

5. Курдюмов В.П., Хромов А.П. О базисах Рисса из собственных и присоединенных функций дифференциально-разностного оператора с многоточечным краевым условием // Математика. Механика: Сб. науч. тр. Саратов: Изд-во Сарат. ун-та, 2004. Вып. 6. С. 80–82.

УДК 681.322

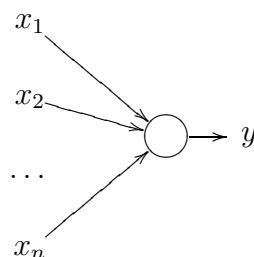
А.В. Малыгин, О.И. Аравин

## РАСПОЗНАВАНИЕ ОБРАЗОВ НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ

В настоящий момент в мире бурно развивается новая прикладная область математики, специализирующаяся на искусственных нейронных сетях. Актуальность исследований в этом направлении обусловлена возможностями применения нейронных сетей в самых разных областях.

Широта задач, решаемых нейронными сетями (НС), объясняется во многом тем, что возможность обучения позволяет сделать функционирование системы на порядок более эффективным.

Нейронные сети могут сильно отличаться друг от друга, однако у них есть несколько общих черт. Основу каждой нейронной сети составляют простые, чаще всего однотипные, элементы (ячейки), имитирующие работу нейронов мозга (аксоны и дендриты). Нейрон может определяться своим текущим состоянием по аналогии с нервными клетками головного мозга, которые могут быть возбуждены или заторможены. Нейрон обладает группой синапсов – входных связей, соединенных с выходами других нейронов, а также имеет аксон – выходную связь данного нейрона, с которой сигнал (возбуждения или торможения) поступает на синапсы следующих нейронов [1]. (рисунок)



Нейрон

Каждый синапс определяется так называемой величиной синаптической связи или ее весом  $w_i$ . Текущее состояние нейрона определяется как сумма его входов:

$$s = \sum_{i=1}^n x_i w_i.$$

Выход нейрона можно определить как функцию от его состояния:  $y = f(s)$ .

Нелинейная функция  $f$  называется *активационной* и может быть представлена по-разному, одной из наиболее распространенных является нелинейная функция с насыщением, или сигмоид (т.е. функция  $S$ -образного вида).

Этап проектирования, в котором происходит корректировка весов, называется *обучением нейронной сети*, и от того, насколько качественно он будет выполнен, зависит возможность сети решать поставленные проблемы. На этом этапе, кроме параметра качества подбора весов, важную роль играет и время обучения. Чаще всего, эти два параметра связаны обратной зависимостью. Нейронную сеть можно обучать с учителем или без него [2]. В первом случае сети предъявляются значения входных и правильных выходных сигналов, и сеть по внутреннему алгоритму подстраивает веса своих синаптических связей. Во втором случае, выходы НС формируются самостоятельно, а веса корректируются по алгоритму, учитывающему только входные сигналы.

Например, большинство известных на сегодняшний день нейронных сетей используют для нахождения взвешенной суммы входов нейрона формулу

$$s = \sum_{i=1}^n x_i w_i,$$

однако в некоторых приложениях нейронных сетей полезно ввести другую запись, например:

$$s = \sum_{i=1}^n x_i^2 w_i$$

или

$$s = \sum_{i=1}^n x_i x_{((i+1) \bmod n)} w_i.$$

Главное, чтобы разработчик нейронной сети четко понимал, для чего эта сеть проектируется и какими полезными свойствами должен обладать нейрон. Введение нелинейности такого рода в общем случае увеличивает вычислительную мощь сети, то есть позволяет из меньшего числа нейронов создать нейронную сеть, выполняющую работу обычной сети.

Задача, стоящая перед нами, включала в себя проектирование нейронной сети, позволяющей идентифицировать образцы почерков разных людей и формирование этих образцов на основе рукописных материалов.

На первом этапе была спроектирована нейронная сеть для распознавания образцов почерка двух людей (А и В). Обучение производилось на основе выборки из образцов почерка людей. Под образцом почерка мы понимаем отдельные растровые изображения букв каждого из людей. Сеть имела 3

слоя, в каждом из которых происходила обработка разных сегментов изображения, каждый участок сети обрабатывал свой кусок изображения и подавал на следующий слой результативный выходной сигнал, на следующем слое происходила обработка этих сигналов, группировка более крупных кусков и так далее, пока на третьем слое не вычислялся выходной сигнал от всего изображения в целом. Сначала обучение сети производилось с помощью лишь одной буквы «И». Впоследствии выборка была расширена до пяти букв «И», «Ж», «О», «А», «Н». В дальнейшем представляется необходимым расширять набор рассматриваемых букв, чтобы при анализе текста можно было бы существенно улучшить качество распознавания. Наша обучающая выборка состояла из 5 букв «И», 5 букв «О», 5 букв «Ж», 10 букв «А» и 5 букв «Н» каждого из людей.

После обучения следовала длительная процедура проверки корректности нейронной сети с помощью большого количества образцов почерков, не участвующих в обучении. На вход нейронной сети подается битовая матрица изображения образца почерка.

На выходе нейронная сеть выдает 1, если указанный образец можно отнести как наиболее близкий к образцам почерка человека А, и 0 — в обратном случае. Путем корректировки активационной функции и пороговых значений, а также скорости обучения, удалось добиться достаточно высокой точности распознавания (менее 5% ошибок).

На следующем этапе необходимо было продолжить работу в направлении сильного снижения процента ошибок и более общей классификации. То есть в случае, когда рассматривалась бы группа людей и сеть была бы обучена образцам почерков каждого из них. Работоспособность нейронной сети проверялась на 11 тестовых примерах. Тестирование показало, что уровень погрешности существенно снижен (менее 1% ошибок). Наибольший процент ошибок был зарегистрирован при распознавании буквы «О», что объясняется достаточно общей манерой ее написания.

После получения работоспособной нейронной сети встала следующая проблема. При криминалистическом анализе почерка необходимо иметь надежный механизм работы с рукописным текстом, который позволил бы выделить из текста отдельные буквы, распознать их и подать на вход нейронной сети [3]. Таким образом, имея в своем распоряжении текст-образец почерка подозреваемого и текст, который необходимо идентифицировать, программа могла бы обучить нейронную сеть на основе букв, выделенных из текста-образца, и сделать вывод о принадлежности его подозреваемому.

Задача, стоящая перед нами, оказалась далеко не тривиальной, потому что необходимо было решить несколько проблем:

- очистить текст от артефактов и корректировка естественных деформаций;

- выделить отдельные буквы из текста и сохранить их в отдельных файлах или битовых массивах, пригодных к тому, чтобы быть поданными на вход нейронной сети;
- распознать эти буквы.

На первом этапе работы с текстом необходимо разбить страницу на блоки, основываясь на особенностях правого и левого выравнивания и возможного наличия нескольких колонок. Затем распознанный блок разбивается на строки. И здесь неизбежны перекосы изображения страницы или фрагментов страницы. Даже небольшой наклон может привести к тому, что левый край одной строки становится существенно ниже правого края следующей, что особенно чувствительно при небольшом межстрочном интервале [4]. В результате возникает проблема определения строки, к которой относится тот или иной фрагмент изображения. Например, для букв «Ё», «Й» при небольшом наклоне уже сложно определить, к какой строке относится верхняя (отдельная) часть символа (иногда ее можно принять за запятую или точку).

Затем происходит разбиение строк на непрерывные области изображения, которые, как правило, соответствуют буквам; алгоритм распознавания делает предположения относительно соответствия этих областей символам; а затем делается выбор каждого символа.

Встречающиеся в реальной жизни тексты обычно далеки от совершенства, и процент ошибок распознавания для «нечистых» текстов часто недопустимо велик. Грязные изображения имеют наиболее очевидную проблему, потому что даже небольшие пятна могут затенять определяющие части символа или преобразовывать один в другой.

Распознавание осложняют особенности почерка отдельных людей. Для улучшения распознавания некачественных текстов часто используют специальные контекстные модули. В систему распознавания имеет смысл заложить словарь языка, модели словоизменения. Например, в слове «дом» буква «о», написанная с дефектом, превратилась в букву «с». Но программа знает, что слова «дсм» в русском языке нет, а слово «дом» есть, значит, программа решает, «скорее всего, это буква "о»».

В настоящий момент нам удалось решить задачу обработки изображения, удаления артефактов и выделения отдельных букв, проблема распознавания решена пока только на уровне распознавания печатного текста.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Ясницкий Л.Н.* Введение в искусственный интеллект. М.: Академия, 2005.
2. *Хайкин С.* Нейронные сети. М.; СПб.; Киев: Вильямс, 2006.
3. *Курьянова А.А.* Теоретические основы и методика судебно-почерковедческих исследований: Автореф. дис. ... канд. техн. наук. М., 1992.
4. *Рутковская Д., Пилиньский М., Рутковский Л.* Нейронные сети, генетические алгоритмы и нечеткие системы. М.: Горячая Линия – Телеком, 2007.