

РАЗРАБОТКА ПРИЛОЖЕНИЯ «ХРАНИЛИЩЕ ВИЗУАЛЬНЫХ ИЗОБРАЖЕНИЙ С ИНТЕЛЛЕКТУАЛЬНОЙ КЛАССИФИКАЦИЕЙ ПО ТЕМАТИЧЕСКИМ КОЛЛЕКЦИЯМ ДЛЯ МАШИННОГО ОБУЧЕНИЯ»

Корлякова Мария Антоновна

студент, ФГБОУ ВО «Пермский государственный гуманитарно-педагогический университет», РФ, г. Пермь

Шаповалова Наталья Олеговна

студент, ФГБОУ ВО «Пермский государственный гуманитарно-педагогический университет», РФ, г. Пермь

Аннотация. В данной статье описывается алгоритм создания классификатора пород собак с помощью сверточной нейронной сети. Данный алгоритм дает представление о том, как происходит подготовка к обучению, создание модели нейронной сети и ее обучение.

Ключевые слова: сверточные нейронные сети, многоклассовая классификация изображений, машинное обучение

Классификация — это самая популярная задача машинного обучения. Ее целью является классифицирование объектов по заранее известному признаку. Для реализации данной задачи широко применяются сверточные нейронные сети.

Разрабатываемый нами классификатор пород собак — это программа, которая представляет собой обученную модель сверточной нейронной сети на основе изображений собак. Помимо этого, в классификаторе содержится информация о породах, которую может получить пользователь. Основная цель программы — определение породы собаки по загруженному пользователем изображению, а также вывод информации о ней. Классификатор реализуется на языке программирования Python, а также с помощью Keras, Tensorflow, NumPy.

В нашей модели нейронной сети мы использовали многоклассовую классификацию. Многоклассовая классификация — это задача классификации с более чем двумя классами.

В нашей работе применяется классификация с использованием признаков, извлеченных из набора изображений собак, где каждое изображение может быть отдельной породой. Каждое изображение представляет собой один образец и помечен как один из 3 возможных классов. Многоклассовая классификация предполагает, что каждому образцу присваивается одна метка.

Для сбора датасета мы написали программу для скачивания изображений формата jpg на языке программирования Python.

Программа скачивает только с веб-страницы Pinterest. Она работает с использованием Chrome Driver. Chrome Driver — это инструмент с открытым исходным кодом для автоматического тестирования веб-приложений. Он предоставляет возможности для навигации по веб-страницам. В нашем случае он открывает страницу с изображениями, прогружает ее до конца

и сохраняет фотографии в созданную папку.

Собранный датасет состоит из трех папок:

Training — содержит изображения для обучения;

Validation — содержит изображения проверки модели;

Test — содержит тестовые изображения.

Топология модели:

Слой свертки: 16 фильтров размером 5 на 5.

Слой подвыборки: max-pooling по каждому квадрату размером 2 на 2 **ПИКСЕЛЯ**.

Слой свертки: 32 фильтра размером 5 на 5.

Слой подвыборки: max-pooling по каждому квадрату размером 2 на 2 **ПИКСЕЛЯ**.

Слой свертки: 64 фильтра размером 2 на 2.

Слой подвыборки: max-pooling по каждому квадрату размером 2 на 2 **ПИКСЕЛЯ**.

Слой свертки: 128 фильтров размером 5 на 5.

Слой подвыборки: max-pooling по каждому квадрату размером 2 на 2 **ПИКСЕЛЯ**.

Полносвязный слой: 1024 и 256 нейронов.

Выходной слой: классы изображений (породы собак).

Наша модель сверточной нейронной сети состоит из сверточных, субдискретизирующих и полносвязных слоев, которые располагают нейроны в 3 измерениях – ширине, высоте и глубине.

Первые два типа слоя, чередуясь между собой, формируют входной вектор признаков для многослойного персептрона.

Для извлечения высокоуровневых признаков применена свертка. Первый её слой условно несёт ответственность за схватывание признаков низкого уровня, таких как кромки, цвет, ориентация градиента и так далее. Через дополнительные слои архитектура адаптируется к признакам высокого уровня.

В свертке имеется ядро (фильтр), которое представляет из себя матрицу весов. Оно продвигается по двумерному изображению и выполняет операцию умножения с той частью данных, над которой оно находится в данный момент, потом суммирует все полученные значения в один выходной пиксель.

Размер ядра сверточной нейронной сети определяет количество признаков, которые будут объединены для получения нового признака на выходе. В нашем случае выбрана размерность ядра 5x5.

Первый слой свертки нашей модели имеет 16 фильтров с ядром 5x5. Постепенно количество фильтров увеличивается, в связи с чем используется паддинг для избежания потери разрешений выходных данных изображений.

Благодаря ему исходное изображение расширяется, добавляя к краям поддельные пиксели нулевого значения. Таким образом, при прохождении ядро позволяет неподдельным пикселям оказываться в своем центре, а затем распространяется на поддельные пиксели за пределами края, создавая выходную матрицу того же размера, что и входная.

В качестве выходного слоя модели мы получаем только один нейрон, который решает к какой категории относится изображение.

В результате тестового обучения на трех породах модель показала, что доля верных ответов на тестовых данных составляет 68%.

Также нами собрана следующая информация, которая будет выводиться при определении породы: продолжительность жизни, характер, особенности содержания, возможные болезни.

В будущем мы планируем увеличить количество классифицируемых пород, а также разработать полноценное пользовательское приложения на ПК.

Список литературы:

1. «Машинное обучение. Нейронные сети (часть 3) — Convolutional Network под микроскопом. Изучение АПИ Tensorflow.js». URL: <https://habr.com/ru/post/518766/> (дата обращения: 21.05.2023)
2. «Как работает свёрточная нейронная сеть». URL: <https://habr.com/ru/company/skillfactory/blog/565232/> (дата обращения: 21.05.2023)