

УДК 004.891.3

# НЕЙРОСЕТЕВАЯ МОДЕЛЬ РАСПОЗНАВАНИЯ РУКОПИСНЫХ СИМВОЛОВ ДЛЯ ПОСТРОЕНИЯ СИСТЕМ БИОМЕТРИЧЕСКОЙ АУТЕНТИФИКАЦИИ

**Катасёв А.С.**, д-р техн. наук, доцент

**Катасёва Д.В.**, ассистент

Казанский национальный исследовательский технический университет  
им. А.Н. Туполева-КАИ (КНИТУ-КАИ)

Контакты: [kat\\_726@mail.ru](mailto:kat_726@mail.ru)

*Данная статья посвящена разработке нейросетевой модели распознавания десятичных рукописных символов. В основу модели положена специально разработанная сверточная нейронная сеть. Построение модели проводилось в среде MatLab. Тестирование и оценка адекватности модели показали ее эффективность и практическую пригодность для построения систем биометрической аутентификации.*

**Ключевые слова:** нейронная сеть, распознавание рукописных символов, биометрическая аутентификация.

## Введение

В эпоху развития информационных технологий с ростом значимости информационной безопасности актуальными стали вопросы идентификации и аутентификации личности. Идентификация позволяет пользователю назвать себя, а аутентификация – подтвердить подлинность. В настоящее время широкое распространение получили биометрические системы идентификации, например, распознавание личности по его рукописному почерку [1].

Распознавание рукописных символов заключается в способности программы относить их к одному из заранее определенных классов. Важным вопросом является разработка устойчивых систем, позволяющих выполнять оперативное распознавание и интерпретацию рукописных данных путем считывания как движения ручки (онлайн метод), так и написанного текста (оффлайн метод) [2].

Цель данного исследования – повышение эффективности биометрической идентификации и аутентификации личности за счет использования методов искусственного интеллекта. Проблема, возникающая при распознавании рукописного текста, обусловлена следующими основными факторами:

- в отличие от машинописного текста, рукописный текст не стандартизирован;
- рукописный текст индивидуален и подвержен различным искажениям, а текст, который напечатан, обладает фиксированным размером и очертаниями, свойственным заданному шрифту;
- в рукописном тексте символы пишутся слитно, а в машинописном тексте между ними присутствует интервал;
- текст, написанный от руки, обладает хаотичностью написания, т.е. не соблюдаются строки и столбцы символов.

Таким образом, возникает сложность в распознавании рукописных символов из-за их неоднозначного написания в слове.

### **Анализ методов распознавания символов**

Существует множество методов распознавания символов, в основе которых лежат свои алгоритмы и принципы. К наиболее распространенным из них относятся: шаблонные методы; структурные методы; признаковые методы; методы структурно-пятенных эталонов; нейросетевые методы.

При использовании шаблонных методов изображение символа преобразуется в растр. Его сравнивают со всеми шаблонами из ба-

зы и выбирают тот из них, который имеет наименьшее количество точек, отличающихся от изображения на входе.

Структурные методы представляют объект в виде графа, узлы которого являются элементами входного объекта, а дуги – пространственными отношениями между ними. Такие методы, как правило, работают с векторными изображениями. Линии, которые составляют символ, являются структурными элементами. Распознаваемый символ подвергается процедуре скелетизации (утоньшению). При скелетизации контур каждого изображения представляется в виде набора последовательных «особых» точек и специального кода, включающего базовую точку и множество направлений из текущей точки к другим точкам.

Отличительной особенностью признаковых методов является то, что каждому распознаваемому изображению ставится в соответствие вектор признаков фиксированной конечной размерности. Процесс распознавания строится на сравнении его с набором эталонных векторов той же размерности. При этом задачу системы распознавания символов можно рассматривать как задачу классификации, решение которой основано на расчете евклидова расстояния между вектором описывающих признаков распознаваемого изображения и вектором признаков эталонного изображения. На качество распознавания влияет и тип, и количество признаков. Текущий вектор формируется при анализе предварительно подготовленного изображения. Этот процесс называется извлечением признаков. Базовый эталон для каждого распознаваемого класса символов формируется аналогично путем обработки изображения символов в обучающей выборке данных.

Метод структурно-пятенных эталонов требует представления каждого изображения в виде множества пятен, каждое из которых связано со всеми остальными пятнами  $n$ -арными отношениями. Это определяет внутреннюю структуру каждого распознаваемого символа. Формируемые отношения структуры расположения пятен об-

разуют структурные элементы распознаваемого изображения. К основным типам  $n$ -арных отношений между двумя пятнами относятся отрезок, эллипс и дуга. Остальные отношения определяют пространственное расположение элементов, составляющих распознаваемый символ.

Для принятия решения при распознавании рукописного текста актуально использовать интеллектуальные системы [3; 4], например, нейронные сети, решающие задачу классификации объектов [5]. Существует множество примеров эффективного использования нейросетевых методов в решении прикладных задач [6; 7; 8].

Алгоритмы нейросетевого распознавания символов, как правило, характеризуются следующими особенностями [9]:

- растр распознаваемого изображения приводится к определенному стандартному нормализованному размеру (как правило, используется маска  $16 \times 16$  пикселей и более);
- уровни яркости каждой точки нормализованного растра используются в качестве значений входных параметров нейросетевой модели;
- количество выходных нейронов в нейросетевой модели соответствует количеству различных символов, подлежащих распознаванию;
- в качестве результата нейросетевого распознавания выступает символ, которому соответствует максимальное значение на выходе нейросетевой модели.

В ходе проведения анализа представленных методов наиболее эффективными оказались нейросетевые методы распознавания символов. Нейронные сети, в отличие от статистических методов, базируются на концепции параллельной обработки информации, что определяет их быстродействие. Кроме того, нейронные сети обладают способностью к самообучению и адаптации к анализируемым данным, что определяет их эффективность при распознавании рукописных символов.

### Реализация нейросетевой модели

Для реализации нейросетевой модели распознавания десятичных рукописных символов была выбрана среда моделирования MatLab, так как этот программный продукт обладает необходимым набором инструментов для создания нейронных сетей, а именно включает пакет Neural Network Toolbox. Для решения поставленной задачи была выбрана сверточная нейронная сеть [10].

В основе данного типа нейронных сетей лежат три главных механизма:

- локальное извлечение анализируемых признаков;
- построение слоев, состоящих из набора карт признаков;
- формирование выборок данных для обучения нейронной сети.

Данные механизмы позволяют получать правильный результат распознавания, несмотря на возможные искажения символов, связанные с масштабированием.

Рассмотрим структуру разработанной сверточной нейронной сети (рис. 1).

На вход нейронной сети подается отцентрированное изображение каждого распознаваемого символа. Данная операция требуется для характерных признаков изображения (дуг, концевых точек), которые должны находиться в центре во время извлечения значений признаков более высокого порядка. В предлагаемой сверточной нейронной сети исходные изображения символов располагались в центре изображения размером  $32 \times 32$  пикселя. Данная нормализация требуется для ускорения сходимости алгоритма обучения нейросетевой модели.

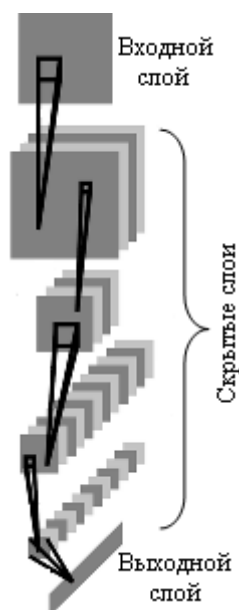


Рис. 1. Структура сверточной нейронной сети

Первый скрытый слой нейронной сети является сверточным и содержит шесть карт признаков размерностью  $28 \times 28$ . Каждый нейрон данного слоя соединен с ограниченной областью входного изображения размером  $5 \times 5$ . Общее число связей составляет 122304, что соответствует 156 настраиваемым параметрам. Эффективное использование памяти и вычислительных ресурсов в данном случае достигается за счет одновременного использования весовых коэффициентов в пределах одной карты.

Второй скрытый слой нейронной сети, являющийся слоем подвыборки, состоит из шести карт признаков размерностью  $14 \times 14$ . Каждый из элементов карт данного слоя связан с областью размером  $2 \times 2$  в соответствующей карте признаков первого слоя. Задачей данного слоя является формирование подвыборки на основе локального усреднения. При этом не требуется хранение карт с

четырьмя весами и сдвигами, а достаточно хранить один общий вес и сдвиг. Следовательно, данный скрытый слой нейронной сети содержит 5880 связей и 12 настраиваемых параметров.

Третий скрытый слой нейронной сети является сверточным и содержит шестнадцать карт признаков размерностью  $10 \times 10$ . Все элементы в каждой карте связаны с несколькими областями размерностью  $5 \times 5$  соответствующих карт предыдущего слоя.

В четвертом слое нейронной сети формируются подвыборки из шестнадцати карт признаков размерностью  $5 \times 5$ . Функционирование данного слоя аналогично функционированию второго скрытого слоя. Все его элементы связаны с соответствующими областями размерностью  $2 \times 2$  предыдущего слоя. Данный слой содержит 2000 связей и 32 настраиваемых параметра.

Пятый слой нейронной сети является полносвязным сверточным слоем, содержащим 120 элементов. Каждый элемент соединен областями размерностью  $5 \times 5$  со всеми картами предыдущего слоя. Данный слой является гибридным, так как содержит сверточные и полносвязные элементы. Общее число настраиваемых параметров данного слоя равно 48120.

Шестой слой нейронной сети содержит 84 нейрона и тоже является полносвязным. Число настраиваемых параметров слоя составляет 10164.

Последний слой содержит десять выходных нейронов значения на выходе которых определяют распознаваемый рукописный символ.

После построения архитектуры нейронной сети ее необходимо обучить и протестировать. В качестве исходных данных была использована база данных MNIST [11], которая содержит 60000 обучающих и 10000 тестовых изображений. Изображения нормализованы по размеру и отцентрованы.

На рис. 2 представлено окно ввода десятичного рукописного символа на базе реализованной в *MatLab* модели сверточной нейронной сети.

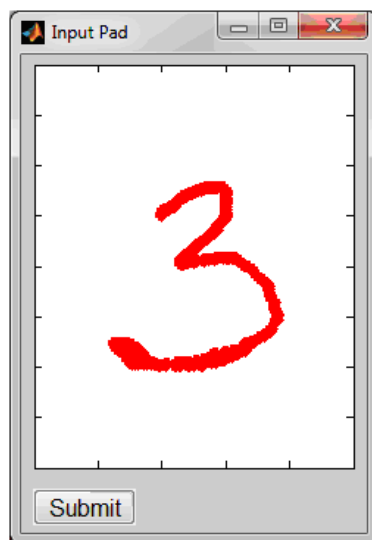


Рис. 2. Пример введенного символа для распознавания

Из данного рисунка видно, что изображение десятичного рукописного символа вводится внутри матрицы размерностью  $5 \times 8$  пикселей. Однако, в реализованной системе возможно задание любой битовой маски, что определяет ее гибкость и возможность производить экспериментальные исследования для тестирования и оценки эффективности модели сверточной нейронной сети при распознавании десятичных рукописных символов.

### **Тестирование и оценка эффективности нейросетевой модели**

Изучение характеристик построенной модели сверточной нейронной сети требуется для оценки ее устойчивости, т.е. способ-



ности правильно классифицировать входные образы в условиях изменения характеристик вводимых символов.

Для практического использования построенной нейросетевой модели требуется определение ее адекватности, т.е. соответствие тому, насколько точно она решает поставленную задачу распознавания десятичных рукописных символов.

В таблице 1 представлены типовые результаты тестирования обученной сверточной нейронной сети на изображениях десятичных символов, сформированных пользователем.

Таблица 1. Результат тестирования модели на исходных изображениях

Символ	Результат распознавания	Время распознавания, секунд
0	верно	0,89
1	верно	0,77
2	верно	0,28
3	верно	0,72
4	верно	0,83
5	верно	0,68
6	верно	0,94
7	верно	0,95
8	верно	0,61
9	верно	0,93

Как видно из представленной таблицы, все изображения сверточная нейронная сеть распознает верно. Причем время распознавания любого символа составляет менее 1 секунды.

Для дальнейшей оценки эффективности модели сверточной нейронной сети произведем искажения распознаваемых символов. В качестве искажений используем изменение масштаба вводимого

символа (уменьшение или увеличение), а также изменение наклона (влево или вправо).

На рис. 3 представлены примеры искаженных изображений распознаваемых символов.



Рис. 3. Примеры искаженных изображений

Как видно из рисунка, рукописная цифра «пять» представлена в четырех вариантах, отличающихся по величине и углу наклона. Подобные искажения применялись ко всем входным изображениям.

В таблице 2 представлены результаты тестирования нейросетевой модели на искаженных изображениях. Как видно из таблицы, не все изображения сверточная нейронная сеть смогла распознать правильно.

Таблица 2. Результат тестирования модели на искаженных изображениях

Символ	Вид искажения	Результат распознавания	Время распознавания, секунд
0	уменьшение	неверно	0,59
	увеличение	верно	0,64
	наклон вправо	неверно	0,62
	наклон влево	неверно	0,51
1	уменьшение	верно	0,97
	увеличение	верно	0,59
	наклон вправо	неверно	0,72
	наклон влево	верно	0,34
2	уменьшение	верно	0,84
	увеличение	верно	0,57
	наклон вправо	верно	0,72
	наклон влево	верно	0,87
3	уменьшение	верно	0,63
	увеличение	верно	0,42
	наклон вправо	верно	0,79
	наклон влево	неверно	0,52
...	...	...	...
9	уменьшение	верно	0,57
	увеличение	верно	0,63
	наклон вправо	верно	0,48
	наклон влево	неверно	0,59

На рис. 4 представлены данные о верно распознанных рукописных символах при различных вариантах их искажений.

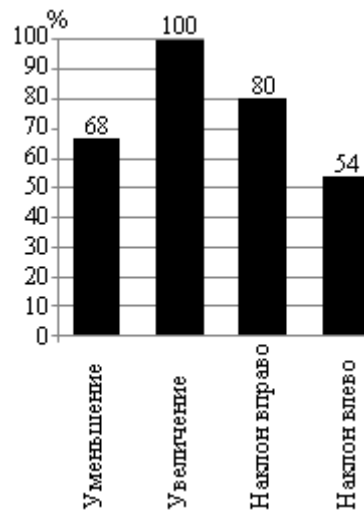


Рис. 4. Гистограмма верно распознанных рукописных символов

Из гистограммы видно, что искажение, связанное с увеличением символа, не повлияло на результат нейросетевого распознавания. При этом такие искажения, как уменьшение и наклон символа, повлияли на количество правильно распознанных символов. При уменьшении правильно распознано только 68% от общего числа изображений, при наклоне вправо – 80%, влево – 54%.

Рассчитаем общую классифицирующую способность  $K$  нейросетевой сверточной модели по следующей формуле:

$$K = \frac{N_{true}}{N} \times 100\%,$$

где  $N_{true}$  – число верно распознанных символов,  $N$  – общее число распознаваемых символов.

Подставляя численные значения, получим:

$$K = \frac{302}{400} \times 100\% = 75,5\%.$$

Таким образом, общая эффективность разработанной нейросетевой модели распознавания рукописных символов составила 75,5%.

### **Заключение**

Как показали результаты проведенных экспериментальных исследований, предложенная нейросетевая сверточная модель имеет высокую эффективность. На исходных изображениях десятичных рукописных символов модель показала 100%-ю эффективность. При этом оценка классифицирующей способности модели на искаженных изображениях составила 75,5%. Несмотря на некоторое снижение эффективности распознавания, можно утверждать, что в целом работа построенной модели является эффективной. Это позволяет использовать разработанную модель в системах биометрической аутентификации.

### **Источники**

1. Зинин А.М. Биометрические системы и идентификация человека // Lex Russica. 2012. Т. LXXI. № 5. С. 1064-1070.
2. Солошенко А.Е., Витовтова О.В., Жилинкова Л.А. Обзор статических методов биометрической идентификации личности // Российская наука и образование сегодня: проблемы и перспективы. 2014. № 3 (2). С. 78-80.
3. Катасёв А.С., Катасёва Д.В. Формирование нечетких правил фильтрации нежелательных электронных сообщений в инфокоммуникационных сетях / Проблемы техники и технологий телекоммуникаций ПТиТТ-2014. Оптические технологии в телекоммуникациях ОТТ-2014. Материалы Международных на-учно-технических конференций. Казань. 2014. С. 320-322.
4. Катасёв А.С., Катасёва Д.В., Кирпичников А.П. Оценка стойкости шифрующих преобразований моноалфавитной замены с использованием генетического алгоритма // Вестник технологического университета. 2015. Т. 18. № 7. С. 255-259.
5. Катасёв А.С., Катасёва Д.В. Разработка нейросетевой системы классификации электронных почтовых сообщений // Вестник Казанского государственного энергетического университета. 2015. № 1 (25). С. 68-78.

6. Катасёв А.С., Катасёва Д.В., Кирпичников А.П. Нейросетевая диагностика аномальной сетевой активности // Вестник технологического университета. 2015. Т. 18. № 6. С. 163-167.
7. Катасёв А.С., Катасёва Д.В., Кирпичников А.П., Евсеева А.О. Нейросетевая модель идентификации ботов в социальных сетях // Вестник технологического университета. 2015. Т. 18. № 16. С. 253-256.
8. Катасёв А.С., Катасёва Д.В., Кирпичников А.П., Семенов Я.Е. Спам-фильтрация электронных почтовых сообщений на основе нейросетевой и нейронечеткой моделей // Вестник технологического университета. 2015. Т. 18. № 15. С. 217-220.
9. Иванов А.И. Нейросетевые алгоритмы биометрической идентификации личности. Кн. 15: монография. М.: Радиотехника, 2004. 144 с.: ил.
10. Петров С.П. Сверточная нейронная сеть для распознавания символов номерного знака автомобиля // Системный анализ в науке и образовании. 2013. № 3 (21). С. 66-73.
11. Kussul E., Baidyk T. Improved method of handwritten digit recognition tested on MNIST database // Image and Vision Computing. 2004. Т. 22. N 12 SPEC. ISS. P. 971-981.

**THE NEURAL NETWORK HANDWRITING RECOGNITION MODEL FOR  
BIOMETRIC AUTHENTICATION SYSTEMS DEVELOPMENT**

**Katasev A.S., Kataseva D.V.**

*This article is developed the neural network decimal handwriting recognition model. This model is based on a specially designed convolutional neural network. Construction of the model was carried out in MatLab environment. Testing and evaluation of the adequacy of the model indicated its efficiency and practical suitability for biometric authentication systems using.*

**Keywords:** *neural network, handwriting recognition, biometric authentication.*

Дата поступления 15.05.2016.