

Моделирование пользовательского интерфейса для сверточной нейронной сети распознавания сложных символов

А.Е. Трубин¹, Ф.А. Мастяев¹, А.В. Батищев¹, А.И. Зайцев¹, С.А. Алексахина¹

*¹Университет «Синергия», Москва, Россия
niburt@yandex.ru*

Аннотация. Данная статья является третьей в цикле работ, посвященных созданию нейросетевых систем анализа текстовой информации. Результаты использования предложенных ранее авторами статьи моделей для распознавания сложных символов японских азбук при решении практических задач (в том числе в рамках учебного процесса) показывали, что отсутствие эффективных и формализованных методик разработки интерфейсов информационных систем, реализующих нейросетевые технологии, приводит к необходимости дополнительного и достаточно длительного обучения пользователей. Это вызвано тем, что бурное развитие нейросетевых методов для распознавания текстовой информации не всегда сопровождается параллельной разработкой удобных интерфейсов пользователя. Основной причиной подобного развития событий является тот факт, что разработка пользовательского интерфейса требует совсем иного набора компетенций и навыков, нежели разработка нейросетевых методов и моделей. В статье предлагается подход к проектированию пользовательского интерфейса для прототипа настольного приложения, реализующего возможности предложенной авторами ранее нейронной сети для распознавания текстов на японском языке, записанных с использованием одной из двух японских азбук – катаканой или хираганой. При описании подхода к проектированию пользовательского интерфейса применялись нотации UML – для моделирования функциональных требований и BPMN – для моделирования логики работы приложения, взаимодействующего с нейронной сетью, модель которой была предложена авторами во второй статье. Определены принципы и инструментальная база проектирования интерфейса программного решения, спроектированы сценарии использования программы, приведены алгоритмы ее функционирования, описан прототип пользовательского интерфейса. Предложенный подход к построению пользовательских интерфейсов для приложений, использующих нейросетевые методы для решения задач, может быть использован в ходе изучения дисциплин, рассматривающих вопросы использования и проектирования информационных систем.

Ключевые слова: человеко-машинный интерфейс, сверточная нейронная сеть, предобработка данных, анализ текстов, машинное обучение

Для цитирования: Трубин А.Е., Мастяев Ф.А., Батищев А.В., Зайцев А.И., Алексахина С.А. Моделирование пользовательского интерфейса для сверточной нейронной сети распознавания сложных символов // Прикладная информатика. 2023. Т. 18. № 3. С. 105–114. DOI: 10.37791/2687-0649-2023-18-3-105-114

User interface modeling for convolutional neural network for complex character recognition

A. Trubin¹, F. Mastyaev¹, A. Batishchev¹, A. Zaytsev¹, S. Aleksakhina¹

¹Synergy University, Moscow, Russia
niburt@yandex.ru

Abstract. In this article, we design a user interface for a prototype desktop application using the capabilities of the author's neural network for recognizing texts in Japanese written by one of the two Japanese alphabets – katakana or hiragana. During the design, the UML notation, a Use-Case Diagram, was used to build scenarios for using the program, and the BPMN notation was used to describe a program's main algorithm. In the beginning of this article short versions of previous two articles were also given – the basics of proposed method for preprocessing of machine learning data and the main parameters of the proposed convolutional neural network model including its efficiency against reference model EfficientNetB0. In the work, the principles and the tool base for designing the interface of the software solution were defined, the scenarios for using the program, the algorithms of the program were designed, a prototype of the user interface was created.

Keywords: neural networks, convolutional neural network model, computer vision, neural network architecture, machine learning, interface

For citation: Trubin A., Mastyaev F., Batishchev A., Zaytsev A., Aleksakhina S. User interface modeling for convolutional neural network for complex character recognition. *Prikladnaya informatika*=Journal of Applied Informatics, 2023, vol.18, no.3, pp.105-114 (in Russian). DOI: 10.37791/2687-0649-2023-18-3-105-114

Введение

В предыдущих работах авторского кол-лектива [2, 3] была предложена методика предобработки данных машинного обучения для решения задач компьютерного зрения и выделены ее последовательные этапы. Данная методика основана на выборе в качестве исходных данных изображений набора ETL, кадрировании входной информации, преобразовании изображений к исходному соотношению сторон и из оттенков серого в черно-белый формат, расширении линий символов для лучшего распознавания печатных азбук и аугментации. Методика позволила улучшить распознавание японских азбук катаканы или хираганы в независимости от положения символов в пространстве. В результате была предложена типовая процедура предобработки данных для задач рас-

познавания сложных символов японских азбук [2]. Проведенная работа основывалась на различных исследованиях ученых в сфере распознавания символов, анализа данных и моделирования нейросетей [1, 4, 5].

В работах авторов [2, 3] также были рассмотрены архитектуры нейронных сетей, которые могут использоваться для распознавания символов японских азбук. В результате предложена модель сверточной нейронной сети на основе платформы с открытым исходным кодом для машинного обучения TensorFlow в связке с надстройкой для оперативной работы с сетями глубинного обучения Keras. Спроектированная модель показала лучшие результаты в сравнении с готовой архитектурой EfficientNetB0, созданной подразделением Google Research компании Google. Данный вывод был подтвержден проведенными экспериментальными исследованиями