

УДК 004.85

**М.С. Мясникова**

студент кафедры вычислительных систем и сетей

**А.М. Сергеев**

кандидат технических наук, доцент

Санкт-Петербургский государственный университет аэрокосмического приборостроения

## **СВЕРТОЧНАЯ НЕЙРОННАЯ СЕТЬ «РАСПОЗНАВАНИЕ РУКОПИСНЫХ ЦИФР ПО ИЗОБРАЖЕНИЮ»**

Рассматривается создание и применение нейронной сети для обработки изображений цифр. Нейронная сеть написана на языке программирования Python и библиотек NumPy и Matplotlib. Нейронная сеть может обрабатывать изображения, созданные пользователем, и классифицировать их.

**Ключевые слова:** нейронная сеть, изображения, слой, цифры

**M.S. Myasnikova**

Student of the Department of Computing Systems and Networks

**A.M. Sergeev**

Candidate of Technical Sciences, Associate Professor

St. Petersburg State University of Aerospace Instrumentation

## **CONVOLUTIONAL NEURAL NETWORK "HANDWRITTEN DIGIT RECOGNITION BY IMAGE"**

The creation and application of a neural network for processing digital images is considered. The neural network is written in the Python programming language and the NumPy and Matplotlib libraries. The neural network can process images created by the user and classify them.

**Keywords:** neural network, images, layer, numbers

Нейронные сети и их применение в области распознавания изображений являются одними из наиболее перспективных направлений в современной науке и технологии. С увеличением объема данных и потребности в автоматизации процессов обработки информации, задачи классификации и распознавания изображений становятся все более востребованными. Особенно актуальными эти задачи являются в контексте рукописного

ввода, где точное и быстрое распознавание цифр и символов может существенно повысить эффективность различных приложений, таких как банковские системы, системы ввода данных и образовательные технологии [1].

Целью данной работы является разработка обучаемой среды, использующей в качестве алгоритма обучения нейросеть с учителем (метод обратного распространения ошибки), анализирующей изображения с целью выявления в них рукописных цифр и их классификации.

Архитектура нейронной сети представляет собой перцептрон, содержащий один входной, один выходной и три скрытых слоя. Распознаватель будет работать с изображениями 16x16 пикселей, что определяет размер входного слоя в 256 нейронов. Количество выходных нейронов соответствует количеству цифр в десятичной системе счисления. Поскольку пользователю неудобно работать с разрешением 16x16 пикселей (большое количество пикселей замедляет процесс обучения нейросети), то для введения символа подлежащего распознаванию будет отведено поле 256x256 пикселей. Для свертки изображения такого размера в формат данных, пригодный для обработки нейросетью, был разработан следующий алгоритм (листинг 1).

Листинг 1

```
def process_image(self):
    image = np.zeros((16, 16))
    chunk_width = self.canvas_width // 16
    chunk_height = self.canvas_height // 16
    for i in range(16):
        for j in range(16):
            x_start = i * chunk_width
            y_start = j * chunk_height
            chunk = []
            if self.canvas.find_overlapping(x_start, y_start, x_start +
chunk_width, y_start + chunk_height):
                chunk.append(1)
            else:
                chunk.append(0)
            image[i][j] = np.mean(chunk)
    return image
```

Входной слой принимает на вход изображение, преобразованное в одномерный массив. Скрытые слои, состоящий из 64 нейронов, выполняют нелинейное преобразование с использованием функции активации сигмоиды. Выходной слой, состоящий из 10 нейронов, соответствует десяти классам цифр и также использует сигмоидную функцию активации [3].

В ходе разработки была получена программа, в которой пользователь имеет возможность нарисовать рукописный символ цифры и выполнить классификацию изображения с помощью нейросети. В случае, если нейросеть совершила ошибку, пользователь может ввести ожидаемый результат, после чего выполняется итерация обучения нейросети. На рисунке 1 представлен графический интерфейс программы.

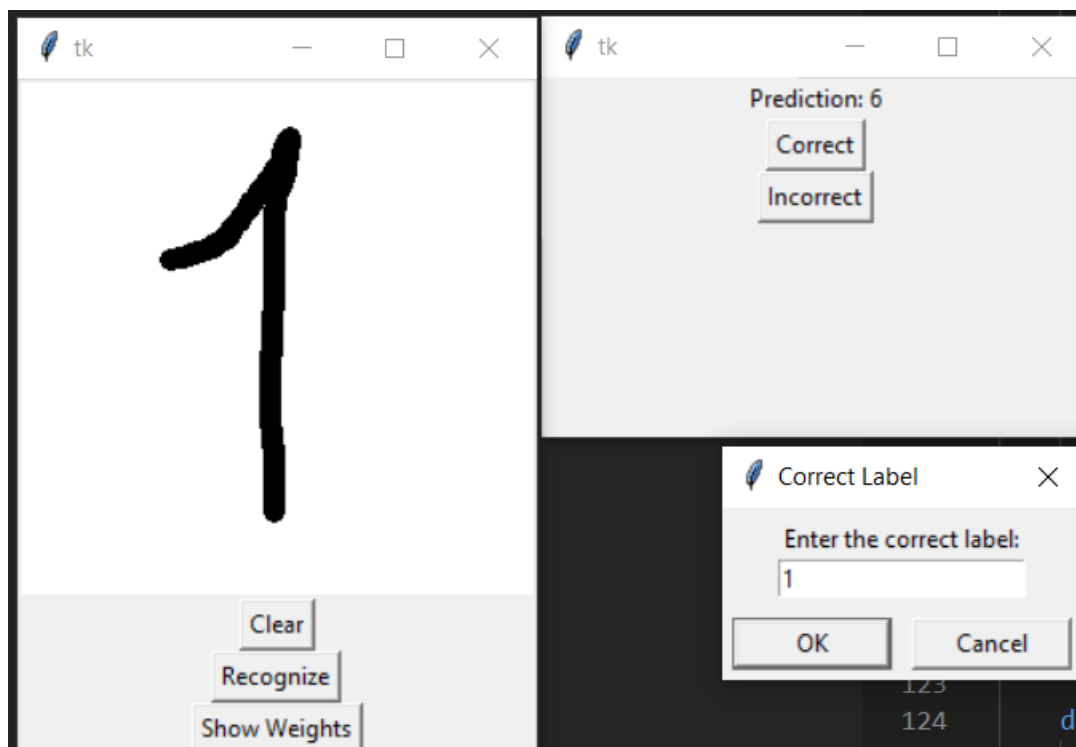


Рисунок 1 – Графический интерфейс программы

В ходе обучения нейросети были выявлены следующие особенности методологии обучения распознавателя:

- 1) рукописный символ должен занимать всю плоскость холста;
- 2) следует располагать символы в разных частях холста;
- 3) следует менять символ для распознавания каждую итерацию.

Поскольку обучение происходило без заготовленного датасета, то результат обучения не сильно заметен после проведения каждой итерации, однако динамику можно наблюдать ввиду возрастания количества верно распознанных символов.

Данный проект может быть использован при исследовании особенностей обучения простейших нейросетей – классификаторов. Результат обучения сохраняется в специальный файл, который может быть доступен для использования другими пользователями. Также данный файл позволяет продолжать обучение при повторном использовании программы. Язык программирования Python обеспечивает гибкую архитектуру кода, чем обусловлена расширяемость программы для дальнейшей интеграции инструментов исследования и статистики.

В ходе данной работы была успешно разработана обучаемая среда для распознавания рукописных цифр с использованием нейронной сети. Полученный проект позволяет пользователю рисовать цифры и классифицировать их с помощью созданного классификатора. Если нейросеть ошибается, пользователь может ввести правильный результат, и нейросеть выполнит дополнительное обучение. Эксперименты показали, что точность распознавания улучшается по мере проведения итераций обучения, несмотря на отсутствие заранее подготовленного набора данных. Созданный проект может быть полезным инструментом для изучения процессов обучения нейросетей и их применения в задачах распознавания изображений. Результаты обучения сохраняются, что позволяет продолжать работу с программой и использовать накопленные данные в дальнейшем.

### **Библиографический список**

1. Philipp Christian Petersen. Neural Network Theory – [Электронный ресурс] URL: [http://pc-petersen.eu/Neural\\_Network\\_Theory.pdf](http://pc-petersen.eu/Neural_Network_Theory.pdf) (дата обращения 16.05.2024)
2. Simeon Kostadinov. Understanding Backpropagation Algorithm – [Электронный ресурс] URL: <https://towardsdatascience.com/understanding-backpropagation-algorithm-7bb3aa2f95fd> (дата обращения 16.05.2024)
3. Sagar Sharma. Activation Functions in Neural Networks – [Электронный ресурс] URL: <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6> (дата обращения 16.05.2024)