

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ**  
**БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ**  
**Факультет прикладной математики и информатики**  
**Кафедра многопроцессорных систем и сетей**

**ПЕРЕСТОРОНИН ПАВЕЛ СЕРГЕЕВИЧ**  
**КАЛАШНИКОВ ИГОРЬ АНАТОЛЬЕВИЧ**

**Применение нейронных сетей для распознавания текста**

Курсовой проект

студентов 4 курса 1 группы

«Допустить к защите»  
**Руководитель работы**

«\_\_» \_\_\_\_\_ 2011 г

**Руководитель**  
*Гусейнова Анастасия Сергеевна*  
ассистент кафедры МСС

**Минск 2011**

## СОДЕРЖАНИЕ

Содержание .....	2
1 Введение.....	3
2 Проблема.....	4
3 Общие сведения о нейронных сетях .....	5
3.1 Основные задачи, решаемые с помощью нейронных сетей.....	5
3.1.1 Классификация образов. ....	5
3.1.2 Кластеризация/категоризация. ....	5
3.1.3 Аппроксимация функций.....	5
3.1.4 Предсказание/прогноз. ....	5
3.1.5 Оптимизация. ....	5
3.2 Основные компоненты нейронной сети .....	5
3.3 Архитектуры нейронных сетей .....	6
3.3.1 Однослойные сети прямого распространения .....	6
3.3.2 Многослойные сети прямого распространения.....	6
3.3.3 Рекуррентные сети.....	6
4 Механизм распознавания символов .....	7
5 Текущая реализация.....	9
6 Заключение .....	10
7 Список использованных источников .....	11

## 1 ВВЕДЕНИЕ

В последнее время в мире очень большое внимание уделяется технологиям распознавания изображений и в частности распознавания текста. Сложно представить сферу человеческой деятельности, где бы ни могли найти применение эти технологии – так как информация сейчас собирается в первую очередь именно в виде фотоснимков и видео. И собранную таким образом информацию часто необходимо приводить к иному виду в целях хранения, демонстрации или поиска.

В настоящее время распознавание обширно используется в целях обеспечения безопасности. Так, распознаванию подвергаются кадры, на которых запечатлены предполагаемые преступники и нарушители. Идет поиск лиц по базам данных, что помогает раскрывать преступления. Также распознаванию подвергаются автомобильные номера, что позволяет отыскивать угнанные автомобили и разыскивать скрывающихся на них людей.

Распознавание текста используется во многих приложениях, целевой аудиторией которых является среднестатистический пользователь. Такие приложения умеют распознавать текст с фотографий, отсканированные документы и текст на рисунках. Эта информация может использоваться для формирования текстовых документов в различных форматах из исходных изображений, может особым образом сохраняться в базу данных для дальнейшего поиска текста. Такие программы и программные комплексы позволяют также осуществлять перевод целых библиотек из книжного формата в формат электронный, что увеличивает доступность литературы для любого человека, имеющего доступ к базе данных оцифрованных книг.

Для распознавания текста сейчас существует огромное число технологий и алгоритмов – это серьезная проблема, затрагивающая области распознавания образом, компьютерного зрения и искусственного интеллекта. Очень многие из алгоритмов основаны на использовании нейронных сетей, которые позволяют обучать себя и постоянно повышать качество распознавания изображений и текста.

В то же время в распознавании текста есть очень много нерешенных проблем и проблем, которые в ближайшее время решить не представляется возможным. Так, человек распознает символы на основе многолетнего опыта, сравнивая картинку с множеством виденных ранее по очень сложным алгоритмам. И все равно часто человек не может понять написанное другим. Компьютер же не обладает таким опытом и такой вычислительной мощностью, как человеческий мозг – потому ошибки появляются достаточно часто.

До сих пор нет точного способа определения слов, символов по определенным изображениям – и такие алгоритмы, вероятно, не существуют из-за сложности и неоднозначности человеческого письма. Без анализа почерка конкретного человека часто невозможно сказать, какой символ он писал.

## 2 ПРОБЛЕМА

В качестве направления для исследований нами выбрано оптическое распознавание, которое в настоящее время является очень актуальным. Развитие и распространение компьютерной обработки информации привели к возникновению потребностей в технологиях, позволяющих машинам осуществлять распознавание в обрабатываемой ими информации.

Разработка методов машинного распознавания позволяет расширить круг выполняемых компьютерами задач и сделать машинную переработку информации более интеллектуальной. Несмотря на то, что некоторые из этих задач распознавания решаются человеком на подсознательном уровне с большой скоростью, до настоящего времени ещё не создано компьютерных программ, решающих их в столь же общем виде. Существующие системы предназначены для работы лишь в специальных случаях со строго ограниченной областью применения.

Примерами задач распознавания могут служить:

- распознавание букв;
- распознавание автомобильных номеров;
- распознавание лиц;
- распознавание речи.

Очевидно, что многие из этих задач могут иметь широкое практическое применение. Например, с помощью системы распознавания автомобильных номеров вы без труда автоматизируете въезд на любую территорию, будь то стоянка у торгового центра, территория промышленного предприятия или платная автострада.

Установка системы распознавания на въездах и выездах позволит контролировать присутствие транспортных средств на территории. Как владелец парковки, вы получите полную статистическую информацию о ее загруженности, проанализировав которую сможете принять меры по повышению эффективности вашего бизнеса.

На сегодняшний день в основе наиболее распространенных технологий верификации и идентификации лежит использование паролей и персональных идентификаторов. Однако такие системы слишком уязвимы и могут легко пострадать от подделки, воровства и просто человеческой забывчивости. Поэтому все больший интерес вызывают методы биометрической идентификации, позволяющие определить личность человека по его физиологическим характеристикам путем распознавания по образцам. Классический пример биометрии – анализ отпечатков пальцев, а к новейшим технологиям относятся распознавание сетчатки и радужной оболочки глаза, что применяется и для идентификация личности, и для расследования преступлений – так как имеется определенная уникальность этих частей тела..

### 3 ОБЩИЕ СВЕДЕНИЯ О НЕЙРОННЫХ СЕТЯХ

Искусственные нейронные сети (ИНС) строятся по принципам организации и функционирования их биологических аналогов [3]. Они способны решать широкий круг задач распознавания образов, идентификации, прогнозирования, оптимизации, управления сложными объектами. Дальнейшее повышение производительности компьютеров все в большей мере связывают с ИНС, в частности, с нейрокомпьютерами (НК), основу которых составляет искусственная нейронная сеть.

#### 3.1 Основные задачи, решаемые с помощью нейронных сетей

##### 3.1.1 Классификация образов.

Задача состоит в указании принадлежности входного образа, представленного вектором признаков, одному или нескольким предварительно определенным классам. К известным приложениям относятся распознавание букв, распознавание речи, классификация сигнала электрокардиограммы, классификация клеток крови, задачи рейтингования.

##### 3.1.2 Кластеризация/категоризация.

При решении задачи кластеризации, которая известна также как классификация образов без учителя, отсутствует обучающая выборка с образцами классов. Алгоритм кластеризации основан на подобии образов и размещает близкие образы в один кластер. Известны случаи применения кластеризации для извлечения знаний, сжатия данных и исследования свойств данных.

##### 3.1.3 Аппроксимация функций.

Предположим, что имеется обучающая выборка  $((X_1, Y_2), (X_2, Y_2), \dots, (X_N, Y_N))$ , которая генерируется неизвестной функцией, искаженной шумом. Задача аппроксимации состоит в нахождении оценки этой функции.

##### 3.1.4 Предсказание/прогноз.

Пусть заданы  $N$  дискретных отсчетов  $\{y(t_1), y(t_2), \dots, y(t_n)\}$  в последовательные моменты времени  $t_1, t_2, \dots, t_n$ . Задача состоит в предсказании значения  $y(t_n + 1)$  в момент  $t_n + 1$ . Прогнозы имеют значительное влияние на принятие решений в бизнесе, науке и технике.

##### 3.1.5 Оптимизация.

Многочисленные проблемы в математике, статистике, технике, науке, медицине и экономике могут рассматриваться как проблемы оптимизации. Задачей оптимизации является нахождение решения, которое удовлетворяет системе ограничений и максимизирует или минимизирует целевую функцию.

#### 3.2 Основные компоненты нейронной сети

Нейронная сеть является совокупностью элементов, соединенных некоторым образом так, чтобы между ними обеспечивалось взаимодействие. Эти элементы, называемые также нейронами или узлами, представляют собой простые процессоры, вычислительные возможности которых обычно ограничиваются некоторым правилом комбинирования входных сигналов и правилом активизации, позволяющим вычислить входной сигнал по совокупности входных сигналов. Выходной сигнал элемента может посылаться другим элементам по взвешенным связям, с каждой из которых связан

весовой коэффициент или вес. В зависимости от значения весового коэффициента передаваемый сигнал или усиливается, или подавляется.

В модели нейрона можно выделить три основных элемента:

1. Набор синапсов или связей, каждый из которых характеризуется своим весом или силой. В частности, сигнал  $x_j$  на входе синапса  $j$ , связанного с нейроном  $k$ , умножается на вес  $w_{kj}$ .
2. Сумматор складывает входные сигналы, взвешенные относительно соответствующих синапсов нейрона. Эту операцию можно описать как линейную комбинацию.
3. Функция активации ограничивает амплитуду выходного сигнала нейрона. Обычно нормализованный диапазон амплитуд выхода нейрона лежит в интервале  $[0, 1]$  или  $[-1, 1]$ .

### 3.3 Архитектуры нейронных сетей

В общем случае можно выделить три фундаментальных класса нейросетевых архитектур.

#### 3.3.1 Однослойные сети прямого распространения

В многослойной нейронной сети нейроны располагаются по слоям. В простейшем случае в такой сети существует входной слой узлов источника, информация от которого передается на выходной слой нейронов (вычислительные узлы), но не наоборот. Такая сеть называется сетью прямого распространения или ациклической сетью.

#### 3.3.2 Многослойные сети прямого распространения

Другой класс нейронных сетей прямого распространения характеризуется наличием одного или нескольких скрытых слоев, узлы которых называются скрытыми нейронами. Функция последних заключается в посредничестве между внешним входным сигналом и выходом нейронной сети. Такая сеть позволяет выделять глобальные свойства данных с помощью локальных соединений за счет наличия дополнительных синаптических связей и повышения уровня взаимодействия нейронов. Способность скрытых нейронов выделять статистические зависимости высокого порядка особенно существенна, когда размер входного слоя достаточно велик.

Узлы источника входного слоя сети формируют соответствующие элементы шаблона активации (входной вектор), которые составляют входной сигнал, поступающий на нейроны второго слоя. Выходные сигналы второго слоя используются в качестве входных для третьего слоя и т.д. Обычно нейроны каждого из слоев сети используют в качестве входных сигналов выходные сигналы нейронов только предыдущего слоя. Набор выходных сигналов нейронов выходного слоя сети определяет общий отклик сети на данный входной образ, сформированный узлами источника входного слоя.

#### 3.3.3 Рекуррентные сети

Рекуррентная нейронная сеть отличается от сети прямого распространения наличием по крайней мере одной обратной связи. Например, рекуррентная сеть может состоять из единственного слоя нейронов, каждый из которых направляет свой выходной сигнал на входы всех остальных нейронов сети.

## 4 МЕХАНИЗМ РАСПОЗНАВАНИЯ СИМВОЛОВ

В качестве механизма распознавания символов выбраны сверточные нейронные сети – одни из самых эффективных на данный момент нейронных сетей. [2] Они, в отличие от обычных полносвязных сетей, позволяют существенно сэкономить на памяти, необходимой для хранения нейронной сети, так как использует ограниченное число весов соединений.

Идея сверточных нейронных сетей заключается в чередовании сверточных слоев (C-layers), субдискретизирующих слоев (S-layers) и наличии полносвязных (F-layers) слоев на выходе.

Такая архитектура включает в себе 3 основных парадигмы [1]:

- Локальное восприятие.
- Разделяемые веса.
- Субдискретизация.

Локальное восприятие подразумевает, что на вход одного нейрона подается не все изображение (или выходы предыдущего слоя), а лишь некоторая его область. Такой подход позволил сохранять топологию изображения от слоя к слою.

Концепция разделяемых весов предполагает, что для большого количества связей используется очень небольшой набор весов. Т.е. если у нас имеется на входе изображение размерами 32x32 пикселя, то каждый из нейронов следующего слоя примет на вход только небольшой участок этого изображения размером, к примеру, 5x5, причем каждый из фрагментов будет обработан одним и тем же набором.

Суть субдискретизации и S-слоев заключается в уменьшении пространственной размерности изображения. Т.е. входное изображение грубо (усреднением) уменьшается в заданное количество раз. Чаще всего в 2 раза, хотя может быть и не равномерное изменение, например, 2 по вертикали и 3 по горизонтали. Субдискретизация нужна для обеспечения инвариантности к масштабу.

Чередование слоев позволяет составлять карты признаков из карт признаков, что на практике означает способность распознавания сложных иерархий признаков.

Сверточные сети обладают еще одним существенным преимуществом – при распознавании они учитывают топологию входного изображения, извлекая локальные связи между пикселями изображения. Это позволяет увеличить точность распознавания, так как учитывается пространственная организация пикселей входного изображения.

В то же время сверточные сети обладают рядом недостатков, например, малой скоростью обучения. На обучение такой сети для распознавания только цифр уходит до десяти часов, время же обучения для множества символов очень велико.

На рисунке 1 можно увидеть варианты символов для обучения нейронной сети распознаванию символа «3».

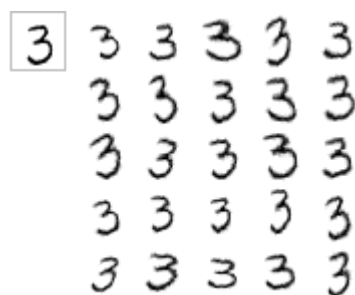


Рисунок 1 – Набор символов для обучения

В данном проекте выбранная нейронная сеть из пяти уровней:

- Входной (нулевой) уровень получает на вход изображение размера  $n \times n$  (точное значение  $n$  будет определено позже, предположительно  $n = 28$ ) в оттенках серого. Соответственно, входной уровень содержит  $n \times n$  нейронов.
- Первый уровень – сверточный уровень с шестью картами признаков. Каждая карта признаков имеет размер  $13 \times 13$ . Таким образом, первый уровень содержит  $13 \times 13 \times 6 = 1014$  нейронов. Каждый элемент карты признаков соединен с областью  $5 \times 5$  на исходном изображении.
- Второй уровень – также сверточный, и содержит 50 карт признаков. Каждая карта признаков имеет размер  $5 \times 5$ , и каждый элемент карты признаков соединен с областью  $5 \times 5$  с предыдущего слоя (уровня) – то есть с областью  $5 \times 5$  на каждой карте предыдущего слоя.
- Третий уровень нейронной сети – полный, и каждый нейрон этого уровня соединен со всеми нейронами предыдущего уровня. На этом уровне находится 100 нейронов, каждый из которых соединен с каждым из 1250 нейронов, находящихся на втором уровне.
- Четвертый уровень – выходной. Это полный уровень с  $m$  нейронами, где  $m$  – число распознаваемых символов. Каждый из нейронов этого уровня соединен с каждым нейроном предыдущего.

На рисунке 2 можно увидеть схему работы нейронной сети при распознавании символа – в данном случае символа «2».

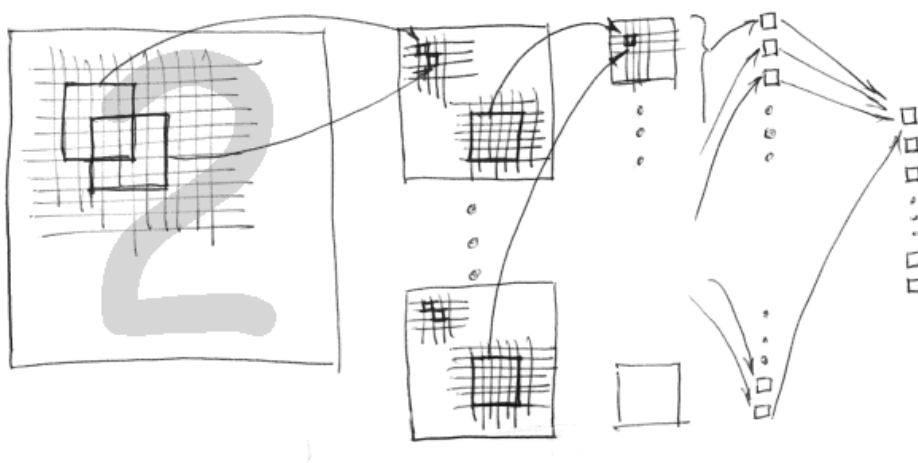


Рисунок 2 – Схема работы сверточной нейронной сети



## 5 ТЕКУЩАЯ РЕАЛИЗАЦИЯ

В текущей реализации последний уровень нейронной сети состоит 10 нейронов, и сеть распознает только символы цифр. Сеть обучена на основе 60000 образцов рукописных символов. При обучении использована уже существующая база изображений MNIST [4]. На рисунке 3 продемонстрирован пример работы этой сети: результат распознавания вместе с промежуточными вычисленными картами признаков.

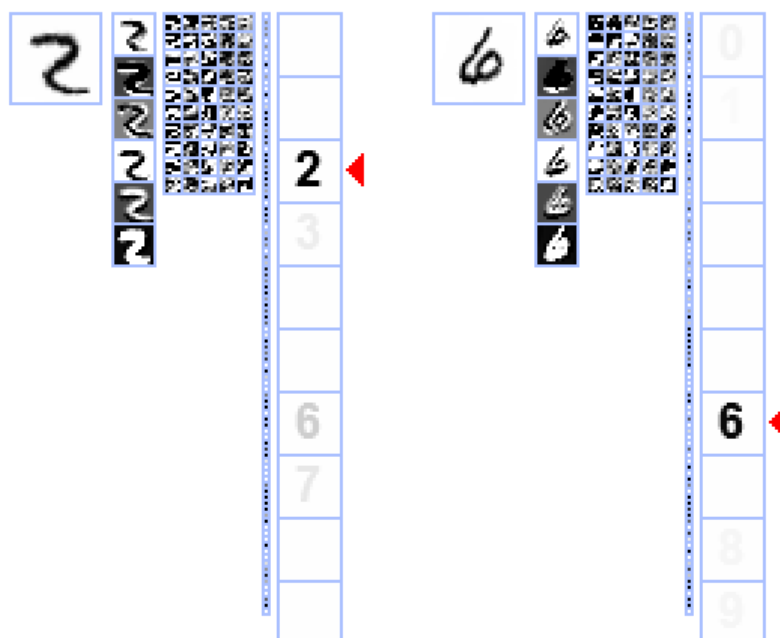


Рисунок 3 – Пример работы сверточной нейронной сети

## 6 ЗАКЛЮЧЕНИЕ

В данном курсовом проекте:

- Поставлена задача распознавания символов с помощью нейронных сетей.
- Изучены основные принципы работы нейронных сетей, их типы.
- Подробно рассмотрен отдельный класс сверточных нейронных сетей.
- Реализовано распознавание рукописных цифр на основе сверточной нейронной сети.

## 7 СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Применение нейросетей в распознавании изображений [Электронный ресурс] / ред. М.Сиротенко. – Habrahabr, 2009. – Режим доступа : [http://habrahabr.ru/blogs/artificial\\_intelligence/74326/](http://habrahabr.ru/blogs/artificial_intelligence/74326/). – Дата доступа: 22.12.2011.
2. Neural Network for Recognition of Handwritten Digits [Electronic resource] / ed. Mike O'Neill. – The Code Project, 2006. – Mode of access : <http://www.codeproject.com/KB/library/NeuralNetRecognition.aspx>. – Date of access : 23.01.2011.
3. Каширина И. Л. Нейросетевые технологии. Учебно-методическое пособие для вузов. – Воронеж: Издательско-полиграфический центр Воронежского государственного университета, 2008. – 72 с.
4. MNIST handwritten digit database [Electronic resource] / ed. Yann LeCun, Corinna Cortes. – Yann LeCun's Home Page, 2004. – Mode of access : <http://yann.lecun.com/exdb/mnist/index.html>. - Date of access : 22.12.2011.