

Применение нейронных сетей для распознавания символов

[Eric W. Brown](mailto:feneric@ccs.neu.edu)
feneric@ccs.neu.edu

Аннотация

В статье исследуются применения нейронных сетей для решения проблемы распознавания печатных символов в автоматическом режиме. В частности, здесь применяется сеть обратного распространения, которая обучена на восьмидесяти четырех символах одного шрифта и протестирована на двух других шрифтах. Его результаты сравниваются с результатами, полученными на тех же данных с применением более традиционного подхода.

Ключевые слова

Распознавание текста, оптическое распознавание символов (OCR), нейронные сети, обратного распространения, выделение признаков, офлайн распознавания символов, машины печатные распознавания символов.

Потребность в той или иной форме автоматизированного или полуавтоматического оптического распознавания символов (OCR) была признана в течение десятилетий. Сегодня существует множество алгоритмов, которые выполняют эту задачу, каждый со своими преимуществами и недостатками. В этой статье исследуются различия между двумя разными алгоритмами распознавания: метод выделения признаков, используя традиционные методы искусственного интеллекта для классификации и нейронных сетей практически без предварительной обработки. Основная идея эксперимента была по существу распознать идентичные данные с помощью двух различных алгоритмов и обратить внимание на различия в каждом прогоне. Метод выделения признаков фактически выполнен автором ранее, в то время как метод нейронных сетей был реализован специально для этой статьи с использованием точно таких же данных и после аналогичной процедуры. Таким образом, сравнение двух методов не зависит от различий в данных или несбалансированных процедур, которые бы дали преимущество одного метода над другим.

Используемые данные включают восемьдесят четыре символа шрифта. Разрешение символа для всех трех шрифтов 8X8. Наборы содержат заглавные и строчные буквы, а также номера и несколько разных знаков препинания. Наборы символов используются как фактические символы дисплея на старом IBM S/360 computer поэтому они представляют реальные данные, они не были предназначены для этого эксперимента. В частности, наборы включают в себя несколько примеров допустимых альтернативных видов символов (т.е. нуля с косой чертой вместо нуля без косой черты и т.д.). Таким образом, общий результат несколько ниже, чем он должны быть, справедливое обращение было бы устранить эти специальные формы символов.

Целью этой статьи не является разрыв первоначального алгоритма выделения признаков в мельчайших подробностях. Общее резюме проектирования и эксплуатации включено здесь для удобства. Идея алгоритма извлечения точечного объекта является определением символа на основе функций, которые чем-то напоминают способности людей используемые для идентификации символов. Смысл заключается в том, что, когда алгоритм делает классификацию символа, он должен выбрать символ. Для реализации этого алгоритм сканирует по всей 8X8 матрице и анализирует каждый непустой пиксель. В непосредственной близости от пикселей будут рассмотрены все пиксели (см. рисунок 1, "Отображение символа"). Для сравнения были вычислены суммы минимального расстояния между точками символа, которые будут определены как функции. Гипотезой является словарь символов с наименьшими суммами.

Обратите внимание, что этот алгоритм не пытается рассмотреть различные типы точек, хотя это может быть, это не так важно. Алгоритм не был оптимизирован для наилучшего результата, это было лишь попыткой проверить общее понятие алгоритма. Поскольку нейронная сеть также не была оптимизирована, общее сравнение между этими двумя методами распознавания можно считать достоверным. Результаты алгоритма выделения признаков приведены в Таблице 2, "Результат распознавания – метод выделения признаков".

Рисунок 1: Отображение символа

```
(D)isplay char, (R)un new set, or (Q)uit: d

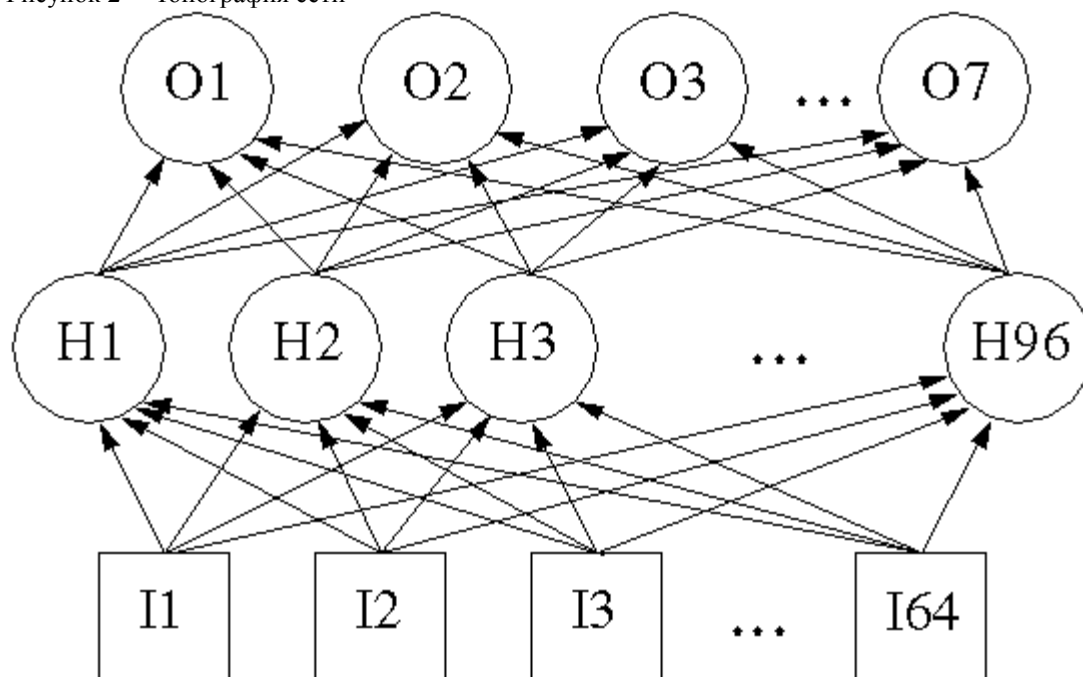
Character number: 68
Character name: K

XX *X
** **
****
**X
****
** **
XX *X
```

Подход с нейронной сетью реализован в три отдельных этапа. На первом этапе просто перевод двоичных данных символов в дружественную форму. На втором этапе приняли на выходе первой обученной сети обратного распространения, выводя все полученные веса и общую информацию о сети. На третьем этапе создали новую сеть. Затем провели полный набор символов по сети и вывели результат классификации. Причины для реализации нейронной сети как три этапа были. Выполняя первый шаг, с предварительной обработки кода с функцией извлечения OCR программа может быть использована, чтобы устранить разницу между двумя алгоритмами. На втором этапе машина должна быть отделена только потому, что обучение было таким медленным процессом. Несколько машин могут, таким образом, выполнять только обучение в то время как другая машина будет использоваться для анализа результатов. Сеть состояла из шестидесяти четырех входов, девяносто шести скрытых узлов, и семи выходов. Это была по существу упреждающая сеть, которая была полностью подключена без самостоятельного входа или отклонения (см. Рисунок 2, "Топография сети"). Это было сделано для подготовки того же набора символов, что алгоритм извлечения признаков (см. Рисунок 3, "СВМ символы и извлеченные функция точек"). Обучение было достигнуто за счет обратного распространения без движения.

Каждый из шестидесяти четырех входов был подключен к одному из пикселей в 8X8. Вход был равным нулю, если пиксель был пуст, или один в противном случае. Семь выходов были использованы, чтобы сделать семь ярлыков, что совпало с упорядочением набора символов.

Рисунок 2 - "Топография сети"



Обучение было чревато трудностями. Небольшие изменения в размер шага, используемые в обновлении весов привело к большим изменениям в поведении обучения в целом. В частности, большинство Размер шага в результате полного провала сходятся в течение разумного числа итераций по набору данных (учебная программа была создана для рассмотрения номер 100000 разумно). Там, казалось, не было априори определить хороший размер шага (или число скрытых узлов использовать) образуют с самого начала. Программы обучения требуется несколько сотен компьютеров часов на специальном Sun SPARCstation бежать, что делает его практически невозможно, чтобы действительно попытаться получить оптимальный результат. Это почти гарантированно, что девяносто шестерых, не оптимальное число скрытых узлов, это число было произведено путем умножения количества входов на 1,5. Вполне возможно использование momentum3 или method5 Ньютона может частично облегчить проблемы с обучением. В настоящее время эта проблема стоит обучение, как нежелательно, что метод выделения признаков не имеет, это крайне затрудняет для оптимизации нейронной сети в классификатор никаких ограничений по времени.

При всех этих трудностей было сказано, сеть было сделано правильно указать шаг 1,05. Она довольно быстро сходится, и требуется только сто тридцать шесть проходов через набор данных. После соответствующей подготовки сеть была получена, она производила неплохие результаты для неоптимизированного алгоритма. Она также подготовила их явно быстрее, чем метод выделения признаков. Эта разница частично объясняется тем, что в то время как метод выделения признаков должен делать посимвольное сравнения для каждого символа, которые будут определены во время выполнения, нейронная сеть узнала все, что нужно знать о символе во время ее подготовки может получить результаты для символа с одним упреждающим проходом.

Таблица 1: Результаты распознавания – метод нейронной сети

Результаты для 'Alt' набора символов		
Всего правильно	66	79%
Всего правильно (без учета одинаковых символов)	15	18%
Всего неизвестных	2	2%
Всего неправильных догадок	16	19%
Всего неправильно или не указано	18	21%
Результаты для 'Ult' набора символов		
Всего правильно	33	39%
Всего правильно (без учета одинаковых символов)	12	14%
Всего неизвестных	3	4%
Всего неправильных догадок	48	57%
Всего неправильно или не указано	51	61%
Всего		
Всего правильно	99	54%
Всего правильно (без учета одинаковых символов)	27	16%
Всего неизвестных	5	3%
Всего неправильных догадок	64	38%
Всего неправильно или не указано	69	41%

Точность этого алгоритма ниже чем у метода выделения признаков, однако, его скорость намного выше, что делает его приоритетней.

Таблица 2: Результаты распознавания – метод выделения признаков

Результаты для 'Alt' набора символов		
Всего правильно	72	86%
Всего правильно (без учета одинаковых символов)	21	25%
Всего неизвестных	5	6%
Всего неправильных догадок	7	8%

Результаты для 'Alt' набора символов		
Всего неправильно или не указано	12	14%
Результаты для 'Ult' набора символов		
Всего правильно	49	58%
Всего правильно (без учета одинаковых символов)	28	33%
Всего неизвестных	12	14%
Всего неправильных догадок	23	27%
Всего неправильно или не указано	35	42%
Всего		
Всего правильно	121	72%
Всего правильно (без учета одинаковых символов)	49	29%
Всего неизвестных	17	10%
Всего неправильных догадок	30	18%
Всего неправильно или не указано	47	28%

Возможно, лучшим вариантом будет сочетание обоих подходов. Хотя в настоящее время, если сеть была обучена на полных данных символов, было бы относительно легко обучить ее только на извлечение данных функций. Изменение было сделано, но после нескольких тысяч минут машинного времени на множестве различных машин вывода не было сделано. Ожидается, что такой комплексный подход позволит сохранить желаемый метод извлечения функций как обычно делает человек, сохраняя при этом скорость идентификации нейронными сетями. Тщательное корректировки сети и процесса выделения признаков может радикально повысить точность.

Общие результаты экспериментов были многообещающими, но далеки от идеальных в их нынешнем виде. Если один из методов должен быть выбран по сравнению с другими, результат здесь будет означать, что метод выделения признаков с использованием стандартных методов ИИ для классификации будет лучшим выбором. Гибридный алгоритм дает другие варианты, и, вполне возможно, наибольший потенциал лежит по этому пути.

Литература

1. E. W. Brown, "Character Recognition by Feature Point Extraction", неопубликованная статья написанная в Northeastern University, 1992
2. The Commodore 128 Computer, Commodore Business Machines, 1984
3. Stephen I. Gallant, Chapters 11 & 12; *Neural Network Learning and Expert Systems*, 1993, The MIT Press, Cambridge, MA
4. D. S. Levine, pp. 58-62; *Introduction to Neural & Cognitive Modeling*, 1991, Lawrence Erlbaum Associates, Publishers; Hillsdale, NJ
5. Patrick K. Simpson, pp. 112-123, *Artificial Neural Systems*, 1990, Pergamon Press, New York, NY

Рисунок 3 - "СВМ символы и извлеченные функция точек"

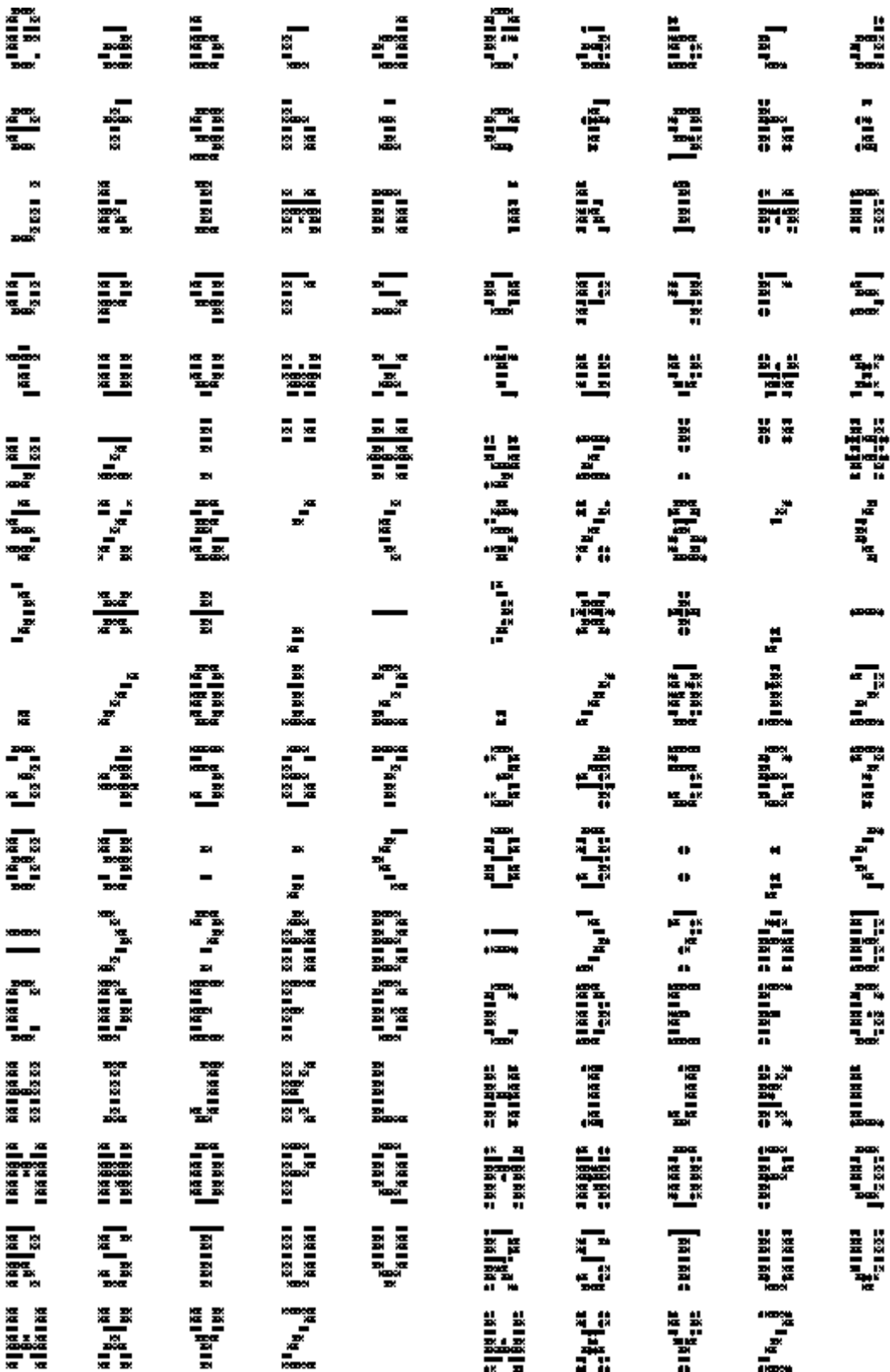


Рисунок 4 – “Alt” набор символов

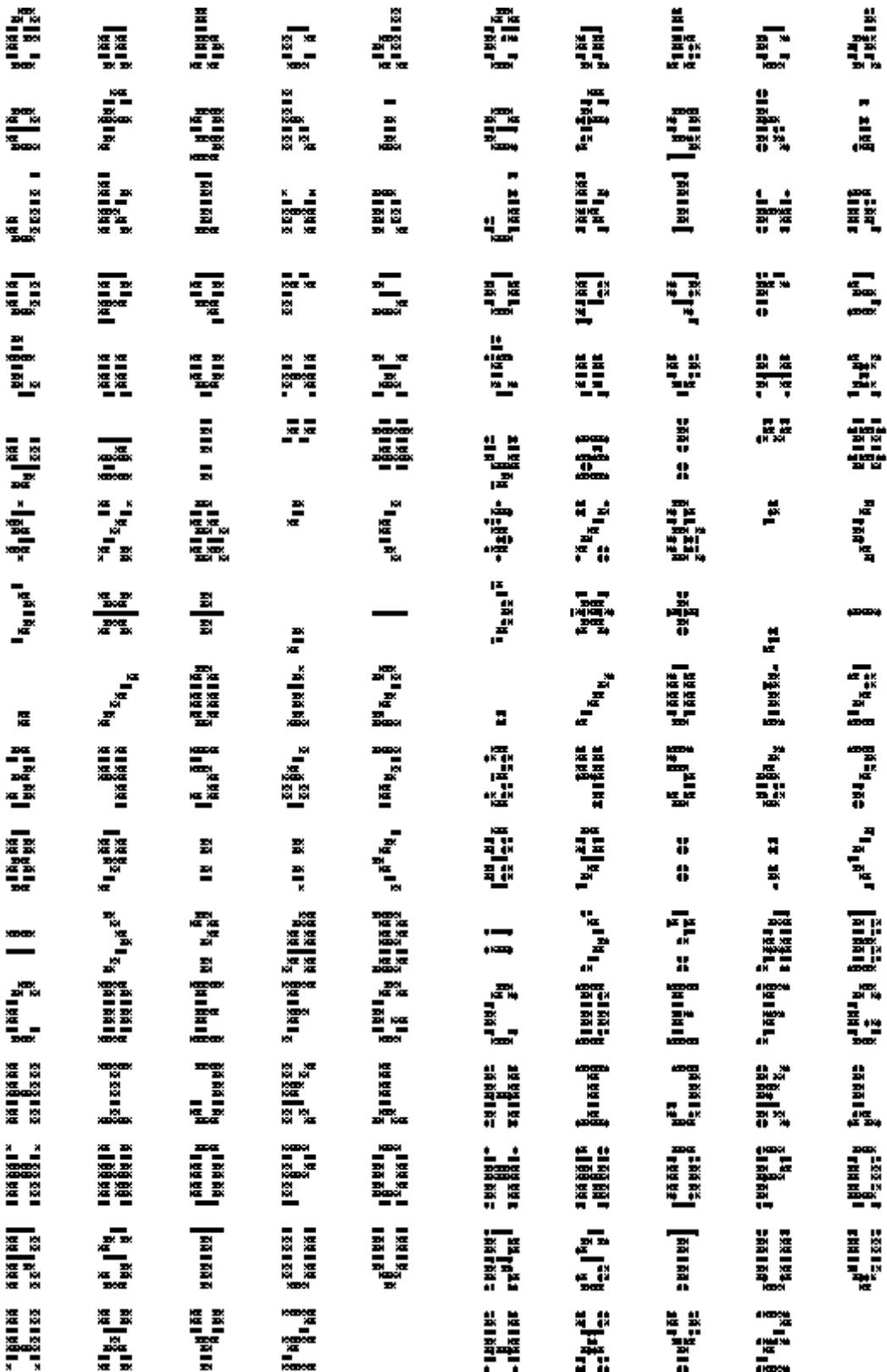


Рисунок 5 – “Ult” набор символов

