

Программная реализация однослойной нейронной сети для распознавания цифровых символов

А.Ю.Кулакович

Донской государственной технической университет, Ростов-на-Дону

Аннотация: Рассмотрены существующие способы для распознавания символов. Выявлены плюсы и минусы применения нейронных сетей для распознавания символов. Приведён обзор методов для распознавания символов. Так же определены достоинства и недостатки алгоритмов для распознавания текста. Реализован алгоритм простейшей однослойной нейронной сети для распознавания каллиграфических и печатных цифр. В работе был проведен анализ обучаемости данной однослойной нейронной сети.

Ключевые слова: распознавание символов, однослойная, нейронная сеть, алгоритмы, персептрон.

Введение

На данный момент проблема распознавания символов является актуальной. В настоящее время разработано и определено довольно много алгоритмов, решающих данную задачу [1, с. 83]. Наряду с печатными текстовыми символами существует и рукописный ввод, который гораздо проще определяется с помощью нейронных сетей. Одной из целей данной статьи является определить достоинства и недостатки качества работы алгоритмов для распознавания символов. В качестве символов будут рассматриваться как печатные символы, так и рукописные символы. Реализовано программное обеспечение на основе однослойной нейронной сети для распознавания символов.

Обзор методов для распознавания текста

Распознавание образов включает в себя тему распознавания текста. Под распознаванием текста выделяется два основных направления:

- 1) Это изучение способности алгоритма к распознаванию символов, как у живых существ, анализ и способность к их моделированию;

2) Развитие и усовершенствование теорий и методов построения алгоритмов, которые будут предназначены для решения определенных задач в прикладных целях.

Данные методы применимы как к распознаванию рукописного ввода, оцифровки книг, так и довольно большого применения в бытовых целях. В настоящее время не прекращаются исследования в данной области, что говорит о ее высоком приоритете и значимости для современного мира.

Следует заметить, что под методами распознаванием символов обычно понимают три основных метода [1, С. 83]:

1) Метод, в результате которого происходит сравнение с уже заранее подготовленным шаблоном;

2) Метод распознавания текста, который использует критерии распознаваемого объекта.

3) Метод распознавания текста, который основан на самообучающихся алгоритмах, к которым можно отнести нейронные сети.

Для каждого метода осуществляется этап предобработки, в результате которого происходит подготовка изображения к распознаванию.

Для начала к изображению применяется *медианный фильтр*, в результате которого минимизируется шум, смазываются острые края букв. Далее приходит на помощь *монохромный фильтр*, в результате которого границы букв четко фиксируются, а так же происходит бинаризация изображения – процесс при котором цветное изображение или градация серого переводится в черно-белое [2, С. 126].

Россия – многонациональное государство, отличающееся большим этнокультурным многообразием.

Рис. 1 –Изображение после обработки медианного и монохромного фильтра.

После предобработки начинается *сегментация* изображения – процесс разбиения текста на символы. Для начала определяется среднее значение расстояния между двумя символами в слове. Далее происходит процесс деления изображения на строки, путём нахождения белых полос полной ширины. Затем, все полосы делятся на слова, путём нахождения белых полос определенной ширины. И, наконец, последним этапом сегментации является разбиение слов на символы. Далее начинается непосредственно процесс распознавания.

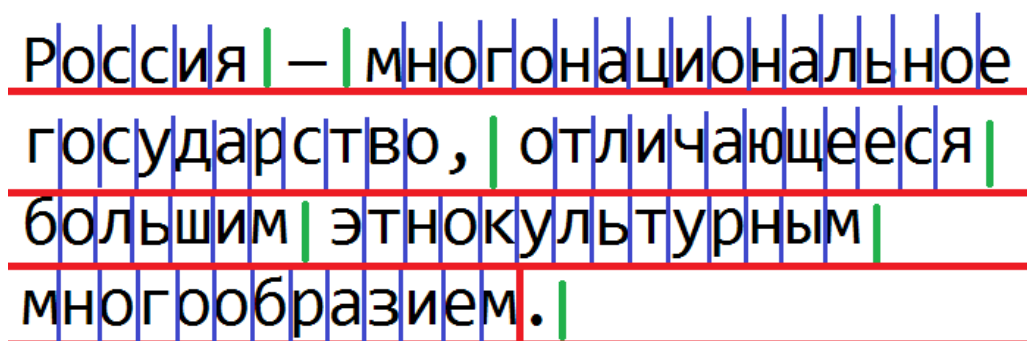


Рис. 2 – Сегментированное изображение.

Рассмотрим первое направление – распознавание с помощью метрик.

Метрика – это определенное условное значение функции, которое определяет объекта в пространстве. Таким образом, если оба объекта будут

похожими друг на друга, например, написаны разным шрифтом, то метрики для таких символов будут похожими или даже будут совпадать.

Для того, чтобы определить букву на изображении нужно определить ее метрику по уже заготовленным шаблонам. Тот шаблон, у которого метрика станет наиболее близкой к 0 и будет ответом.

На практике было определено, что подсчет метрики скорее всего не даст положительного результата, ввиду схожести некоторых букв друг на друга, к примеру «Ь» «ъ».

Экспериментом было выяснено, что такие буквы как «Н» «I» «i» «O» «o» «X» «x» «l» обладают суперсимметрией, из-за этого они были отделены в отдельный класс, что должно будет уменьшать перебор всех метрик в 6 раз. Аналогичные действия были проведены в отношении других букв.

Нейронная сеть. Принцип работы.

Искусственная нейронная сеть (ИНС) — это математическая модель, которая в дальнейшем описывается программно. ИНС построена на основе принципа организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма [3, С. 2].

В настоящее время в области технологии нейронных сетей ведутся обширные работы. Модернизируются и улучшаются методы распознавания и способы обучения ИНС.

Нейронная сеть состоит из нейронов, которые в свою очередь имеют определенное ограниченное количество входов и к каждому из входов соответствует некоторый вес (рис. 3). Задачей нейрона является определить взвешенную сумму своих входов. В данном случае на входе будут пиксели изображения с тем или иным символом. Для того, чтобы нейрон обработал и адекватно сформировал выходной сигнал применяется целевая функция - $out = \phi(net)$. Значение этой функции и является выходом нейрона [3, С. 3].

$out = \phi(\sum_{i=1}^n x_i w_i)$, где ϕ – функция активации, $\sum_{i=1}^n x_i w_i$ – взвешенная сумма, n – произведений входных сигналов на соответствующие веса.

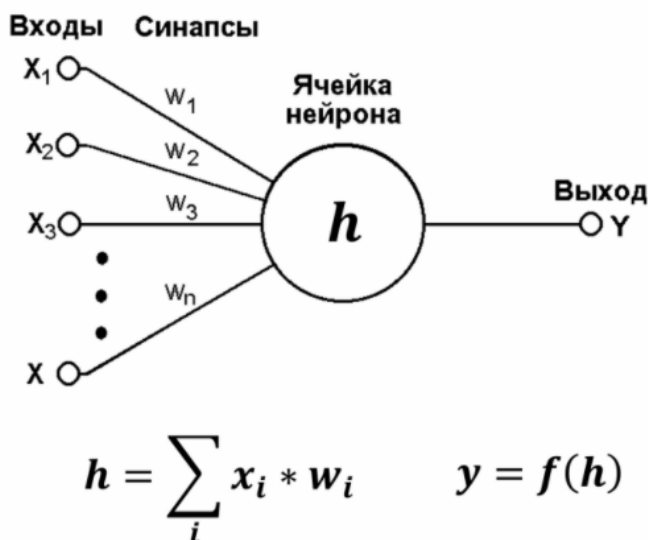


Рис. 3 – Устройство нейрона

В свою очередь нейронные сети можно классифицировать по способу обучения на ИНС с учителем и ИНС без учителя. В первом случае обучения присутствует учитель, который подает на вход нейронной сети образы, производит сравнение результирующих выходов с требуемыми, а затем настраивает веса ИНС сети таким образом, чтобы уменьшить различия. Трудно представить такой обучающий механизм в биологических системах; следовательно, хотя данный подход привел к большим успехам при решении прикладных задач, он отвергается теми исследователями, кто полагает, что искусственные нейронные сети обязательно должны использовать те же механизмы, что и человеческий мозг.

Во втором случае обучение происходит без учителя. В результате обработки входных образов сеть самоорганизуется, в результате чего происходит настройка весов согласно определенному алгоритму. Требуемый выход в процессе обучения не указан, поэтому результаты определения возбуждающих образов для конкретных нейронов непредсказуемы. При этом, однако, сеть организуется в форме, отражающей существенные

характеристики обучающего набора. Например, входные образы могут быть классифицированы согласно степени их сходства так, что образы одного класса активизируют один и тот же выходной нейрон.

Нейронные сети так же подразделяются на однослойные и многослойные.

Однослойная нейронная сеть – это простейшая ИНС, состоящая из группы нейронов, образующих слой (рис. 3). В данной нейронной сети сигналы от входного слоя сразу подаются на выходной слой, который преобразует сигнал, выводя результат.

Многослойные нейронные сети – обладают большими вычислительными способностями. Данная нейронная сеть состоит из входного и выходного слоя, а так же расположенными между ними скрытыми слоями (рис. 4).

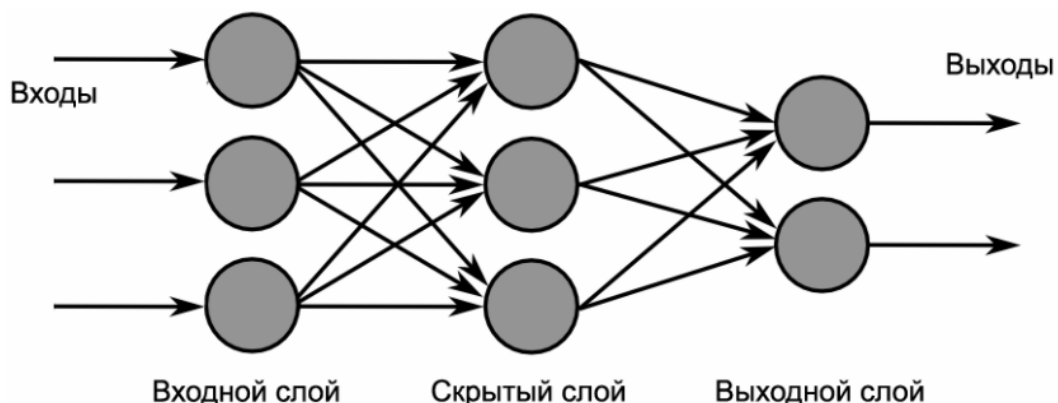


Рис. 4 – Многослойная нейронная сеть

Разработанное программное обеспечение.

В среде Microsoft Visual Studio 2012 разработано программное обеспечение на основе простейшей однослойной нейронной сети для распознавания цифр от 0 до 9. Структурная схема алгоритма работы предложенной программы приведена на рис. 5.

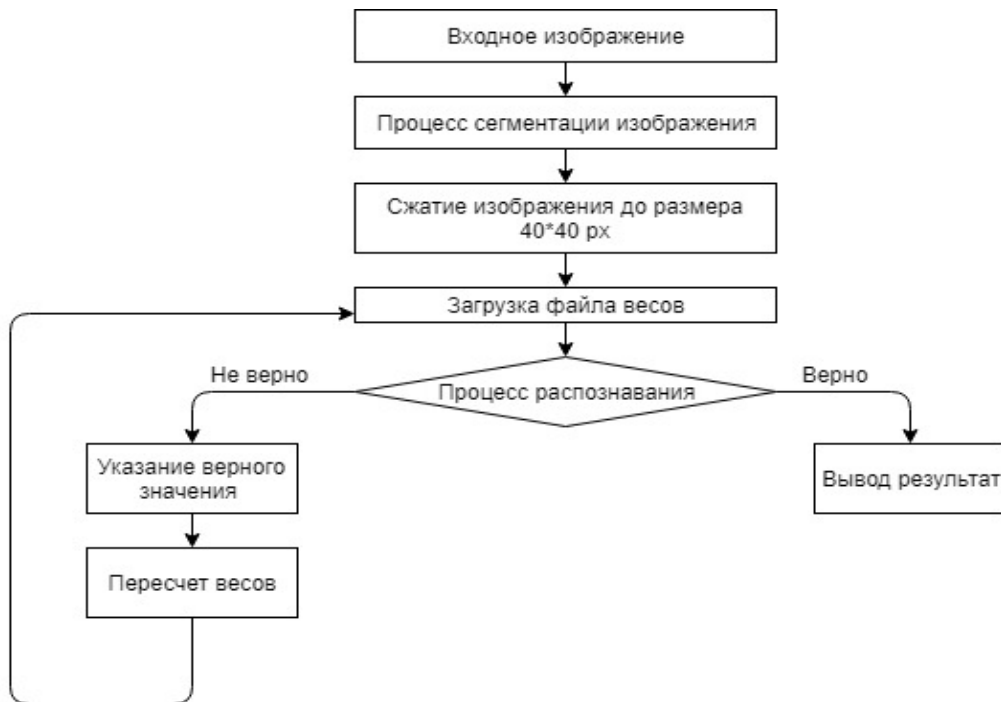


Рис. 5 - Структурная схема работы алгоритма

Интерфейс программы приведен на рис. 6. Изображенные на рис. 6 кнопки предназначены для:

«*Getweight*» – загрузка файла весов, «*train*» и «*correct*» – тренировка нейронной сети или коррекция весов ИНС, «*clear*»– визуальная очистка формы, «*open*» –загрузка изображения с символом, «*recognize*»– распознавание изображения.

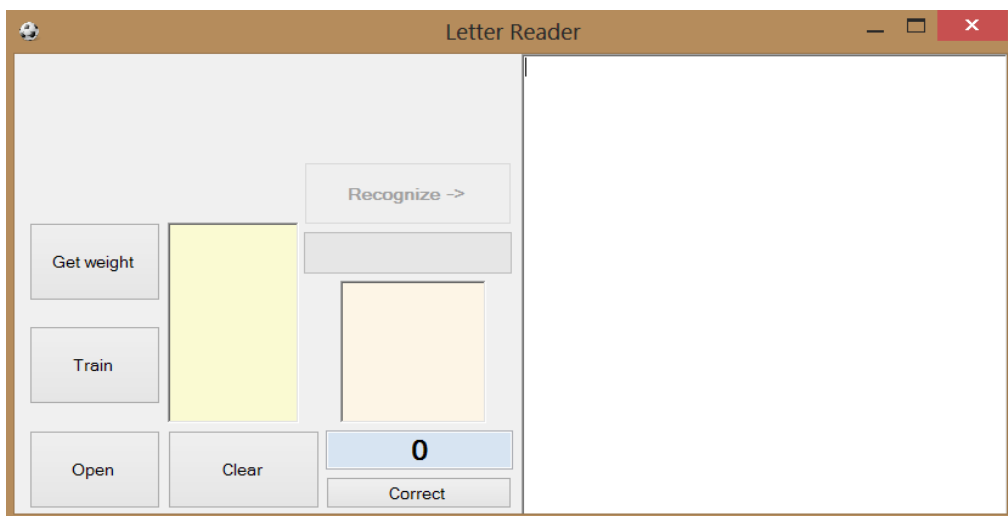


Рис. 6 – Программное обеспечение для распознавания численных символов

В данном программном обеспечении используется 10 нейронов по количеству цифр. При считывании изображения каждому пикселю входное значение 0 или 1 в зависимости от интенсивности цвета пикселя. Далее для каждого нейрона происходит процесс умножение веса на входное значение. Позже результат весов передаются на сумматор, где для каждого нейрона определяется сумма весов. Функцию активации – функция единичного скачка. Допустим, если взвешенная сумма больше определенного порога b , то выход нейрона равен 1. Если ниже, то 0(рис. 7).

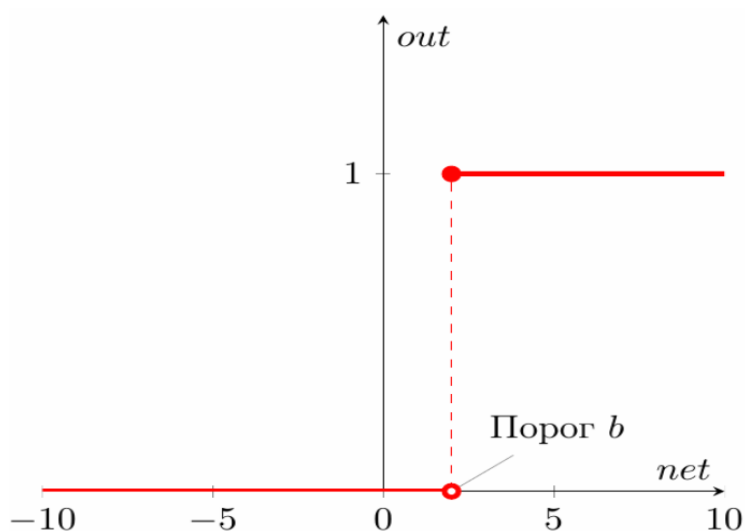


Рис. 7 – Функция единичного скачка

Данная простейшая однослойная нейронная сеть была обучена с помощью учителя. В процессе обучения с учителем, программе подается на вход изображение с цифрой и в случае неверного результата, происходит корректировка (перерасчет) весов путем указания правильного результата учителем. Процент удачного распознавания равен 92%.

Результат работы программного обеспечения для распознавания печатного символа «4» представлен на рисунке 8.

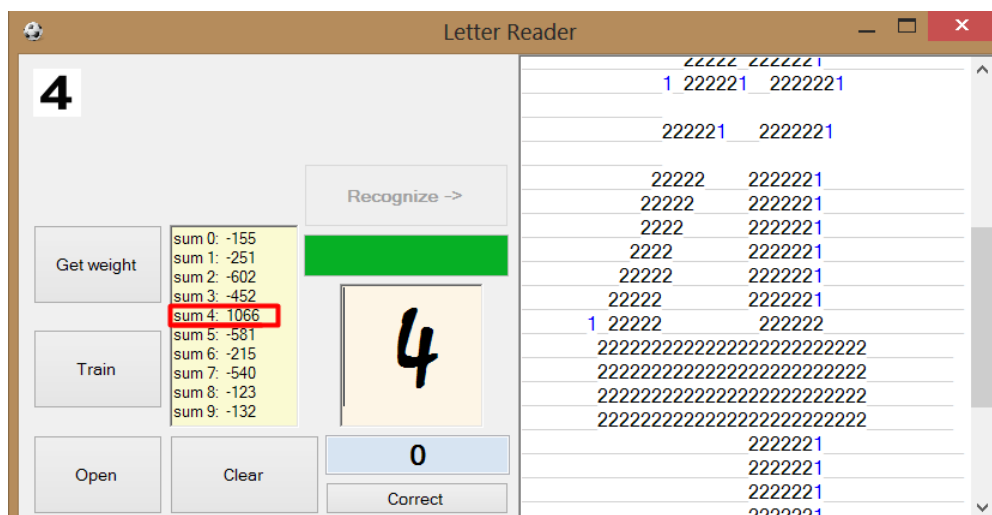


Рис. 8 – Результат работы программного обеспечения

Далее программе были предоставлены обучающие выборки с изображениями цифр рукописного ввода. Количество предоставленных выборок и процент верного распознавания представлены на графике 1. Ввиду своей сложной структуры процент верного распознавания значительно уменьшился. Для повышения точности распознавания необходимо увеличение обучающих выборок.

Результат работы программного обеспечения для распознавания каллиграфических символов представлен на рисунке 9.

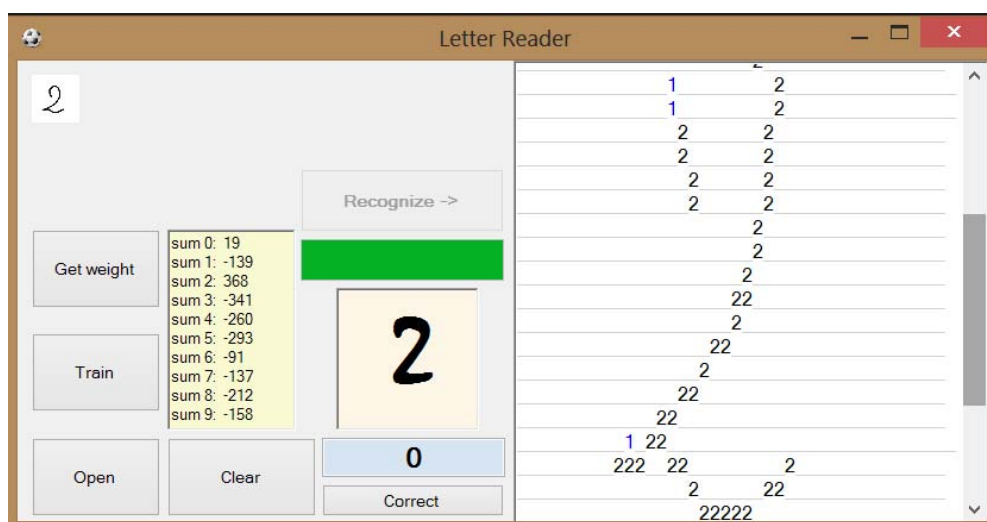


Рис. 9 – Результат работы программного обеспечения

В левом верхнем углу расположено сжатое изображение символа «4». Следует заметить, что максимальный вес находится у нейрона, отвечающий за символ «4» (выделено красным прямоугольником), соответственно результатом является цифра «4».

В правой части программы для наглядности отображена интенсивность оттенков пикселей входящего изображения рис.8, рис.9.

Разработанная программа позволяет загружать фотографии формата jpeg, png,jpg и размерности до 1024*1024. Вычислительный эксперимент производился на компьютереASUSN56(Inte® Core™ i7-3630QM, CPU @2.40 GHz, ОЗУ 8 ГБ, 64 – разрядная ОС), под управлением ОС Windows 8.

Таблица – Анализ обучаемости нейронной сети

	Количество представленных ИНС объектов для изучения (для каждого символа)						
	10	20	40	70	100	200	300
% верного распозн-ия печатных символов	≈9	≈14	≈17	≈26	≈38	≈65	≈82
% верного распозн-ия каллигр-их символов	≈1	≈2	≈3	≈6	≈9	≈18	≈29

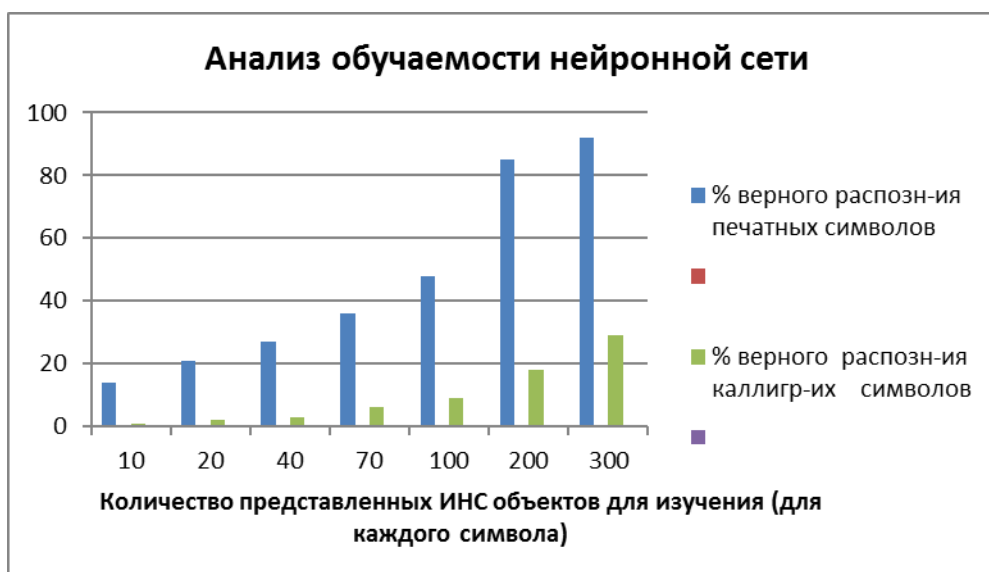


График 1 - Анализ обучаемости нейронной сети

За оптимальный размер изображения, в котором будет распознаваться символ, был взят 40*40px. Программа автоматически трансформирует изображение в указанный выше размер. Если символ занимает меньше 80% изображения $X < 80\%$, то происходит процесс сегментации – выделение символа, обрезание изображения, так чтобы символ занимал больше 80% изображения $X > 80\%$, рис. 9. Ввиду вышеприведенного алгоритма сегментации можно сделать вывод, что скорость распознавания не зависит от размерности изображения.



Рис. 9 – Сегментация символа $X < 80\%$

Заключение

- 1) Было реализовано программное обеспечение на основе простейшей однослойной нейронной сети для распознавания цифровых символов.
- 2) Реализован метод, позволяющий сегментировать конкретный символ с последующим сжатием до входного размера – 40 * 40 px.
- 3) В результате анализа обучаемости простейшей однослойной нейронной сети было установлено, что для удовлетворительного результата (свыше 90%) необходимо как минимум 300 обучающих выборок на один символ.

- 4) Из 1 графика видно, что качество распознавания нейронной сети каллиграфических символов меньше, а, следовательно есть необходимость увеличение итераций обучения.
- 5) К недостаткам нейронной сети можно отнести: достаточно долгий период обучения, высокие требования производительности, сложность в реализации. К достоинствам относятся: при достаточно долгом обучении – вероятность верного ответа, стремящийся к 99%, способность распознавать нестандартные образы, большой интерес в мире к данным технологиям.
- 6) К недостаткам алгоритмов для распознавания символов можно отнести: подготовка шаблонов и алгоритмов для определенных видов символов, отсутствие способности к обучению. К достоинствам относятся: скорость распознавания.

Литература

1. Афонасенко А.В., Елизаров А.И. Обзор методов распознавания структурированных символов. - М.: Известия Южного федерального университета. Технические науки, 2015 -с. 83 - 88.
 2. Шалымов Д. С. Автоматическое распознавание печатных текстов арабского языка. – М.: Издательство Санкт-Петербургского государственного университета (Санкт-Петербург), 2007 -с.124-137.
 3. Петров С.П. Сверточная нейронная сеть для распознавания символов номерного знака автомобиля. – М.: Системный анализ в науке и образовании, 2013. – с. 1 – 8.
 4. Артемьев И.С., Лебедев А.И., Долгий А.И., Хатламаджиян А.Е., Меерович В.Д. Метод блочного оптического распознавания инвентарных номеров железнодорожных подвижных единиц на основе комитетной нейроиммунной модели классификации // Инженерный вестник Дона, 2014, №1 URL:ivdon.ru/ru/magazine/archive/n1y2014/2259.
-

5. Кулакович А.Ю. Венцов Н.Н. Краткий обзор и программная реализация избранных методов для деконволюции изображений // Инженерный вестник Дона, 2017, №4 URL: ivdon.ru/ru/magazine/archive/n4y2017/4468.
6. Luoting Fu, Levent Burak Kara. Neural network-based symbol recognition using a few labeled samples. Computers&Graphics. 2011; pp. 955 – 966.
7. Gennari L, Kara LB, Stahovich TF. Combining geometry and domain knowledge to interpret hand-drawn diagrams. Computers&Graphics. 2005; pp. 547– 562.
8. Э.М. Браверман. Автоматический анализ сложных изображений. - М.: Мир, 1969. - 310 с.
9. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005. –с. 1072
10. Андерсон Т. Статистический анализ временных рядов. – М.: Мир, 1976. – с.756

References

1. Afonassenko A.V., Elizarov A.I. Izvestiya Yuzhnogo federal'nogo universiteta. Tekhnicheskie nauki, 2015; pp. 83 – 88.
2. Shalymov D. S. Avtomaticheskoe raspoznavanie pechatnykh tekstov arabskogo yazyka [Automatic recognition of printed texts of the Arabic language]. Izdatel'stvo Sankt-Peterburgskogo gosudarstvennogo universiteta (Sankt-Peterburg), 2007; pp.124-137.
3. Petrov S.P. Svertochnaya neyronnaya set' dlya raspoznavaniya simvolov nomernogo znaka avtomobilya [Convolutional neural network for recognition of license plate symbols of a car]. Sistemnyy analiz v nauke i obrazovanii, 2013; pp. 1 – 8.



4. Artem'ev I.S., Lebedev A.I., Dolgiy A.I., Khatlamadzhyan A.E., Meerovich V.D. Inzhenernyj vestnik Dona (Rus), 2014, №1. URL: ivdon.ru/ru/magazine/archive/n1y2014/2259.
5. Kulakovich A.Yu. Ventsov N.N. Inzhenernyj vestnik Dona (Rus), 2017, №4. URL: ivdon.ru/ru/magazine/archive/n4y2017/4468.
6. Luoting Fu, Levent Burak Kara. Neural network-based symbol recognition using a few labeled samples. Computers & Graphics. 2011; pp. 955 – 966.
7. Gennari L, Kara LB, Stahovich TF. Combining geometry and domain knowledge to interpret hand-drawn diagrams. Computers & Graphics. 2005; pp. 547 – 562.
8. E.M. Braverman. Avtomaticheskij analiz slozhnykh izobrazheniy [Automatic analysis of complex images]. Mir, 1969. 310p.
9. Gonsales R., Vuds R. Tsifrovaya obrabotka izobrazheniy [Digital image processing]. Tekhnosfera, 2005. 1072 p.
10. Anderson T. Statisticheskij analiz vremennykh ryadov [Statistical analysis of time series]. Mir, 1976. 756 p.