

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ ИЗ МУЛЬТФИЛЬМОВ НА ОСНОВЕ ТЕХНОЛОГИИ ГЛУБОКОГО ОБУЧЕНИЯ TENSORFLOW¹

CARTOON IMAGE RECOGNITION BASED ON TENSORFLOW DEEP LEARNING TECHNOLOGY

*Jiang Xichun
Yao Yan
Shen Qi*

Summary. The article is devoted to the consideration of the features of recognizing cartoon images on the basis of deep learning technology TensorFlow. In the process of research the distinctive features and capabilities of TensorFlow technology are outlined. Also formalized algorithm for recognizing images from cartoons on the basis of convolutional neural network. Separately described is the code that uses Python and TensorFlow to implement neural networks designed to recognize hand-drawn cartoon images. In addition, the paper summarizes the basic steps required for neural networks to recognize hand-drawn expressions.

Keywords: image, cartoon, deep learning, recognition, performance.

Цзян Сичунь

*Старший преподаватель,
Хэйхэский университет, г. Хэйхэ, КНР
hhxujxc@163.com*

Яо Янь

*Профессор, Хэйхэский университет, г. Хэйхэ, КНР
19616261@qq.com*

Шэнь Ци

*Старший преподаватель,
Хэйхэский университет, г. Хэйхэ, КНР
56385355@qq.com*

Аннотация. Статья посвящена рассмотрению особенностей распознавания изображений из мультфильмов на основе технологии глубокого обучения TensorFlow. В процессе исследования обозначены отличительные черты и возможности технологии TensorFlow. Также формализован алгоритм распознавания изображений из мультфильмов на основе сверточной нейронной сети. Отдельно описан код, который использует Python и TensorFlow с целью реализации нейронных сетей, предназначенных для распознавания нарисованных вручную изображений из мультфильмов. Кроме того, в статье приведены основные шаги, необходимые нейронным сетям для распознавания нарисованных вручную выражений.

Ключевые слова: изображение, мультфильм, глубокое обучение, распознавание, производительность.

TensorFlow — это технология глубокого обучения, разработанная компанией Google, которая широко используется в области распознавания изображений. Она использует схемы потоков данных для реализации сложных вычислительных процессов глубокого обучения и предоставляет расширенные интерфейсы, которые позволяют облегчить обучение пользователей [1].

TensorFlow предоставляет множество инструментов обработки изображений и манипулирования ими, таких как сверточные нейронные сети (CNN) и рекуррентные нейронные сети (RNN). С помощью этих инструментов мы можем выполнять операции извлечения признаков, классификации и распознавания изображений.

Технология распознавания изображений TensorFlow, основанная на методах глубокого обучения, обладает высокой вычислительной мощностью и гибкостью, может широко использоваться в классификации изображений, распознавании лиц, обнаружении объектов и других областях [2].

Рассмотрим более подробно особенности распознавания изображений их мультфильмов на основе модели CNN.

Алгоритм основан на модели сверточной нейронной сети (CNN).

Во-первых, извлечение объектов изображения. Модель CNN выполняет операции свертки над входным изображением через слой свертки для извлечения локальных объектов на изображении [3].

Во-вторых, уровень объединения. Уровень объединения также является важной частью модели CNN. Его роль заключается в уменьшении размерности объектов, извлеченных из предыдущего сверточного слоя, и снижении вычислительной сложности [4]. Объединяющий слой представляет всю область путем усреднения или максимизации части области входного объекта, тем самым реализуя дальнейшую абстракцию и обобщение.

¹ Научный проект по финансированию фундаментальных научных исследований 2022 г. высших учебных заведений в провинции Хэйлунцзян «Исследования в области глубокого обучения на основе фреймворка TensorFlow», номер проекта: 2022-KYYWF-0374.

В-третьих, полностью подключенный слой: после прохождения через несколько сверточных слоев и их объединения в пул, модель CNN сопоставляет ранее извлеченные объекты с предварительно обученным классификатором через полностью подключенный слой, чтобы получить окончательный результат классификации [5].

С помощью описанных выше шагов модель CNN может реализовать автоматическую классификацию и распознавание изображений. Преимущество этого метода заключается в том, что он позволяет избежать громоздкого процесса ручного извлечения объектов, и в то же время может автоматически адаптироваться к различным сложным изменениям изображения [6, 7].

На следующем этапе исследования опишем конкретные программные процедуры.

Импорт склада и модулей

```
import tensorflow as tf
from tensorflow.keras import datasets, layers, models
import matplotlib.pyplot as plt
```

Подготовка набора данных CIFAR-10.

Он содержит 6000 цветных изображений размером 32x32 в 5 категориях, из которых 2000 используются для обучения, а остальные 4000 — для тестирования. (train_images, train_labels), (test_images, test_labels) = datasets.cifar10.load_data().

Предварительная обработка данных.

Предварительная обработка данных изображений, включая нормализацию, стандартизацию и другие операции. train_images, test_images = train_images / 255.0, test_images / 255.

Определим структуру модели: в этом примере мы используем простую модель сверточной нейронной сети (CNN) для классификации и распознавания изображений из мультфильмов. Модель содержит два сверточных слоя, объединяющий слой, два полностью связанных слоя и выходной слой [8]. Процедура заключается в следующем.

```
model = models.Sequential()
model.add(layers.Conv2D(32, (3, 3), activation='relu',
input_shape=(32, 32, 3)))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Flatten())
model.add(layers.Dense(64, activation='relu'))
model.add(layers.Dense(10))
```

Распознавание изображений из мультфильмов на основе нейронной сети

Построим нейронную сеть для распознавания четырех нарисованных руками выражений: удивление + открытые глаза, удивление + закрытые глаза, улыбка + открытые глаза, улыбка + закрытые глаза. Набор данных взят из 1200 картинок, нарисованных вручную учащимися 1-ой и 2-ой групп 3-го курса Хэйхэского университета по специальности больших данных. Размер картинок в пикселях составляет 100*100. Принцип заключается в использовании алгоритмов глубокого обучения по нарисованным вручную картинкам, чтобы нейронная сеть могла научиться извлекать признаки и классифицировать картинки.

Ниже приведены основные шаги, необходимые нейронным сетям для распознавания нарисованных вручную выражений.

Сначала создаем ру-файл в рамках проекта, созданного в ruchart, затем напишем программу для генерации обучающего подмножества из набора изображений, которые мы подготовили. После завершения выполнения перейдем в каталог, чтобы посмотреть, сгенерировано ли тестовое подмножество. Таким же образом рисуется несколько наборов изображений для создания нескольких обучающих подмножеств. Позже будет создан более крупный тренировочный набор. Теперь у нас есть 10 обучающих подмножеств, и мы объединяем эти 10 подмножеств в более крупный обучающий набор, содержащий в общей сложности 1200 выборок. Перейдем в каталог, чтобы посмотреть, сгенерирован ли набор тестовых подмножеств train.txt. Когда будет успешно сгенерирован большой обучающий набор, мы можем начать обучение модели и сохранить параметры сети.

Затем в рамках проекта создаем новый ру-файл, используемый для написания программ и сохранения сетевых параметров в file.txt. Когда мы получим сетевые параметры, мы можем использовать программу ввода сетевых параметров для идентификации изображения.

В завершении, создаем в рамках проекта новый ру-файл, используемый для написания программ. Введем тестовое изображение test0.jpg, и обработаем его. Введем сетевые параметры и перестроим сеть, обозначим веса в параметрах сети в file.txt в W01, смещения будут импортированы в b01. Запустим модель и получим результаты.

Ниже приведен код, который использует Python и TensorFlow для реализации нейронных сетей для распознавания нарисованных вручную изображений из мультфильмов.

```

import tensorflow as tf
import numpy as np
import matplotlib.pyplot as plt
(x_train, y_train), (x_test, y_test) = tf.keras.datasets.
mnist.load_data()
x_train = x_train / 255.0
x_test = x_test / 255.0
model = tf.keras.Sequential([
    tf.keras.layers.Conv2D(32, (3, 3), activation='relu', input_
shape=(28, 28, 1)),
    tf.keras.layers.MaxPooling2D((2, 2)),
    tf.keras.layers.Flatten(),
    tf.keras.layers.Dense(64, activation='relu'),
    tf.keras.layers.Dense(10)])
model.compile(optimizer='adam', loss=tf.keras.losses.
SparseCategoricalCrossentropy(from_logits=True),
metrics=['accuracy'])
model.fit(x_train, y_train, epochs=5, validation_
data=(x_test, y_test))
predictions = model.predict(x_test)
print('Accuracy:', tf.keras.metrics.
SparseCategoricalAccuracy().compute(y_test, predictions))

```

Приведенный выше код использует модель сверточной нейронной сети для распознавания изображений в наборе данных, состоящих из нарисованных вручную картинок из мультфильмов. Также этот код позволяет оценить производительность модели с помощью обучающих и тестовых данных. Проведенные тесты позво-

лили установить, что уровень достоверности составляет 0,8762.

Заключение

В этом исследовании, с использованием технологии TensorFlow, рассматриваются особенности применения глубокого обучения для распознавания изображений из мультфильмов. Путем экспериментальной проверки мы обнаружили, что глубокое обучение позволяет эффективно повысить точность распознавания изображений из мультфильмов, демонстрируя сильную способность к обучению и обобщению. Кроме того, мы также обнаружили, что благодаря использованию передовых архитектур нейронных сетей, таких как ResNet точность распознавания может быть дополнительно повышена, и это способно обеспечить мощную поддержку для разработки анимации, игр и других областей.

В этом исследовании все еще существуют некоторые ограничения, такие как небольшой размер набора данных, а также сложность модели и вычислительные ресурсы, которые не учитываются в полной мере. Будущие исследования могут еще больше расширить набор данных, оптимизировать структуру модели и улучшить способность к обобщению и надежность модели, чтобы более эффективно решать задачу распознавания мультяшных изображений.

ЛИТЕРАТУРА

1. Захватов Д.М., Лукиных А.А. Распознавание изображений с использованием нейронных сетей: проблемы и перспективы // Современные научные исследования и инновации. 2023. № 9 (149).
2. Стеценко А.И., Смагулова А.С. Разработка нейронной сети для распознавания изображений на базе TENSORFLOW // E-Scio. 2023. № 3 (78). С. 133–139.
3. Винокуров И.В. Использование свёрточной нейронной сети для распознавания элементов текста на отсканированных изображениях плохого качества // Программные системы: теория и приложения. 2022. Т. 13. № 3 (54). С. 29–43.
4. Гребенникова Н.И., Злобин А.Б., Сергеев М.Ю., Сергеева Т.И. Автоматизированная система детектирования и распознавания графических изображений применением сверточных нейронных сетей // Экономика и менеджмент систем управления. 2022. № 4 (46). С. 84–91.
5. Годунов А.И., Балаян С.Т., Егоров П.С. Сегментация изображений и распознавание объектов на основе технологии сверточных нейронных сетей // Надежность и качество сложных систем. 2021. № 3 (35). С. 62–73.
6. Домбаян Г.С., Куликова О.В., Шпаковский В.П. Распознавание изображений при помощи алгоритмов, основанных на нечетких нейронных сетях // Труды Северо-Кавказского филиала Московского технического университета связи и информатики. 2021. № 1. С. 20–21.
7. Зиядинов В.В., Курочкин П.С., Терешонок М.В. Оптимизация обучения сверточных нейронных сетей при распознавании изображений с низкой плотностью точек // Радиотехника и электроника. 2021. Т. 66. № 12. С. 1207–1215.
8. Калашников А.О., Барабанов В.Ф., Нужный А.М., Барабанов А.В. Поддержка принятия решений в системе распознавания изображений с применением искусственной нейронной сети // Вестник Воронежского государственного технического университета. 2021. Т. 17. № 2. С. 19–23.

© Цзян Сичунь (hxyjxc@163.com); Яо Янь (19616261@qq.com); Шэнь Ци (56385355@qq.com)
Журнал «Современная наука: актуальные проблемы теории и практики»