

УДК 004.932.75

СИСТЕМА РАСПОЗНАВАНИЯ ЗАШУМЛЁННЫХ И ИСКАЖЁННЫХ ГРАФИЧЕСКИХ ОБРАЗОВ НА ОСНОВЕ НЕЙРОННОЙ СЕТИ ТИПА НЕОКОГНИТРОН

О.И.Федяев¹, Ю.С.Махно²

Разработана программная модель искусственной нейронной сети типа неокогнитрон для распознавания графических образов при наличии различного вида искажений. Рассмотрен алгоритм обучения и особенности функционирования неокогнитрона, отмечены его достоинства и недостатки при распознавании образов, подверженных искажениям.

Введение

Не снижающийся интерес к задачам распознавания обусловлен необходимостью автоматизации как функций контроля и управления сложными динамическими объектами в реальном времени, так и образных процессов коммуникации (визуальных, речевых) в интеллектуальных системах. Поэтому до сих пор продолжается поиск и реализация эффективных принципов передачи распознавательной функции человека компьютеризированным системам. Одно из перспективных направлений решения данной проблемы основано на применении нейрокомпьютеров, как наиболее адекватных классу трудно формализуемых задач распознавания образов. В настоящее время предложено большое количество нейросетевых парадигм для решения задач распознавания образов. Значительные затруднения при распознавании вызывают образы, подверженные какому-либо искажению (зашумлению, смещению, повороту, изменению образа в размерах).

Перспективу в преодолении этих трудностей связывают с применением новой нейросетевой парадигмы – модели неокогнитрона, использующей качественно новую архитектуру и неконтролируемое обучение [Fukushima, 1980, Уоссермен, 1992, Федяев, 2002, 2003]. В основу архитектуры неокогнитрона положена организация зрительной системы человека. Первая (классическая) модель неокогнитрона была

¹ Украина, 83000, Донецк, ул. Артёма 58, ДонНТУ, fedyayev@r5.dgtu.donetsk.ua

² Украина, 83000, Донецк, ул. Артёма 58, ДонНТУ, ymakhno@gmail.com

матрице реагируют на один и тот же образ. То есть, каждый из них настраивается на один специфический входной образ (шаблон).

В неокогнитроне выделяют два типа слоёв: «простые» (S-слои, от слова simple) и «комплексные» (C-слои, от слова complex), которые состоят соответственно из «простых матриц» (распознают входную информацию) и «комплексных матриц» (обобщают распознанную информацию). Каждый нейрон как простого, так и комплексного слоя получает сигналы не от всех нейронов, а лишь от некоторых, с которыми он связан. Такие нейроны образуют «область видимости» [Sato, 1998, Fausett, 2000].

При определении математической модели неокогнитрона для каждого нейрона простой матрицы вводится дополнительный тормозящий (или подавляющий) нейрон v . На тормозящий нейрон возлагается функция подавления выходного сигнала нейрона при несоответствии входного образа образу, на распознавание которого настроен нейрон. Подавляющий нейрон связан с нейроном простого слоя и с теми же нейронами комплексного слоя, с которыми связан сам простой нейрон.

Для математического описания модели введём два вектора. Вектор n задаёт положение нейрона в стеке матриц (рис.1):

$$\bar{n} = (n_x, n_y, n_q). \quad (1.1)$$

Вектор k задаёт смещение координат нейрона относительно рассматриваемого нейрона (вектор n) в стеке матриц предыдущего слоя. Он также имеет 3 координаты:

$$\bar{k} = (k_x, k_y, k_q). \quad (1.2)$$

Между этими двумя векторами определим оператор \oplus_M следующим образом:

$$\left\{ \begin{array}{l} n_x \oplus_M k_x = n_x + k_x \\ n_y \oplus_M k_y = n_y + k_y \\ n_q \oplus_M k_q = (n_q + k_q) \bmod M \end{array} \right. \quad (1.3)$$

Обозначим через параметры T_c и T_s количество матриц в стеках матриц предыдущего слоя. С помощью множества A определим все нейроны, входящие в область видимости рассматриваемого нейрона. Это множество состоит из трёхмерных векторов, каждый из которых задаёт относительное смещение нейрона области видимости от вектора n .

Для расчёта выходного сигнала подавляющего нейрона используется следующая формула:

$$v(\bar{n}) = \sqrt{\sum_{k \in A} t(\bar{k}) [c(\bar{n} \oplus_{T_c} \bar{k})]^2}, \quad (1.4)$$

где c – выходные сигналы нейронов предыдущего комплексного слоя; t – весовые коэффициенты от нейронов C -слоя к подавляющему нейрону.

Сигнал, подаваемый на вход простого нейрона, вычисляется по формуле:

$$x(\bar{n}) = \frac{1 + \sum_{k \in A} c(\bar{n} \oplus_{T_c} \bar{k}) \cdot w(\bar{k})}{1 + v(\bar{n}) \cdot w_v} - 1, \quad (1.5)$$

где w – настраиваемые коэффициенты; v – выходной сигнал подавляющего нейрона (1.4).

Функция активации нейронов простого слоя имеет вид:

$$S(x) = \begin{cases} x, & x \geq 0 \\ 0, & \text{иначе} \end{cases}, \quad (1.6)$$

Входной сигнал для нейронов комплексного слоя вычисляется по формуле:

$$y(\bar{n}) = \sum_{k \in A} s(\bar{n} \oplus_{T_s} \bar{k}) u(\bar{k}), \quad (1.7)$$

где s – выходной сигнал простых нейронов; u – весовые коэффициенты от простых нейронов к комплексному нейрону [Fausett, 2000].

Для нейронов комплексных слоёв использовались различные функции активации.

2. Обучение неокогнитрона

Для обучения неокогнитрона применялась стратегия «обучение без учителя». Настраиваемыми коэффициентами выступали 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. Предполагается, что значения этих коэффициентов должны уменьшаться с удалением от центра области видимости.

При настройке весовых коэффициентов использовалась общепринятая схема расчёта параметра Δ , зависящего от величины входных сигналов [Satoh, 1998, Fausett, 2000]. В итоге применялись следующие формулы корректировки коэффициентов для нейрона заданного вектором n :

$$\Delta w(\bar{k}) = a \cdot t(\bar{k}) c(\bar{n} \oplus_{T_c} \bar{k});$$

$$w^{(m+1)}(\bar{k}) = w^{(m)}(\bar{k}) + \Delta w(\bar{k})$$
(2.1)

где a – коэффициент скорости обучения.

Величина Δ для весового коэффициента подавляющего нейрона v определялась так, чтобы процесс обучения мог завершиться. Поскольку нет единого мнения по поводу наиболее подходящей формулы для расчёта Δ , то оценивались следующие два возможных способа.

Первый способ основывается на формуле $\Delta w_v = a c(\bar{n} \oplus_{T_c} \overline{mid})$, где \overline{mid} – вектор смещений координат центрального нейрона области видимости относительно подавляющего нейрона \bar{n} [Fausett, 2000].

Второй способ основывается на формуле $\Delta w_v = a \frac{v}{2}$ [Sato, 1998]. В этом случае величина Δ зависит от выхода тормозящего нейрона, а сам выход тормозящего нейрона зависит от всех нейронов предыдущего слоя, с которыми связан обучаемый нейрон. В программной модели применялся второй способ из-за его универсальности.

Обучение неокогнитрона происходит слой за слоем. При обучении нейронов использовался метод, при котором в каждой матрице обучается только один нейрон (обычно используется центральный нейрон). Критерий завершения обучения простого нейрона имеет следующий вид:

$$|x(\bar{n})^{(k+1)} - x(\bar{n})^{(k)}| \leq eps, \quad (2.2)$$

где eps – заданное пороговое число. После завершения обучения значения весовых коэффициентов этого нейрона распространяются на все остальные нейроны в этой матрице [Fausett, 2000].

Для обучения нейронов первого слоя шаблон обучения (типовой фрагмент распознаваемого образа) записывается на входном слое. Затем начинается обучение центрального нейрона очередной матрицы. После того, как матрицы первого слоя подобным образом будут настроены на распознавание всего заданного набора шаблонов, переходят к обучению второго слоя.

3. Анализ процесса распознавания

Для исследования процессов обучения и распознавания была разработана программная модель, эмулирующая работу нейронной сети типа R-неокогнитрон на персональном компьютере со следующими характеристиками: процессор Intel Core 2 Duo E6750 с тактовой частотой

2,66 GHz; оперативная память 2 Gb; операционная система Windows XP Professional.

Главная форма программы-эмулятора представлена на рис. 2.

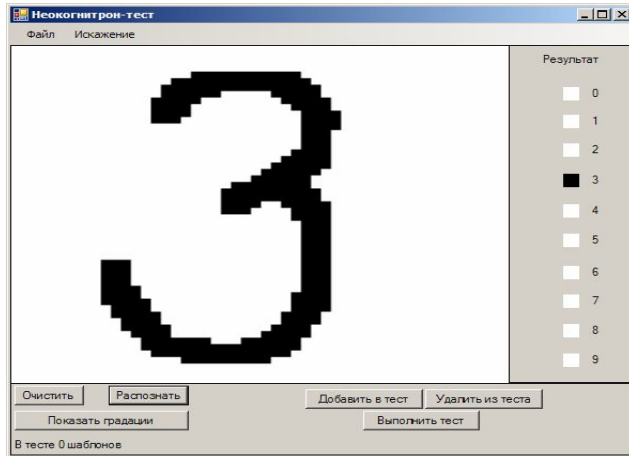


Рис. 2. Окно программного эмулятора R-неокогнитрона

Архитектура R-неокогнитрона определялась следующими параметрами: число слоёв – 4; количество стеков матриц в слое S1 – 16, слое S2 – 26, слое S3 – 78, слое S4 – 42, слое C1 – 12, слое C2 – 26, слое C3 – 54, слое C4 – 10.

Для обучения неокогнитрона распознаванию рукописных арабских цифр использовались шаблоны обучения из открытой базы данных ETL-1 [Sato, 1998, ETL-1]. Обучающее множество включало в себя 500 графических образов (по 50 образов для каждой цифры).

Анализ процесса распознавания был проведен на классах образов, обеспечивающих проверку чувствительности неокогнитрона к следующим искажениям: изменению масштаба; изменению местоположения; присутствию хаотических шумов; повороту исходного образа; наложению образов. Тестовое множество состояло из 400 образов.

При изменении масштаба (рис.3) точность распознавания составила 92%. Ошибки в распознавании происходили в основном при сильном уменьшении масштаба.

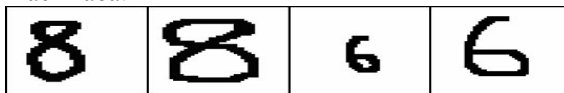


Рис. 3. Распознавание при изменении масштаба

Неокогнитрон является нейронной сетью, нечувствительной к местоположению образов (рис.4). Это было подтверждено результатами моделирования. Неокогнитрон показал 100% распознаваемость на предложенных образах.

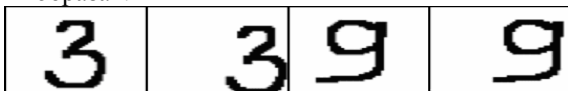


Рис. 4. Распознавание при изменении местоположения

Отличные характеристики неокогнитрон показал и при наложении случайных шумов (рис.5). Были правильно распознаны 94% образов. Ошибки возникали лишь в случаях заполнения случайными шумами более 20% всей рабочей области.

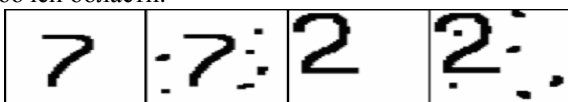


Рис. 5. Распознавание при добавлении хаотических шумов

Данная модель неокогнитрона также позволяет распознавать образы, повернутые не более чем на 70 градусов от своего исходного положения (рис.6). При повороте образов неокогнитрон показал результат распознавания, равный 79%.

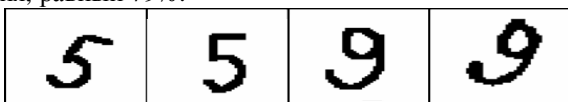


Рис. 6. Распознавание при повороте исходного образа

Одним из недостатков неокогнитрона является то, что он неправильно реагирует на образы, которые были получены путём наложения одного образа на другой.

Как показало моделирование, точность распознавания значительно зависит от состава обучающего множества и параметров процесса обучения. Как было описано выше, для обучения тормозящего нейрона можно использовать два способа. В табл. 1 представлена зависимость качества распознавания от выбранного способа обучения тормозящего нейрона.

Скорость обучения, которая определяется параметром α , имеет особую важность при обучении неокогнитрона на большом количестве образов. В табл. 2 представлена зависимость скорости обучения от значения

параметра α . Стоит заметить, что значение данного параметра не оказывает влияния на точность распознавания.

Табл. 1. Зависимость качества распознавания от выбранного способа обучения тормозящего нейрона

Способ обучения нейрона ν	Точность распознавания, %
Способ 1	84
Способ 2	98

Табл. 2. Зависимость скорости обучения от значения параметра α

Параметр α	Время обучения, с
0,05	94
0,10	91
0,25	102
0,50	129

В табл. 3 показано влияние мощности обучающего множества на точность работы системы распознавания. Обучающие образы из этого множества были неоднородными и охватывали характерные черты каждого из классов.

Табл. 3. Зависимость качества распознавания от мощности обучающего множества

Мощность обучающего множества	Точность распознавания, %
100	90
200	93
300	96
400	97
500	98

Выводы

Моделирование процессов обучения и распознавания на программной модели R-неокогнитрона показало перспективы (по сравнению с обычным неокогнитроном [Федяев, 2008]) распознавания графических образов в условиях различных видов помех, в том числе и при изменении ориентации распознаваемых образов. R-неокогнитрон можно рассматривать как систему индуцирования знаний из обучающего

множества графических данных посредством выделения типовых признаков и их обобщения на последующих уровнях.

Обобщая результаты исследований, отметим ряд свойств R-неокогнитрона, благодаря которым он превосходит другие нейронные сети, ориентированные на распознавание образов:

- малая чувствительность к основным искажениям, которым подвергаются графические изображения (масштабирование, перенос, хаотические шумы);

- способность распознавать двумерные графические образы, повернутые вокруг своей оси;

- высокая скорость обучения и распознавания графических образов.

Вышеперечисленные качества R-неокогнитрона позволяют строить на его основе эффективные системы обработки образной информации в различных предметных областях: распознавание сигналов, идентификация графических изображений, управление с использованием средств визуального наблюдения за объектами.

Список литературы

- [Федяев, 2002] Федяев О.И., Куликов С.А. Многоуровневые нейронные модели типа неокогнитрон. Архитектура, обучение и распознавание // Труды 8-й Всероссийской конф. «Нейрокомпьютеры и их применение» (НКП-2002). – М.: Век книги, 2002. – С.1046-1052.
- [Федяев, 2003] Федяев О.И., Куликов С.А. Применение многоуровневых нейронных сетей типа неокогнитрон для распознавания символов // Сб. тр. междунар. науч. конф. «Нейросетевые технологии и их применение». – Краматорск: ДГМА, 2003. – С. 247-256.
- [Федяев, 2008] Федяев О.И., Махно Ю.С. Распознавание графических образов при наличии искажений с помощью неокогнитронных нейросетей // Тр. VIII междунар. науч. конф. «Интеллектуальный анализ информации ИАИ-2008». – К.: Просвіта, 2008. – С. 512-521.
- [Уоссермен, 1992] Уоссермен Ф. Нейрокомпьютерная техника. Теория и практика. – М.: Мир, 1992. – 260 с.
- [ETL-1] ETL-1 character database, Electrotechnical Laboratory, Japan. – <http://www.is.aist.go.jp/etlcdb/#English>
- [Fausett, 2000] Fausett L. Fundamentals of Neural Networks: Architectures, Algorithms and Applications. Prentice Hall, New Jersey, 2000.
- [Fukushima, 1980] Fukushima K. Neocognitron: a self-organising neural network for mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics 36, 1980. – P. 193-202.
- [Satoh, 1998] Satoh S., Kuroiwa J., Aso H. and Miyake S. Recognition of Hand-written Patterns by Rotation-invariant Neocognitron – Proc. of ICONIP'98, 1998. – P. 295-299.