

DOI: 10.24411/1993-8314-2020-10008

*М. И. Дли, докт. техн. наук, профессор
филиал Национального исследовательского университета «МЭИ»
г. Смоленск, midli@mail.ru*

*А. Ю. Пучков, канд. техн. наук, доцент
филиал Национального исследовательского университета «МЭИ»
г. Смоленск, putchkov63@mail.ru*

*Е. И. Лобанева, аспирант
Национальный исследовательский университет «МЭИ»
г. Москва, lobaneva94@mail.ru*

Анализ влияния архитектуры входных слоев свертки и подвыборки глубокой нейронной сети на качество распознавания изображений¹

Представлены результаты исследования влияния характеристик входных слоёв свертки и подвыборки глубокой свёрточной нейронной сети на качество распознавания изображений. Для слоя свёртки изменяемым параметром являлся размер ядра свёртки, варьируемым параметром архитектуры субдискретизирующего слоя являлся размер рецептивного поля. Все перечисленные параметры, определяющие архитектуру входных слоёв свёртки и подвыборки, разработчикам нейронных сетей приходится подбирать на основе своего опыта. В данной работе излагается способ, позволяющий частично автоматизировать этот процесс в результате предварительного анализа характеристик изображения — гистограмм и дисперсий интенсивности цветов пикселей. На основе этих сравнений выработаны рекомендации для выбора размеров ядра свёртки. Приведены итоги апробации указанного способа с помощью программы, написанной на языке Python с использованием библиотек Keras и Tensorflow.

Ключевые слова: компьютерное зрение, оптимизация гиперпараметров, глубокие свёрточные нейронные сети

Введение

Свёрточные нейронные сети (convolutional neural network, CNN) сейчас наиболее успешный аппарат реализации систем компьютерного зрения. Вычислительная структура CNN обеспечивает построение высокоуровневых обобще-

ний в данных на основе множества нелинейных преобразований и позволяет решать относительно простые задачи, например распознавание цифр, и выполнять значительно более сложные интеллектуальные алгоритмы, такие, как постановку медицинских диагнозов на основе обработки многоканальной разнородной информации [1, 2] и проведение интеллектуального мониторинга. К задачам последнего относятся: анализ поведения толпы [3], детектирование смены сцены [4, 5], рас-

¹ Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-01-00425.