ЛУКОЙЛ-Инжиниринг», Москва, Россия

\* niburt@yandex.ru

**Аннотация.** В сфере машинного обучения не существует единой методологии предобработки данных, так как все этапы этого процесса являются уникальными, под конкретную задачу. Однако в каждом направлении используется определенный тип данных. В гипотезе исследования предполагается, что можно четко структурировать последовательности и фазы подготовки данных для задач распознавания текстов. В статье рассмотрены основные принципы предобработки данных и выделение последовательных этапов как конкретной методики для задачи распознавания символов азбук. В качестве исходных данных были выбраны изображения набора ETL. Предобработка включала в себя этапы работы с изображениями, на каждом из которых в исходные данные вносились изменения. Первым шагом являлось кадрирование, которое позволило избавиться от лишней информации на изображении. Далее был рассмотрен подход преобразования изображения к исходному соотношению сторон и определен метод преобразования из оттенков серого в черно-белый формат. На следующем этапе были искусственно расширены линии символов для лучшего распознавания печатных азбук. На последнем этапе предобработки данных была произведена аугментация, которая позволила лучше распознавать символы азбук независимо от их положения в пространстве. Как результат, была выстроена общая структура методики предобработки данных для задач распознавания текстов.

**Ключевые слова:** нейронные сети, сверточная нейронная сеть, предобработка данных, компьютерное зрение, машинное обучение

**Для цитирования:** Трубин А. Е., Морозов А. А., Зубанова А. Е., Ожередов В. А., Корепанова В. С. Методика предобработки данных машинного обучения для решения задач компьютерного зрения // Прикладная информатика. 2022. Т. 17. № 4. С. 47–56. DOI: 10.37791/2687-0649-2022-17-4-47-56

# The method of preprocessing machine learning data for solving computer vision problems

A. Trubin<sup>1\*</sup>, A. Morozov<sup>2</sup>, A. Zubanova<sup>2</sup>, V. Ozheredov<sup>1,3</sup>, V. Korepanova<sup>1,4</sup>

<sup>1</sup> Synergy University, Moscow, Russia

<sup>2</sup> Orel State University named after I. S. Turgenev, Orel, Russia

<sup>3</sup> Space Research Institute of the Russian Academy of Sciences, Moscow, Russia

<sup>4</sup> LLC LUKOIL-Engineering, Moscow, Russia

\* niburt@yandex.ru

**Abstract.** In the field of machine learning, there is no single methodology for data preprocessing, since all stages of this process are unique for a specific task. However, a specific data type is used in each direction. The research hypothesis assumes that it is possible to clearly structure the sequences and phases of data preparation for text recognition tasks. The article discusses the basic principles of data preprocessing and the allocation of successive stages as a specific technique for the task of recognizing ABC characters. ETL set images were selected as the source data. Preprocessing included the stages of working with images, at each of which changes were made to the source data. The first step was cropping, which allowed to get rid of unnecessary information in the image. Next, the approach of converting the image to the original aspect ratio was considered and the method of converting from shades of gray to black and white format was determined. At the next stage, the character lines were artificially expanded for better recognition of printed alphabets. At the last stage of data preprocessing, augmentation was performed, which made it possible to better recognize ABC characters regardless of their position in space. As a result, the general structure of the data preprocessing methodology for text recognition tasks was built.

**Keywords:** neural networks, convolutional neural network, preprocessing, computer vision, machine learning

**For citation:** Trubin A., Morozov A., Zubanova A., Ozheredov V., Korepanova V. The method of preprocessing machine learning data for solving computer vision problems. *Prikladnaya informatika*=Journal of Applied Informatics, 2022, vol.17, no.4, pp.47-56 (in Russian). DOI: 10.37791/2687-0649-2022-17-4-47-56

## Введение

Распознавание образов – одна из самых распространенных задач в машинном обучении. Человек почти всегда может однозначно интерпретировать изображенное на рисунке. Для компьютера, в свою очередь, это изображение не имеет никакого смысла. С помощью алгоритмов машинного обучения человечество научило машины решать задачи различного рода, в том числе и распознавание образов. До недавнего времени мы действительно справлялись с этой задачей лучше компьютеров. Американская компания

Vicarious научила искусственный интеллект решать reCAPTCHA от Google (система, разработанная для защиты веб-сайтов от интернет-ботов и одновременной помощи в оцифровке текстов книг)<sup>1</sup>. Вместо анализа тысяч примеров готовых капч систему научили видеть и анализировать буквы. Vicarious AI протестировали эффективность сети на примере reCAPTCHA. Результат оказался даже лучше, чем у человека – точность 94% против 87%.

<sup>1</sup> URL: <https://www.vicarious.com/posts/vicarious-ai-passes-first-turing-test-captcha>