

**Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
Московский государственный технический университет им.
Н. Э. Баумана**

23 0102

Согласовано

К.Т.Н., доцент

_____ Терехов В. И.

“ ____ ” _____

Утверждаю

К.Т.Н., доцент

_____ Терехов В. И.

“ ____ ” _____

**Лабораторная работа №3
по курсу «Интеллектуальные системы»**

Решение задач с использованием искусственных нейронных сетей

Студент группы ИУ5-82

_____ Гуца А. В

“ ____ ” _____

Содержание

1	Цель работы	1
2	Задание	1
3	Описание предметной области и выбранной задачи	2
4	Структура, основные параметры выбранной ИНС	3
4.1	Блок схема алгоритма обучения	5
5	Описание программы, ее ключевые особенности и новшества	6
6	Протоколы проведенных экспериментов	8
6.1	Обучение сети	8
6.2	Режим распознавания	8
7	Выводы	11
8	Используемая литература	12

1 Цель работы

Целью лабораторной работы является углубление и закрепление теоретических знаний, полученных на лекциях, приобретение практических навыков самостоятельного исследования при решении задач выбора, обучения и работы ИНС.

В процессе выполнения лабораторной работы по теме «Решение задач с использованием искусственной нейронной сети» студенты решают следующие задачи (задания):

- описывают предметную область и выбирают решаемую задачу (предпочтение должно отдаваться задачам практической направленности);
- определяют множество обучающих примеров;
- в зависимости от решаемой задачи выбирают структуру ИНС;
- выбирают алгоритм обучения ИНС;
- проводят обучение ИНС на тестовом множестве примеров с помощью выбранного алгоритма обучения;
- исследуют работу обученной ИНС в режиме распознавания.

2 Задание

Разработать программу, которая обучает ИНС распознавать прописные английские буквы на изображении любого размера и цвета, но состоящее не менее чем из 35 пикселей (матрица 5x7). При этом, ИНС должна иметь входы, ассоциированные с пикселями матрицы, и выходы, количество которых соответствует количеству прописных латинских букв.

В написанной или выбранной программе должна быть реализована возможность задания множества обучающих примеров в виде образов, а также изменения величины коэффициента скорости обучения. Программа должна предусматривать два режима работы: обучения и распознавания. Обучение должно

производиться с использованием алгоритма, соответствующего архитектуре выбранной для решения задачи ИНС. Вероятность распознавания обученной ИНС должна быть не менее 65%.

3 Описание предметной области и выбранной задачи

Задача распознавания рукописного текста является одной из важнейших задач для интеллектуальных систем (распространенная аббревиатура OCR - optical character recognition). Успешное решение данной задачи применяется в следующих областях:

- оцифровка печатных книг, газет, журналов
- автоматизированная обработка рукописных бланков
- системы дополненной реальности

Входными данными являются растровые изображения рукописных (или печатных) символов прописных букв английского алфавита. Данные изображения имеют любой размер и цветовую информацию, но все изображения проходят предварительную обработку для изменения размера и бинаризации, чтобы данные изображения возможно было подать на вход разработанной ИНС. Ограничение на входные данные: прописные буквы должны быть вписаны в изображение, без наличия значительных пустых полей.

Выходными данными являются: распознанный символ и уверенность нейронной сети в данном ответе.

Искусственная нейронная сеть (ИНС) — математическая модель, а также её программная или аппаратная реализация, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма.

ИНС представляют собой систему соединённых и взаимодействующих между собой простых процессоров (искусственных нейронов). Такие процессоры обычно довольно просты (особенно в сравнении с процессорами, используемыми в персональных компьютерах). Каждый процессор подобной сети имеет дело только с

сигналами, которые он периодически получает, и сигналами, которые он периодически посылает другим процессорам. И, тем не менее, будучи соединёнными в достаточно большую сеть с управляемым взаимодействием, такие локально простые процессоры вместе способны выполнять довольно сложные задачи.

4 Структура, основные параметры выбранной ИНС

Для решения данной задачи был выбран многослойный перцептрон, состоящий из четырех слоев: входного, выходного и двух скрытых. Каждый нейрон в слое l имеет связи с каждым нейроном из слоя $l + 1$.

Число нейронов в каждом слое:

- 1) Соответствует количеству пикселей во входном бинарном изображении $5 * 7 = 35$
- 2) Размер данного слоя был выбран с учетом необходимости способности распознавать большое количество типов символов, исходя из практических наблюдений, размер первого скрытого слоя должен на порядок превышать входной: $35 * 35 = 1225$
- 3) Нейроны в данном слое отвечают за группировку сигналов из предыдущего слоя, в результате экспериментов был выбран размер в $35 * 2 = 70$ нейронов.
- 4) Каждый нейрон данного слоя соответствует одной букве английского алфавита: 26

Как функция активации используется сигмоида:

$$F(x) = \frac{1}{1 + e^{-x}}$$

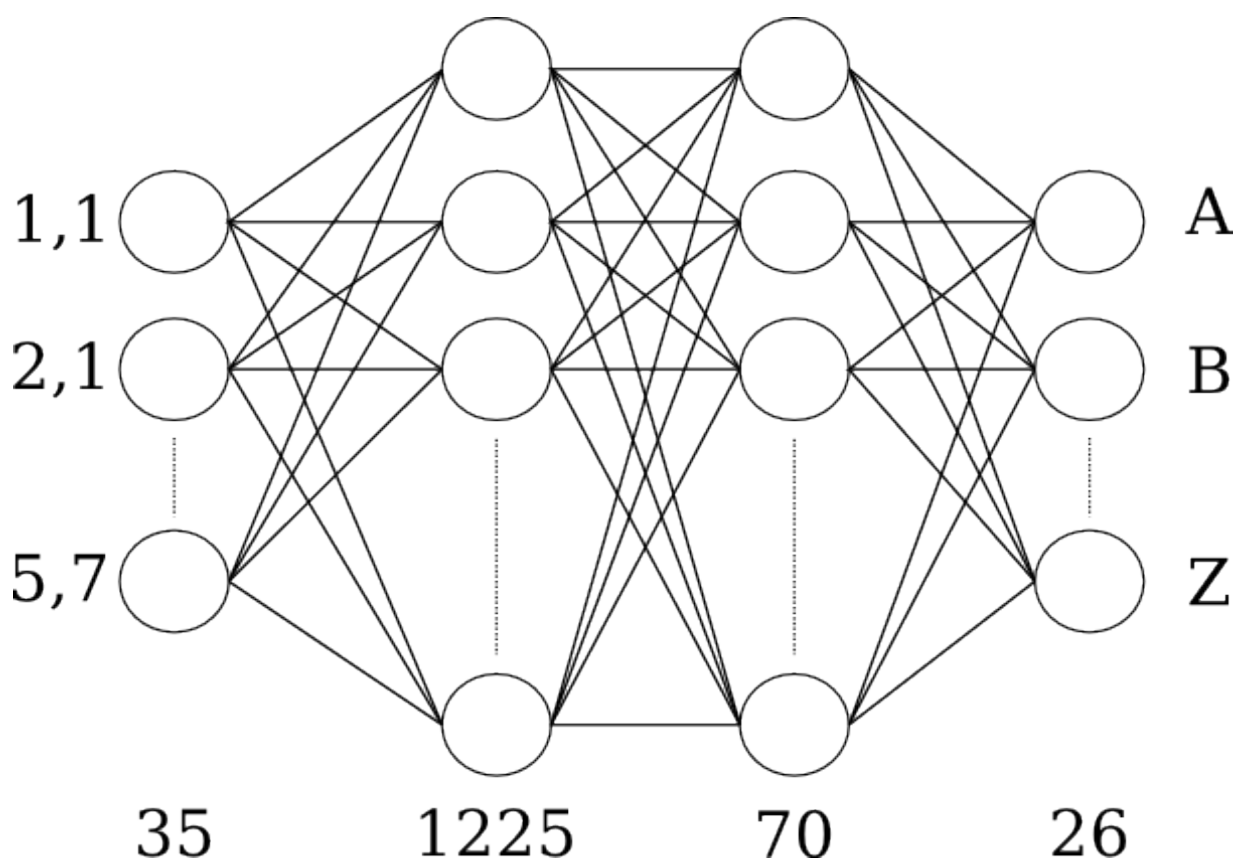


Рисунок 1 – Структура ИНС

4.1 Блок схема алгоритма обучения

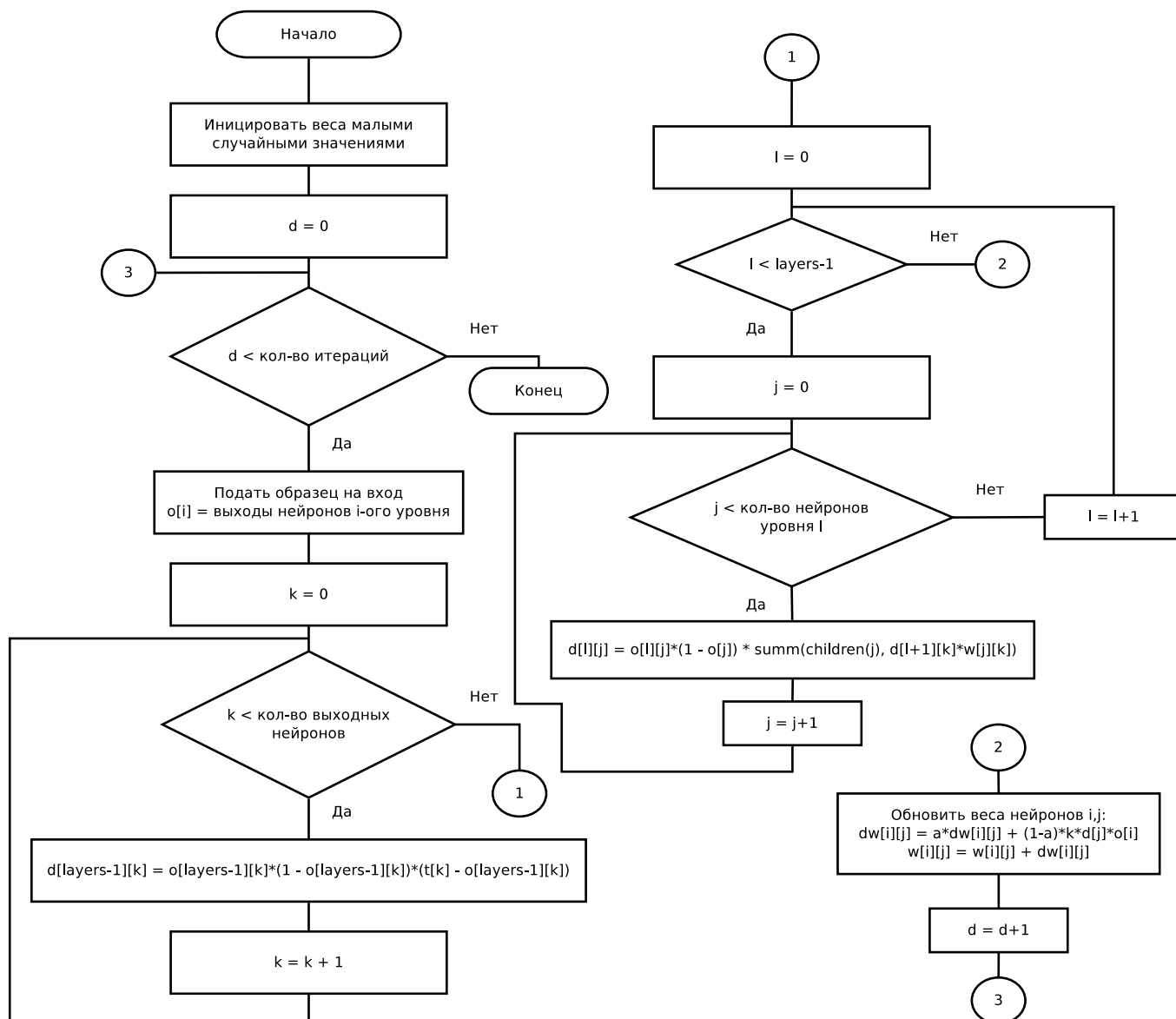


Рисунок 2 – Блок схема алгоритма обратного распространения ошибки

5 Описание программы, ее ключевые особенности и новшества

Одной из отличительной особенностью является предварительная обработка изображений:

- Исходное изображение преобразуется в изображение 5х7 пикселей
- Проводится перевод цветовых каналов в серые тона
- Бинаризация по критерию Оцу - позволяет разделить пиксели двух классов («полезные» и «фоновые»), рассчитывая такой порог, чтобы внутриклассовая дисперсия была минимальной.

Для обработки изображений был применен метод функциональной обработки изображений (и применена соответствующая библиотека, разработанная Владимиром Пантелеевым).

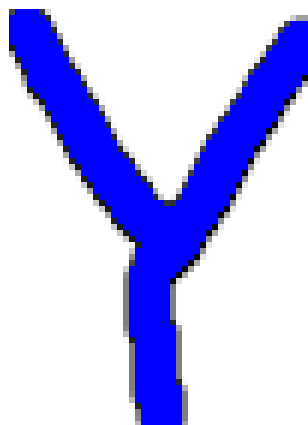


Рисунок 3 –
Исходное
изображение

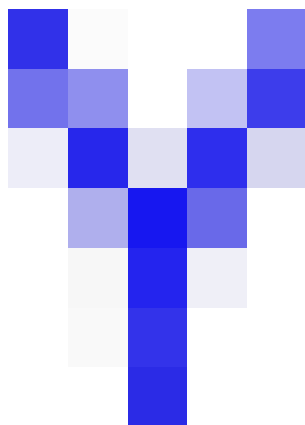


Рисунок 4 –
Изменение
размера

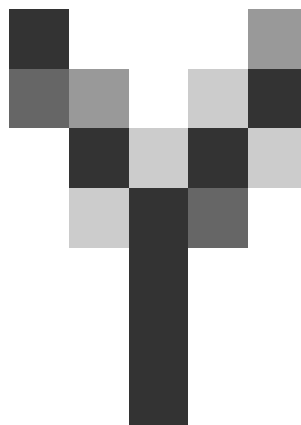


Рисунок 5 –
Перевод в серые
тона

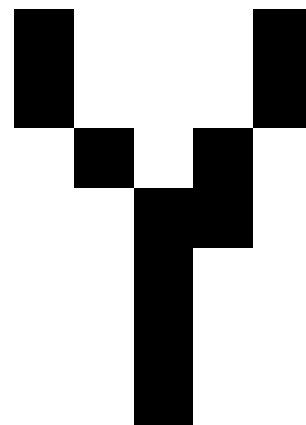


Рисунок 6 –
Бинаризация по
критерию Оцу

Все компонентные системы написаны на языке D - современный мультипарадигменный системный язык программирования. Исключительной особенностью является то, что архитектура нейронной сети задается на этапе компиляции. Генерация кода, описывающего ИНС, на этапе компиляции позволяет значительно оптимизировать скорость выполнения программы.

Пример генерации описания перцептрона, используемого в системе:

```
enum INPUT_SIZE_X = 5;
enum INPUT_SIZE_Y = 7;
enum INPUT_SIZE = INPUT_SIZE_X * INPUT_SIZE_Y;
enum OUTPUT_SIZE = 26;
alias TestNet = Perceptron!(INPUT_SIZE, INPUT_SIZE*INPUT_SIZE
    ↪ , 2*INPUT_SIZE, OUTPUT_SIZE);
```

Программа работает в консольном режиме, справку по пользованию программы можно получить, набрав в терминале команду:

```
$ perceptron --help
perceptron [args]
```

```
args: --learning      - defines learning mode for neural network
      --recognition   - defines recognition mode for neural network.
                        default flag.
      --config=path   - configuration file, default is 'config.json'.
      --genconfig      - if set, new config file is generated in
                        --config location + .example extention.
      --input          - test input to feed trained network. Default
                        is 'test.png'
```

Программа может работать в двух режимах: режим обучения и режим распознавания. Все необходимые параметры для обучения и описание образцов хранится в текстовом конфигурационном файле:

```
{
    "recogintionFolder": "data/recognition",
    "recognitionSamples": [],
    "logFile": "perceptron.log",
    "learnFolder": "data/learn",
    "saveInput": false,
    "controlPart": "0.0",
    "networkFile": "network.json",
    "trainingFactor": "0.8",
    "inertiaFactor": "0.0",
    "iteratesCount": "50",
```

```

    "learnSamples ": [
        {
            "path ": "a",
            "symbol ": "A"
        },
        {
            "path ": "b",
            "symbol ": "B"
        }
        ...
    ]
}

```

6 Протоколы проведенных экспериментов

6.1 Обучение сети

Были подготовлены 260 обучающих изображений размером 50x70 пикселей каждый, по 10 изображений на каждую букву.

Обучение запускалось следующей командой:

```
$ ./perceptron --learning
```

Обучение длилось около часа, ниже представлен график точности распознавания на обучающей выборке в зависимости от времени:

В результате обучения удалось достичь точности распознавания в 98.46% на обучающей выборке.

6.2 Режим распознавания

Для проверки обученной ИНС (и для практического использования) в программе присутствует режим распознавания. Вызов режима распознавания:

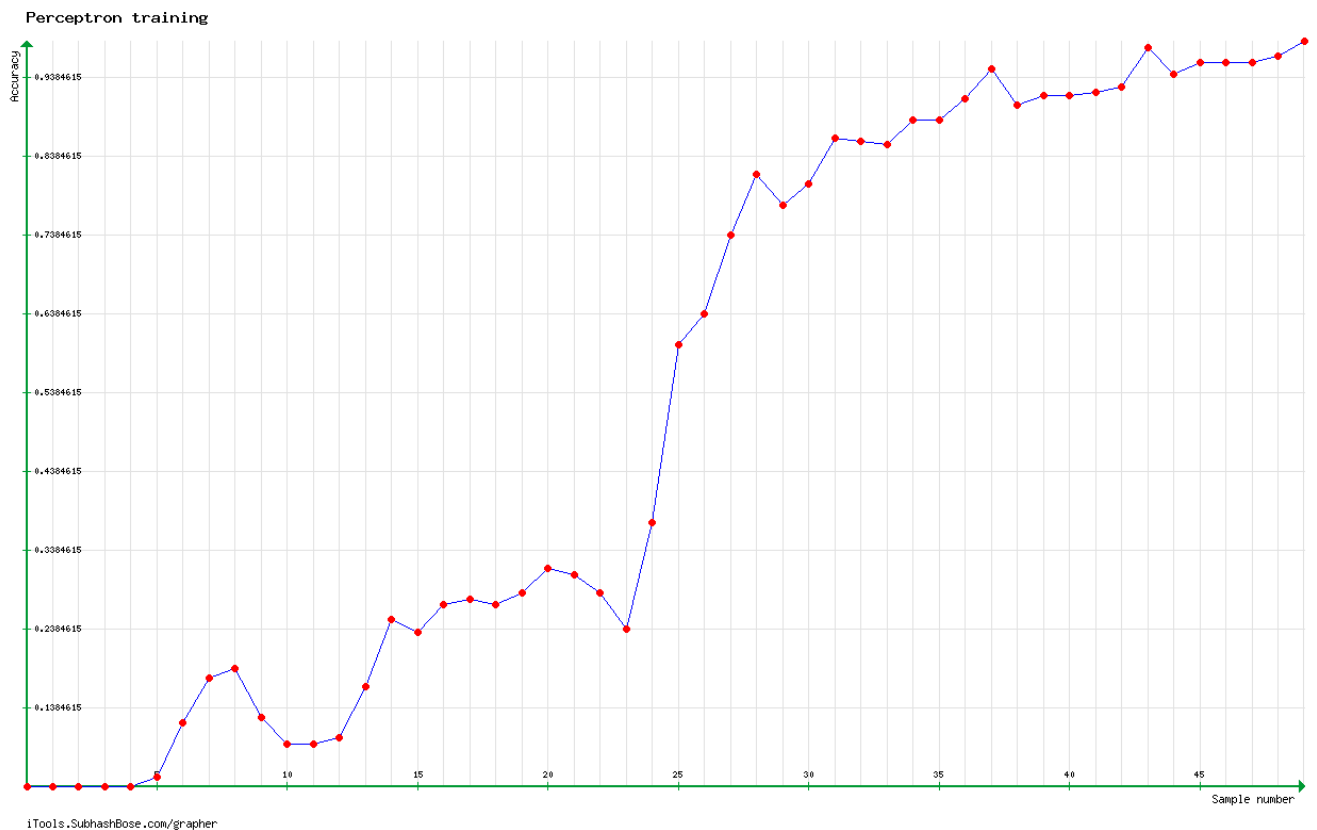
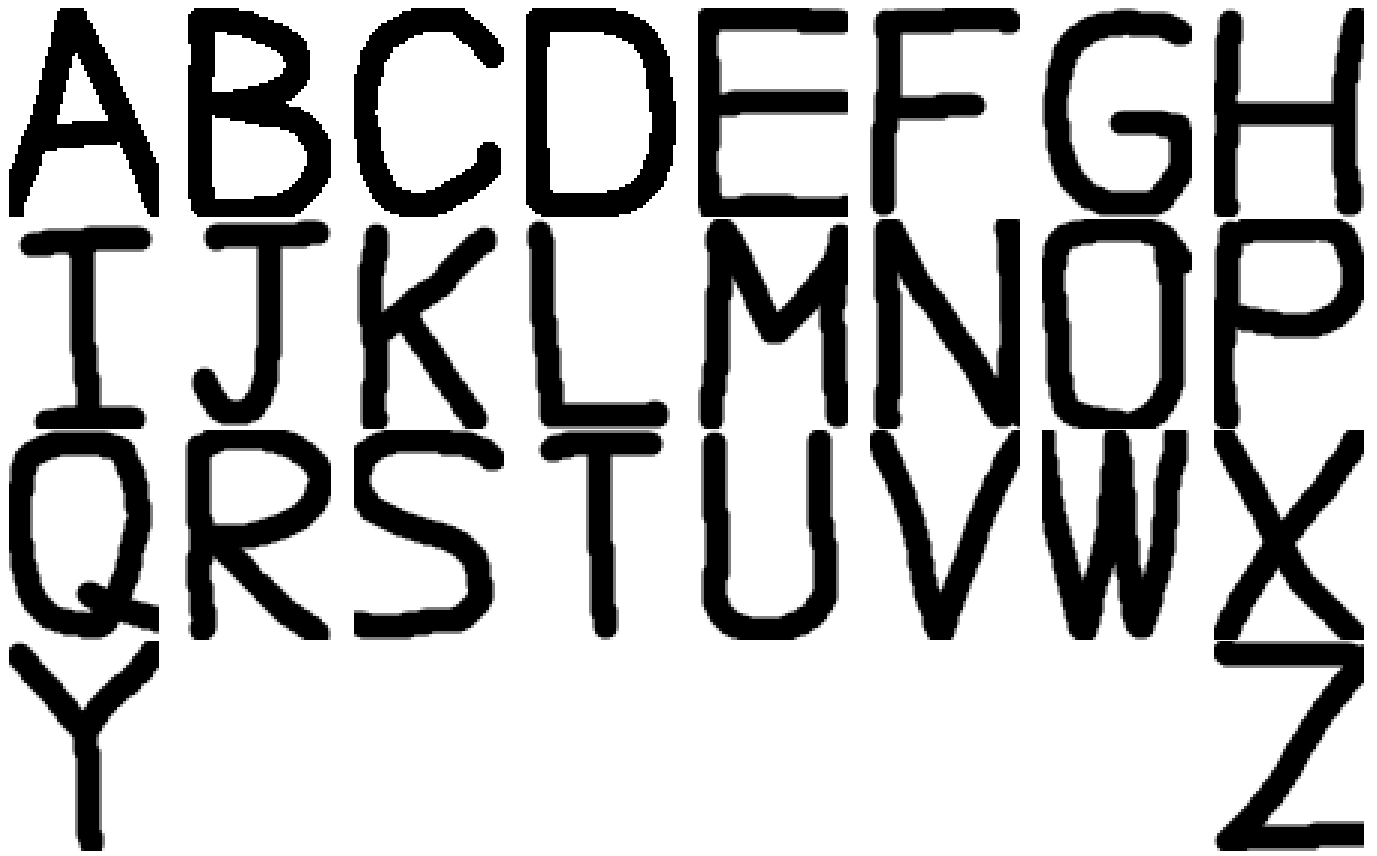


Рисунок 7 – График точности распознавания во время обучения

```
$ ./perceptron --recognition --input=my_image.png
Notice: Start initialization is finished
Notice: Application operates in recognition mode
```

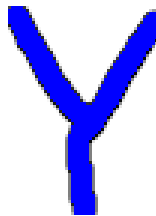


Рисунок 8 – Входное изображение для тестирования ИНС my_image.png

```
Notice: Loading trained network from network.json
```

```
Notice: Loading test sample from test.png
```

```
Notice: Loading symbol map from config
```

```
Notice: Getting answer
```

```
Notice: Answer is: Y
```

```
Notice: Assurance is: 0.946363
```

Программа выводит символ и уверенность сети в данном решении. В данном случае ответом является символ 'Y' и уверенность 94.64%

7 Выводы

В результате работы над лабораторной работы были усвоены следующие знания и навыки:

- Обработка изображений, изменение масштаба, бинаризация и фильтрация.
- Постановка задачи в области OCR
- Структура и логика работы многослойных ИНС, алгоритмы их обучения
- Анализ работы ИНС, контроль процесса обучения
- Навык скоростного рисования изображений прописных букв английского алфавита

В результате исследования работы ИНС после обучения было выявлено некоторые особенности выбранной архитектуры ИНС:

- Перцептрон не может распознать изображения после сдвига, поворота, изменений масштаба.
- ИНС может обрабатывать зашумленные данные на достойном уровне
- Наиболее проблемной буквой является 'Q' из-за ее близости к букве 'O'. Стоит заметить, что частотное появление 'Q' в текстах на английском языке составляет всего 0,01%, что значительно понижает вероятность появления ошибки распознавания текстов.

8 Используемая литература

- 1) Терехов В.И. Курс лекций по предмету «Интеллектуальные системы», 2014
- 2) Wikipedia.org Искусственная нейронная сеть
- 3) Wikipedia.org Перцептрон
- 4) Wikipedia.org Метод обратного распространения ошибки
- 5) Wikipedia.org Метод Оцу
- 6) Vladimir Panteleev, Functional image processing in D