

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ РАСПОЗНАВАНИЯ И ОБУЧЕНИЯ НЕОКОГНИТРОНА

Сова А.А., Федяев О.И.

Донецкий национальный технический университет
кафедра прикладной математики и информатики

E-mail: ann-sov@yandex.ru

Аннотация

Сова А.А., Федяев О.И. Математическая модель распознавания и обучения неокогнитрона. Рассмотрена архитектура многослойной нейронной сети типа неокогнитрон, описана математическая модель процессов обучения и распознавания. Разработаны алгоритмы для программной реализации неокогнитрона, ориентированного на распознавание различных рукописных символов.

Проблемы распознавания образов, подверженных искажениям

Большинство методов распознавания образов являются чувствительными к сдвигам в позиции и искажениям формы образов распознавания. Поэтому ведутся поиски новых подходов, которые преодолевают эти проблемы. Одним из перспективных способов распознавания искаженных образов считается применение специальных нейросетевых структур типа неокогнитрон. Это обусловлено особой структурой данного типа нейросетей, в определённой степени имитирующей работу зрительной системы человека [1]. В работе закладываются математические основы для построения эффективных алгоритмов и программных средств распознавания графических образов с помощью нейросети типа неокогнитрон. Потребность в таких системах постоянно возрастает, поэтому создание на неокогнитроне системы распознавания, которая не уступает в функциональном и экономическом отношении существующим аналогам, является актуальной научно-технической задачей.

Основной задачей данной работы является разработка математической модели неокогнитрона, которая позволит создать программный эмулятор неокогнитрона с возможностью его обучения на распознавание определенного класса графических образов. В перспективе на основе математического описания можно будет точно сформулировать и решить задачу оптимизации многочисленных параметров модели неокогнитрона, адекватных природе распознаваемых графических образов.

Структура нейросети типа неокогнитрон

Неокогнитрон имеет иерархическую структуру. Он состоит из последовательности обрабатывающих слоев (S-слоя и C-слоя). На рис. 1 приведена структура неокогнитрона, состоящего из трех модулей, образованных парами S и C слоев. Входной образ подается на первый слой и передается через плоскости, соответствующие последующим слоям, до тех пор, пока не достигнет выходного слоя, в котором идентифицируется распознаваемый образ [2, 3].

Входной, сенсорный, слой неокогнитрона является прямоугольным полем, которое состоит из светочувствительных клеток. Каждый последующий слой состоит из групп нейронов, причем нейроны одной группы имеют одинаковые веса и распознают одни и те же части образа. Эти нейроны образуют так называемые S-ячейки, расположенные в соответствующих S-слоях. S-ячейки служат для выделения свойств образов, поступающих с предыдущего слоя [1].

Выбор числа слоев неокогнитрона зависит от сложности решаемой задачи, т.к. сеть с большим числом слоев способна распознавать более сложные изображения. Такое

чередование слоев, при котором в слое S выделяются основные признаки изображения, а в слое C происходит коррекция искажений изображений, оказалось наиболее эффективным при синтезе систем распознавания образов, инвариантных к деформациям изображений [4].

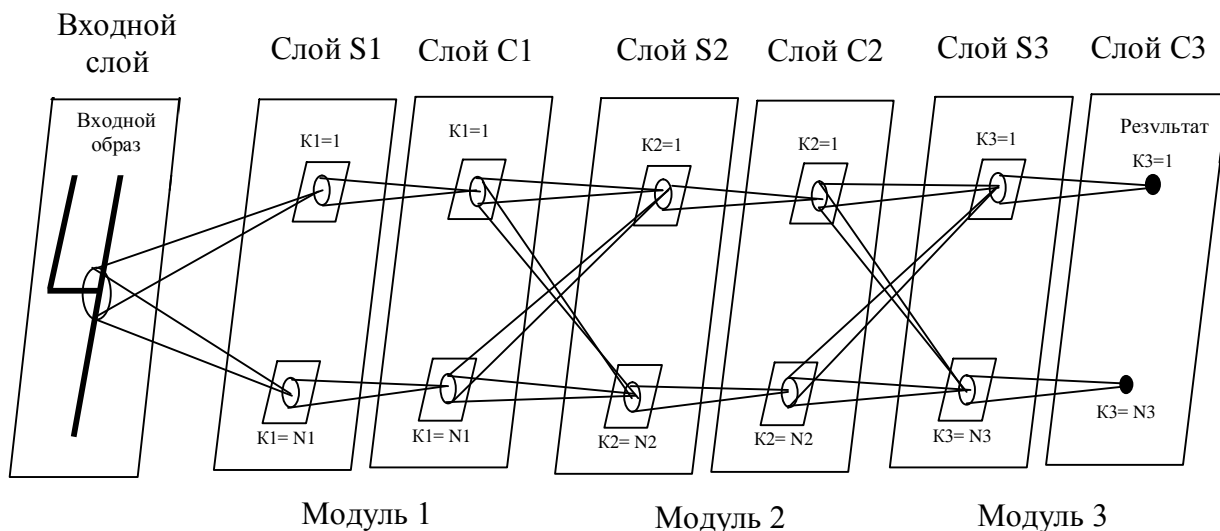


Рис. 1 – Иерархическая структура неокогнитрона со связями между слоями

Выходы нейронов S-слоя соединены с нейронами C-слоя. При этом все ячейки в отдельной плоскости, имеют входные взаимосвязи того же самого пространственного распределения и только позиции предшествующих клеток, от которых их входные взаимосвязи прибывают, сдвинуты параллельно. Эта ситуация проиллюстрирована на рис. 2. Даже в процессе обучения, в котором значения входных взаимосвязей S-клеток различны, переменные взаимосвязи всегда изменяются согласно этому ограничению [3].

Связи, соединяющие S-ячейки и C-ячейки одного модуля, являются жёсткими и не изменяются в процессе обучения. Связи, соединяющие нейроны S-слоя с нейронами предыдущего C-слоя, являются изменяющимися и могут изменяться в процессе обучения. S-ячейка активизируется только тогда, когда найден определенный образ в соответствующем этой клетке месте предыдущего слоя. Образ, на который реагирует ячейка, определяется в процессе обучения [1, 2].

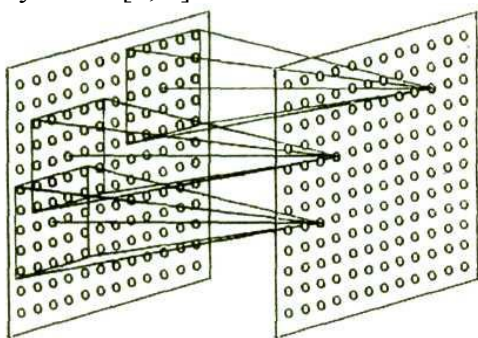


Рис. 2 – Входные связи между нейронами в соседних плоскостях

Ячейки неокогнитрона

Все ячейки, используемые в неокогнитроне, имеют аналоговый тип: т.е. сигналы ввода и вывода ячеек принимают неотрицательные непрерывные значения. Каждая ячейка имеет характеристики, аналогичные биологическому нейрону.

В неокогнитроне используются четыре различных вида ячеек, то есть, S-ячейки, C-ячейки, Vs-ячейки и Vc-ячейки. Для примера сначала обсудим характеристики S-ячейки.

Как показано на рис. 3, S-ячейка имеет много входных терминалов, возбуждающих и подавляющих. Если ячейка получает сигналы из возбуждающих входных терминалов, выход ячейки увеличится. С другой стороны, сигнал из подавляющего входного терминала подавит выходной сигнал. Каждый входной терминал имеет свой собственный коэффициент соединения, значение которого положительно. Хотя ячейка имеет только один терминал вывода, она может посылать сигналы множеству входных терминалов других ячеек. S-ячейка имеет подавляющий вход, который производит эффект шунтирования. Пусть $u_1, u_2 \dots u_N$ - возбуждающие входы, а v - подавляющий вход. Для расчёта выходного сигнала подавляющего нейрона используется следующая формула:

$$v = \sqrt{\sum_i (b_i u_i)^2} \quad 1)$$

где: v – выход тормозящего узла;

i – область над всеми комплексными узлами, с которыми связан тормозящий узел;

b_i – вес i -й связи от возбуждающего узла к тормозящему узлу;

u_i – выход i -го возбуждающего узла.

Веса b_i выбираются монотонно уменьшающимися с увеличением расстояния от центра области реакции, при этом сумма их значений должна быть равна единице.

Тогда выход w этой S-ячейки определен так:

$$w = \varphi \left[\frac{1 + \sum_{i=1}^N a_i * u_i}{1 + b * v} - 1 \right] \quad 2)$$

где: a_i и b представляют возбуждающие и подавляющие коэффициенты соединения, соответственно. Коэффициент b рассчитывается по формуле (1) [3, 5, 6].

Функция $\varphi[]$ определяется следующим уравнением:

$$\varphi[x] = \begin{cases} x, & (x \geq 0); \\ 0, & (x < 0). \end{cases} \quad 3)$$

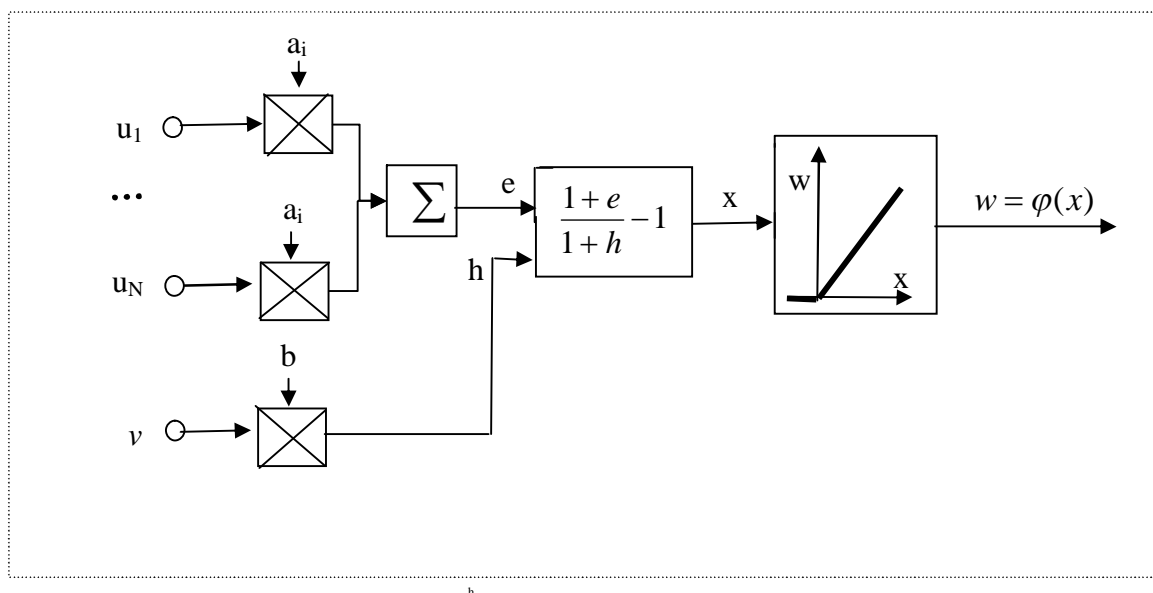


Рис. 3 – Выходная характеристика нейронов S-слоя

Пусть e – сумма всех эксцитатных вводов перемноженных с коэффициентами соединения и h – запрещающий ввод, умноженный на коэффициент соединения; т.е.:

$$e = \sum_{i=1}^N a_i * u_i, \quad (4)$$

$$h = b * v \quad (5)$$

Таким образом, подставляя в уравнение (2) формулы (4) и (5), получим:

$$w = \varphi \left[\frac{1+e}{1+h} - 1 \right] = \varphi \left[\frac{e-h}{1+h} \right] \quad (6)$$

Входной сигнал для нейронов С-слоя вычисляется по формуле:

$$y = \sum_i s_i u_i \quad (7)$$

где s_i – выходной сигнал простых нейронов;

u_i – весовые коэффициенты от простых нейронов к комплексному [4, 7].

Для нейронов комплексных слоёв использовались различные функции активации.

Алгоритм обучения неокогнитрона

Перед началом обучения весам S-слоя присваивают случайным образом определенные положительные значения и задаются неизменные в дальнейшем веса С-слоя. Настраиваемыми коэффициентами выступают коэффициенты от нейронов комплексного слоя к нейронам простого слоя и коэффициент от подавляющего нейрона к нейрону простого слоя. Коэффициенты от комплексных нейронов к подавляющему нейрону и от простых нейронов к комплексным являются фиксированными. В связи с этим возникает задача определения значений фиксированных коэффициентов. В самом начале предполагается, что значения коэффициентов должны уменьшаться с удалением от центра области видимости.

При настройке весовых коэффициентов использовалась простая формула для расчета коррекции (Δ), которая зависит от величины входных сигналов [1, 2, 4, 7]:

$$\Delta w_{i,j} = \alpha w_{i,j} * y_i; \quad (8)$$

$$\Delta w_{v,j} = \alpha v = \alpha \sqrt{\sum_i x_i^2 w_{i,v}}. \quad (9)$$

где α – параметр обучения.

Величина Δ для весового коэффициента подавляющего нейрона выбирается так, чтобы процесс обучения мог завершиться. По поводу формулы, которую лучше использовать, нет единого мнения. Поэтому исследовались следующие возможные два способа.

Первый способ основывался на формуле $\Delta w_0 = \alpha c_{i,j}$. В этом случае настраиваемый коэффициент зависит только от центрально элемента предыдущего слоя. Это накладывает некоторые ограничения на расположение обучающего образа, то есть образ должен быть расположен так, чтобы нейроном предыдущего слоя с максимальным возбуждением был нейрон, который находится в центре области видимости [1, 2].

Второй способ основывался на формуле $\Delta w_0 = \alpha \frac{v}{2}$ [1, 4]. В этом случае дельта зависит от выхода тормозящего нейрона, а сам выход тормозящего нейрона зависит от всех нейронов предыдущего слоя, с которыми связан обучаемый нейрон.

Обучение неокогнитрона происходит слой за слоем. При обучении нейронов использовался метод, при котором в каждой матрице обучается только один нейрон (обычно используется центральный нейрон). После завершения обучения весовые коэффициенты

этого нейрона переносятся на все остальные нейроны в этой матрице. Таким образом, процесс обучения производится достаточно быстро, так как на каждом слое необходимо обучить лишь столько нейронов, сколько матриц в данном слое, а их количество обычно не превосходит 100 [2, 3, 4].

Связи в неокогнитроне должны быть установлены так, чтобы нейроны последнего слоя “видели” полностью входной образ, иначе теряется смысл распознавания. Это условие является обязательным, но, как показывают исследования, размер рецептивного поля не обязательно должен стать равным размеру входного изображения лишь на последнем слое, это может происходить и раньше [4].

Процесс обучения заканчивается тогда, когда будут обучены все S-слои неокогнитрона распознавать эталонные образы.

Выводы

В данной работе была представлена и проанализирована математическая модель неокогнитрона, которая описывает процессы обучения и распознавания. Благодаря своей структуре неокогнитрон способен распознавать образы с различными видами помех и искажений. Несмотря на свою сложность, он является достаточно эффективным и быстродействующим распознавателем.

Анализируя модель неокогнитрона, можно отметить следующие его преимущества над другими нейросетями, которые также ориентированы на распознавание образов:

- малая чувствительность к основным искажениям, которым подвергаются графические изображения (масштабирование, перенос, хаотические шумы);
- способность решать задачу распознавания графических образов за короткое время;
- нейросети типа неокогнитрон самоорганизующиеся;
- наличие в неокогнитроне достаточного количества параметров для его настройки на различные типы образов;
- неокогнитрон сам выделяет отличительные особенности во входном наборе образов, которые в дальнейшем используются при идентификации каждого образа независимо от деформации.

Эти достоинства послужили основанием для разработки программной модели рассмотренной архитектуры неокогнитрона, чтобы с её помощью авторам статьи продолжить исследование процессов обучения и распознавания не только графических образов, но и лиц человека.

Литература

1. Федяев О.И., Махно Ю.С. Система распознавания зашумленных и искаженных графических образов на основе нейронной сети типа неокогнитрон. // Одиннадцатая национальная конференция по искусственному интеллекту с международным участием КНИ-2008: труды конференции, Т.3.-М.: ЛЕНАНД, 2008. – С. 75-83.
2. Уоссермен Ф. Нейрокомпьютерная техника. Теория и практика. – М.: Мир, 1992. – С. 274.
3. K. Fukushima: Neocognitron capable of incremental learning. Tokyo University of Technology, Tokyo, Japan – 2003. – P. 37-43.
4. Freeman J., Skapura D. Neural Networks. Algorithms, Applications and Programming Techniques. // Addison-Wesley. – 1991. – P. 373-392.
5. Fukumi, S. Omatu, and Y. Nishikawa, Rotation-Invariant Neural Pattern Recognition System Estimating a Rotation Angle, *IEEE, Trans., Neural Network*, 8, 1997. – P. 568–581.
6. Hubel D.H., Wiesel T.N. Receptive fields and functional architecture in two nonstriate visual area (18 and 19) of the cat. *J. Neurophysiol.* 28, 1965. – P. 229-289.
7. Shunji Saton, Jousuke Kuroiwa, Hiroto Aso, Shogo Miyake Recognition of rotated

patterns using neocognitron. , IEEE, *Trans., Neural Network*, 9, 1997. – P. 588–597.