

Министерство образования Республики Беларусь
Учреждение образования
Белорусский государственный университет
информатики и радиоэлектроники

УДК _____

Кустов
Антон Юрьевич

Распознавание рукописных цифр на основе нейронных сетей

АВТОРЕФЕРАТ

на соискание степени магистра технических наук
по специальности 1-40 80 01 Элементы и устройства вычислительной
техники и систем управления

Научный руководитель
Вашкевич Максим Иосифович
Доц., канд. техн. наук,

Минск 2015

ВВЕДЕНИЕ

Задача по распознаванию рукописных цифр широко распространена и является активной темой в приложениях ОРС и в исследованиях по классификации/обучению образов. В ОРС приложениях распознавание цифр применяется для сортировки почтовых отправлений, обработки банковских чеков, ввода данных с форм и т.д. Для этих приложений производительность (точность и скорость) распознавания цифр имеет решающее значение для эффективной работы. Среди разработчиков алгоритмов по классификации образов и машинного обучения проблема распознавания рукописных цифр является хорошим примером для тестирования производительности классификации.

Производительность распознавания символов в значительной степени зависит от подхода выделения признаков и схемы классификации/обучения. Для получения хорошего результата требуется построение нейронной сети со сложной архитектурой. У однослойных сетей плохая способность обобщения. Многослойные сети показывают хороший результат, когда они организованы в структуру, невосприимчивую к смещению цифры на изображении. В данной работе показано, что уменьшение свободных параметров в сети улучшает обобщающую способность.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Генерализация – это главное свойство, которое должно быть принято во внимание. Оно определяет количество данных, необходимых для обучения системы так, чтобы она показывала хорошую точность при распознавании наборов данных, не участвующих в обучающей выборке. В работе исследуется вопрос максимизации значения этого параметра.

В литературе описано много успешных приложений, использующих для обучения метод обратного распространения ошибки. В связи с этим, в этой работе использован этот метод. Некоторые проблемы могут быть решены с использованием неструктурированных сетей, но нельзя надеяться, что неструктурированная сеть покажет хорошие обобщающие способности в любой задаче. Главная цель работы – показать, что хорошая обобщающая способность может быть достигнута в том случае, если в сети будут присутствовать основные знания о задаче.

При разработке сети предпринята попытка уменьшить количество используемых функций, которые сеть может генерировать. Теоретические исследования показали, что хорошая генерализация зависит от размера гипотетического пространства, количества обучающих примеров. Если гипотетическое пространство слишком велико или количество обучающих наборов слишком мало, тогда будет большое количество сетей, которые будут одинаково работать с обучающими данными, где только маленькое их количество будет являться решением общей проблемы, из чего следует низкая обобщающая способность.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

В диссертации рассмотрены проблемы, связанные с распознаванием символов. Рассмотрена краткая история ОРС (оптического распознавания символов), значимый вклад в развитие данной темы, стимулировавший развитие роста исследований распознавания образов, а также обзор методик по проблеме диссертации. Описаны признаки и наиболее распространенные методы извлечения признаков, которые могут быть классифицированы как геометрические признаки, конструктивные признаки и методы преобразования пространства признаков.

В 3-ем разделе представлены методы классификации образов, которые, в том числе, использовались в данной работе. В этом разделе описаны методы классификации на основе нейронных сетей, которые позволяют получить хорошие результаты при распознавании символов. Среди этих методов следующие: однослойная сеть; многослойный персептрон; сверточная нейронная сеть.

В 4-м разделе описаны методы обучения и оптимизации нейронной сети. В этом разделе описан быстрый метод для расчета градиентов, известный как «Метод обратного распространения ошибки». Целью обучения сети алгоритмом обратного распространения ошибки является такая подстройка ее весов, чтобы приложение некоторого множества входов приводило к требуемому множеству выходов. В этом же разделе описан метод ускорения обучения, известный как «классический метод инерции». Метод инерции или классический метод инерции – это метод ускорения градиентного спуска, который сохраняет вектор скорости в направлениях уменьшения требуемой величины в течение повторений.

В 5-м разделе представлены в работе и результаты экспериментов, которые показали, что наилучшая обобщающая способность была достигнута сверточной сетью с несколькими детекторами. В этом разделе представлены графики обучения сетей, матрицы спутывания. Проведено сравнение нескольких архитектур, отличающихся количеством и типом связей.

Архитектура сети	Количество весов	Точность, %
1-слойная сеть	7850	87,23
2-слойная сеть	9550	93,26
2-слойная сверточная	6277	90,72
2-слойная сверточная с 2-мя детекторами	12554	94,83

ЗАКЛЮЧЕНИЕ

В диссертационной работе представлены примеры работы сетей, по которым видно, что более сложные нейронные сети показывают лучшие результаты. Однако разработка и программирование таких сетей требует немалых затрат времени и знаний. Тем не менее, в данной работе показано, что задачу можно решить, по крайней мере для распознавания изображений, используя локальные детекторы инвариантные к сдвигу. Подобные методы могут быть применены для других целей, например, распознавания речи.

Для использования методик, описанных в данной работе, могут понадобиться комплексные программные средства с дополнительным пользовательским интерфейсом для решения реальных задач. Для разработок сетей при использовании в реальных задачах могут понадобиться значительные улучшения алгоритмов обучения, рассмотренных в данной работе.

Библиотека БГУИР

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

Библиотека БГУИР