

УДК 004.04

## Анализ алгоритма самоорганизации нейронной сети типа «неокогнитрон» при распознавании образов

Федяев О.И., Труханов Т.И.

Донецкий национальный технический университет  
fedyaev@r5.dgtu.donetsk.ua

### Abstract

*Fedyayev O., Truhanov T. Analysis of neural network algorithm self-organization of the «neocognitron» in Pattern Recognition. This paper studies the principles of setting a multilayer neural network neocognitron to recognize different patterns. For this purpose, a software model of neocognitron with self-organization learning algorithm was developed. Neocognitron with competitive learning algorithm acquires the ability to correctly recognize patterns at various distortions in their form.*

### Введение

В настоящее время предложено большое количество нейросетевых парадигм для решения задачи распознавания различных изображений. Значительные затруднения при распознавании вызывают образы, подверженные какому-либо искажению (зашумление, смещение, поворот, изменение образа в размерах). Эту проблему решают путём выбора соответствующей архитектуры и способа обучения. Анализ работ показывает, что пока не существует такой модели, которая была бы не чувствительна ко всем четырём видам искажений. 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. Неокогнитрон получает на входе двумерные образы, аналогичные изображениям на сетчатой оболочке глаза, и обрабатывает их в последующих слоях зрительной коре головного мозга человека [1,2].

Анализ существующих методов распознавания показывает, что большинство из них чувствительны к позиционным сдвигам входного образа и его искажениям по форме. Для устранения этого недостатка необходима предварительная нормализация позиции и формы распознаваемого образа. Однако, на сегодняшний день пока не найден хороший метод нормализации.

В 1980 году японским учёным К.Фукушимой (K.Fukushima) [1] предложена серия новых нейросетевых архитектур типа «когнитрон» (cognitron) с алгоритмом их

обучения по стратегии «обучение без учителя». Благодаря специфичной архитектуре и новому принципу самоорганизации мультислойного неокогнитрона он приобретает способность правильно распознавать образы в условиях различного вида помех. Для самоорганизации неокогнитрона на основе неконтролируемого обучения достаточно иметь только набор образов (например, рукописных букв или снимков лиц людей) и не иметь информации о классах (букв, людях), к которым принадлежат эти образцы. Путём многократного предъявления образов нейроны неокогнитрона приобретают селективные способности обнаруживать характерные признаки и их расположение во входных образцах. Нейроны последующих слоёв сети формируют знания о более сложных особенностях образцов из комбинаций простых признаков, выделенных на предшествующих слоях сети. В конце обучения всей сети каждый выходной нейрон неокогнитрона будет настроен реагировать только на определённый образец (или близкий к нему), несмотря на позицию или размер распознаваемого образца.

Основная цель данной работы заключается в разработке программного эмулятора неокогнитрона с возможностью его обучения на распознавание определенного класса образов. Такой эмулятор позволит путём моделирования более детально изучить принципы самоорганизации многосвязной мультислойной структуры неокогнитрона и подбирать значения параметров прикладной модели нейросети для качественного распознавания изображений заданного класса.

### Постановка задачи

На теоретико-множественном уровне параметрическую модель многослойного

неокогнитрона NC (NeoCognitron) можно представить в виде кортежа:

$$NC = \langle X, L, \{ \langle K^l, \{ a_{v,m}^{l,s,p,(i,j)} \}, \varphi^{l,s}, \{ b^{l,s} \}, \psi^{l,s}, \{ d_{v,n}^{l,c,p,(i,j)} \}, \varphi^{l,c}, \{ c^{l,c} \}, \psi^{l,c} \rangle \}, Y \rangle,$$

где  $L$  – количество модулей в неокогнитроне;  $l$  – номер модуля,  $1 \leq l \leq L$ ;  $K^l$  – количество плоскостей в одном слое ( $S$  или  $C$ )  $l$ -го модуля;  $X$ ,  $Y$  – соответственно входы (выходы) неокогнитрона для приёма (снятия) входных (выходных) сигналов;  $a_{v,m}^{l,s,p,(i,j)}$  – весовой коэффициент связи  $m$ -го входа ( $ij$ ) нейрона  $p$ -й плоскости  $S$ -го слоя модуля  $l$  с выходом  $v$ -го нейрона ( $v \in C_p^{l-1}, 1 \leq p \leq K^{l-1}, v$  – координаты нейрона в области связи);  $C_p^{l-1}$  – область связи на  $p$ -й плоскости  $C$  слоя ( $l-1$ -го модуля);  $d_{v,n}^{l,c,p,(i,j)}$  – весовой коэффициент связи  $n$ -го входа ( $ij$ ) нейрона  $p$ -й плоскости слоя  $C$  модуля  $l$  с выходом  $v$ -го нейрона из его рецептивной области ( $v \in S_p^l, 1 \leq p \leq K^l$ );  $S_p^l$  – область связи для ( $ij$ ) нейрона, расположенная на  $p$ -й плоскости  $S$ -го слоя модуля  $l$ .

В слоях  $S$  и  $C$  каждого модуля неокогнитрона имеются тормозящие нейроны, которые в кортеже описываются параметрами  $b^{l,s}, c^{l,c}$ . Это постоянные коэффициенты тормозящего входа для всех нейронов соответственно слоёв  $S$  и  $C$  модуля  $l$ . Функции активации базовых нейронов обозначены символами  $c^{l,c}$  и  $\varphi^{l,c}$ , а функции активации тормозящих нейронов –  $\psi^{l,s}$  и  $\psi^{l,c}$ .

Параметрическая модель и многослойная структура неокогнитрона с последовательными связями позволяет в явном виде получить функциональную зависимость выходных сигналов  $Y$  нейросети от её входных сигналов  $X$ :

$$Y = f(X, NC(a, b)).$$

Задача состоит в разработке такой имитационной (программной) модели неокогнитрона и алгоритма его обучения, которые в совокупности минимизируют функционал  $\Phi(a, b)$  и тем самым настраивают (обучают) его распознавать множество требуемых образов  $\bar{Y}$ :

$$\Phi(a, b) = \| Y - \bar{Y} \|_{a, b, X \in \{(X, \bar{Y})\}} \Rightarrow \min,$$

где  $a, b$  – настраиваемые весовые коэффициенты неокогнитрона;  $Y, \bar{Y}$  – реальные и желаемые выходные сигналы неокогнитрона;  $\| \cdot \|$  – выбранная норма вектора.

### Структура неокогнитрона

Неокогнитрон является иерархической нейронной сетью, состоящей из некоторого числа идущих друг за другом слоёв и имеющих неполные (можно сказать достаточно редкие) связи между слоями. На рис.1 представлена упрощенная структура неокогнитрона.

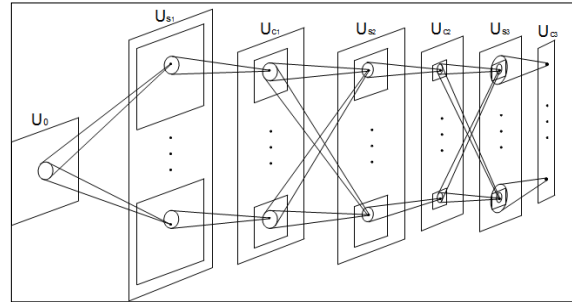


Рисунок 1 – Структура многослойной нейросети типа «неокогнитрон»

На рисунке изображены три модуля нейросети, в каждом из которых имеются два типа слоёв: «простые» ( $S$ -слои, от слова simple) и «комплексные» ( $C$ -слои, от слова complex), которые состоят соответственно из «простых матриц» (распознают характерные признаки во входных образах) и «комплексных матриц» (обобщают распознанную информацию). С целью простоты рисунка на нём не показаны плоскости с тормозящими нейронами. Из приведенной структуры видно, что размер первого слоя совпадает с размером входного образа ( $U_0$ ) и к последнему слою размеры плоскостей уменьшаются. В последнем слое в каждой плоскости находится только по одному нейрону.

Входной слой  $U_0$  обычно представляется как нулевой комплексный слой. 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».

Для построения программной модели неокогнитрона была проведена его объектно-ориентированная декомпозиция. Основные абстракции неокогнитрона взяты из параметрической модели, которая подробно описана в предыдущем подразделе статьи. В результате выделены следующие основные объектные компоненты нейросети:  $S$ -слой,  $C$ -слой, Модуль и Нейросеть.

Логическая структура программной модели неокогнитрона показана на рис. 2 в виде UML-диаграммы классов. На ней каждая отдельная сущность представлена в виде отдельного класса со своими атрибутами и методами. Управляющим классом является класс Нейросеть, который агрегирует в себе объекты класса Модуль,

который в свою очередь агрегирует в себе объекты классов S-слой и C-слой. Обучение нейросети происходит с помощью вызова методов *обучить()*, *выбрать представителя()*, *усилить*

*коэффициенты()*, а распознавание инициируется методами *распознать()* и *посчитать выходы()*.

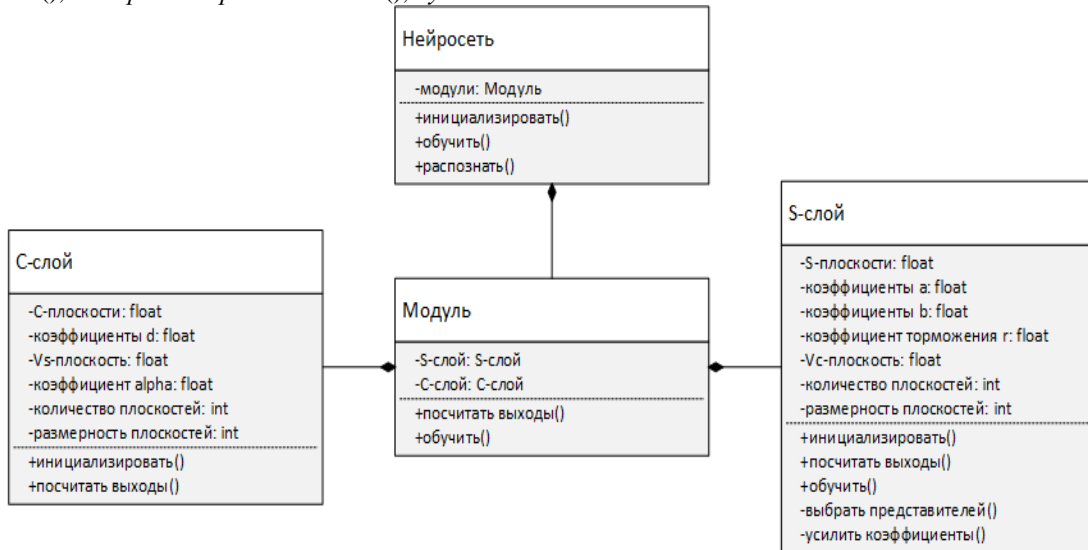


Рисунок 2 - Диаграмма классов программной модели многослойной нейросети типа «неокогнитрон»

**Математическая модель неокогнитрона**

Для вычисления выходных сигналов  $u_{sl}$  нейронов S-плоскости использовалась следующая формула [1]:

$$u_{sl}(k_l, n) = r_l \varphi \left( \frac{1 + \sum_{k_{l-1}=1}^{K_{l-1}} \sum_{v \in S_l} a_l(k_{l-1}, v, k_l) u_{cl-1}(k_{l-1}, n + v)}{1 + \frac{r_l}{1 + r_l} b_l(k_l) v_{cl-1}(n)} - 1 \right)$$

В этой формуле учитываются выходные сигналы тормозящей плоскости  $v_{cl-1}$ , которые вычисляются по формуле:

$$v_{cl-1}(n) = \sqrt{\sum_{k_{l-1}=1}^{K_{l-1}} \sum_{v \in S_l} c_{l-1}(v) u_{cl-1}^2(k_{l-1}, n + v)}$$

где  $L$  – количество модулей в неокогнитроне;  $l$  – номер модуля,  $1 \leq l \leq L$ ;  $a$ ,  $b$  – настраиваемые весовые коэффициенты неокогнитрона;  $\varphi$  – функция активации S-нейронов;  $r_l$  – параметр управления интенсивностью запрещения;  $n$  – позиция нейрона в плоскости  $k_l$ ;  $c_{l-1}$  – фиксированные весовые коэффициенты C-нейронов в модуле  $l-1$ ;  $S_l$  – область связи для S-нейрона, расположенного в позиции  $n$  модуля  $l$ ;  $v$  – позиции нейронов в области связи  $S_l$ ;  $u_{cl-1}$  – выходные сигналы C-нейронов модуля  $l-1$ .

**Алгоритм обучения неокогнитрона на**

**основе самоорганизации**

Для обучения неокогнитрона применяют одну из двух классических стратегий: «обучение с учителем» или «обучение без учителя» [2,3]. Независимо от применяемой стратегии настраиваются только весовые коэффициенты S-плоскостей. Веса нейронов C-плоскостей задаются перед обучением и не изменяются [1]. В целом процесс обучения проводится последовательно, начиная с обучения S-плоскостей 1-го модуля и заканчивая обучением последнего S-слоя.

При распознавании, например, людей по изображениям их лиц необходимо иметь набор характерных признаков, чтобы проводить классификацию лиц с видеокамеры. Такой набор признаков всегда уникальный для каждого множества людей, на распознавание которых будет настроен неокогнитрон. Кроме того, признаки классификации априорно не известны, поэтому нами неокогнитрон рассматривался как самоорганизующаяся нейронная сеть, которая должна уметь самостоятельно в процессе обучения выделять характерные особенности, присущие всем распознаваемым образам, и представлять их в виде иерархической структуры. Способ самоорганизации неокогнитрона, предложенный К.Фукушимой [1], достаточно сложный для исследования аналитическими методами. Поэтому в работе особое внимание уделено программной реализации алгоритма самоорганизации с целью его всестороннего анализа методами имитационного моделирования.

В алгоритме самоорганизации можно условно выделить четыре этапа:

- подача входного образа и вычисление выходных сигналов обучаемого S-слоя;
- выбор кандидатов в представители по S-цилиндрам фиксированной ширины;
- выбор представителей из кандидатов;
- обучение нейронов тех плоскостей, в которых были отобраны представители.

Эти этапы повторяются на каждой итерации процесса обучения. В качестве критерия окончания алгоритма выбран модуль разности между максимальными сигналами S-слоя для каждого распознаваемого образа. Если эта величина не превышает заданное малое значение, то алгоритм завершает обучение текущего модуля и переходит к обучению следующего. При этом входными сигналами каждого следующего модуля служат выходные сигналы предыдущего. В каждой эпохе обучения входные образы подаются по очереди по одному разу.

Этап 1. Вычисление выходных сигналов нейронов на каждом S-слое происходит по формулам из математической модели. Эти вычисления на каждой итерации необходимы, потому что при обучении меняются коэффициенты  $a$  и  $b$ , и, следовательно, меняются и выходные сигналы обучаемого слоя.

Этап 2. После вычисления выходных сигналов S-нейронов находятся кандидаты в представители. Для этого в алгоритме используются, так называемые, S-цилиндры (или S-колонки [1]) (рис. 3). По сути они образованы подмножеством выходных сигналов тех нейронов, которые попадают внутрь цилиндра. Если расположить S-плоскости одна над другой, то получится цилиндр, фактически S-цилиндры - это срез цилиндра по высоте (определяемой количеством плоскостей в S-слое), заданной ширины и с заданными координатами. Эти цилиндры могут быть шириной от одного нейрона до ширины всей плоскости. Они могут пересекаться. Далее находятся нейроны с максимальными выходными сигналами в каждом S-цилиндре. Эти нейроны становятся кандидатами в представители. В одной плоскости могут оказаться более одного кандидата.

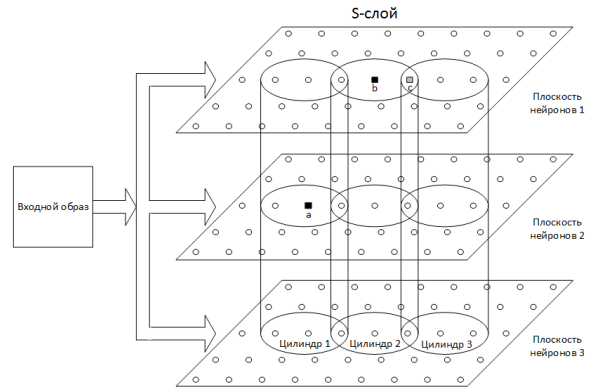


Рисунок 3 - Схема самоорганизации S-плоскостей нейронов на выделение характерных признаков во входном образе:

- - нейроны-кандидаты (a, b, c) в цилиндрах;
- - нейроны-представители (a, b) на плоскостях

Этап 3. После того как определены кандидаты в представители необходимо из них выбрать собственно представителей на обучение. Для этого все нейроны-кандидаты разбиваются на группы по плоскостям и в этих группах находится нейрон с максимальным выходным сигналом. Таким образом, должно получиться не более одного нейрона-представителя на каждой S-плоскости. Поскольку нейроны-представители могут появиться сразу в нескольких плоскостях, то в течение одной итерации можно сразу обучить несколько плоскостей, что значительно ускоряет процесс обучение в целом.

Этап 4. Плоскости, в которых есть нейроны-представители, обучаются по следующим формулам:

$$\Delta a_l(k_{l-1}, \vartheta, \hat{k}_l) = q_l c_{l-1}(\vartheta) u_{cl-1}(k_{l-1}, \hat{n} + \vartheta),$$

$$\Delta b_l(\hat{k}_l) = q_l v_{cl-1}(\hat{n}),$$

где  $q_l$  – положительная константа, определяющая скорость обучения;  $\Delta a$ ,  $\Delta b$  – величины корректировки весовых коэффициентов S-нейронов по возбуждающему и тормозящему каналам;  $\vartheta$  - позиции нейронов в области связи  $S_l$ .

После этого алгоритм неконтролируемого обучения повторяется с первого этапа, пока не выполнится критерий окончания.

### Анализ алгоритма самоорганизации

Программная модель неокогнитрона написана на языке программирования Python [6]. Он хорошо подходит как для быстрого написания прототипов программ, так и для разработки полноценно работающих приложений. Python широко используется в научно-исследовательских проектах. Это объясняется в первую очередь простотой языка, а также наличием хорошо протестированных библиотек, которые позволяют

быстро и эффективно решать прикладные задачи. В работе использовались следующие библиотеки: NumPy (для эффективных вычислений), Matplotlib (для визуализации информации), PIL (для работы с изображениями).

Разработанная программная модель подробно визуализировала динамику обучения и конечное состояние сети. В частности на экран выводились результаты обучения сети на распознавание конкретного изображения, в частности, текущие значения выходных сигналов и весовых коэффициентов во всех слоях нейронной сети. Настраиваемыми коэффициентами выступали коэффициенты от нейронов комплексного слоя к нейронам простого слоя и коэффициент от тормозящего нейрона к нейрону простого слоя. Значения коэффициентов от комплексных нейронов к тормозящему нейрону и от простых нейронов к комплексным остаются фиксированными.

Таким образом, можно увидеть, какие признаки выделяются S-плоскостями на том или ином этапе обучения или распознавания. Для детального изучения внутренней самоорганизации неокогнитрона в данной работе были исследованы модели неокогнитрона, настроенные на распознавание простейших печатных символов Т, С, Н, Г, П.

Во всех описанных ниже экспериментах использовался неокогнитрон, структура которого состояла из трёх модулей. В начале он обучался на символы Т, С, Н, Г и П, расположенные по центру входного слоя  $U_0$ , которые не имели искажений. Рассмотрим детально несколько примеров распознавания буквы «Н». На рис. 4 показаны изменяемые коэффициенты  $a$  для тех плоскостей, которые отреагировали на образ буквы «Н».

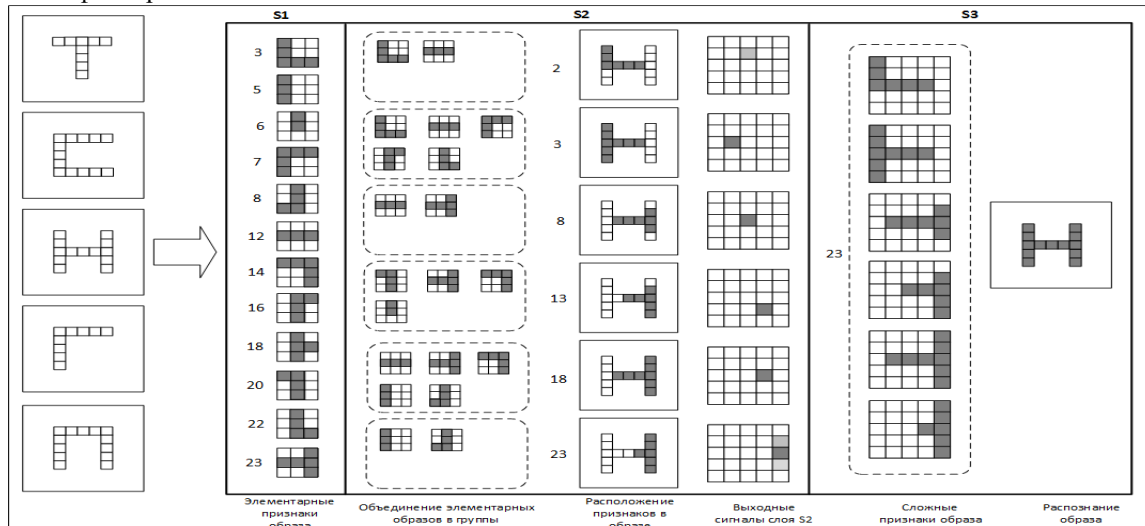


Рисунок 4 – Иллюстрация процесса распознавания символа «Н»

Как видно из рисунка, на первом модуле нейросеть выделила из образа простейшие характерные признаки. На втором модуле показаны выходные сигналы S-слоя и характерные признаки первого S-слоя, объединенные в группы. Второй и последующие слои получают входные сигналы со всех плоскостей предыдущего модуля, за счёт чего они формируют комплексные характерные признаки, состоящие из более простых признаков, выделенных на первом модуле.

На третьем модуле нейроны также получают входные сигналы со всех плоскостей предыдущего модуля, поэтому формируют более сложные характерные признаки из комплексных признаков второго модуля. В данном случае на третьем модуле нейросеть уже полностью распознаёт образ.

За счёт наличия S-слоя неокогнитрон способен различать образы даже со смещениями и искажениями. На рис. 5 приведен пример распознавания символов со сдвигами.

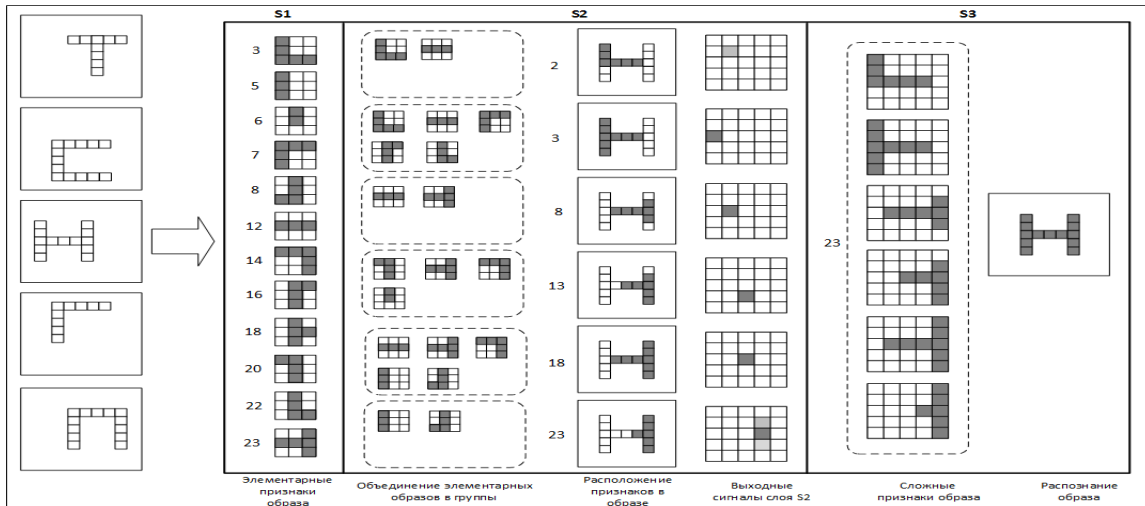


Рисунок 5 – Иллюстрация процесса распознавания символа «Н» со сдвигом влево

Из рисунка видно, что все характерные признаки выявились как в первом, так и во втором модуле. Однако следует отметить, что выходные сигналы второго модуля заметно слабее по сравнению с распознаванием образа без сдвига (см. рис. 4). Сила сигнала на рисунке изображена градацией серого - чем светлее сигнал, тем он слабее.

Однако, несмотря на слабость сигнала, нейросеть всё равно распознала символ «Н» хотя и с меньшей уверенностью. На сопротивление сдвигам и искажениям влияет большое количество различных параметров неокогнитрона, но сильнее всего: коэффициент торможения ( $r$ ),

размерность области видимости С-слоя ( $D$ ) и значения фиксированных коэффициентов С-слоя ( $d$ ).

Кроме того, чтобы в данном случае повысить точность распознавания нужно увеличить размерность плоскостей на внутренних модулях, что позволит охватить все сигналы первых модулей, не «обрезая» их.

В результате проведенного следующего опыта, показанного на рис. 6, выяснилось, что при внесении искажений во входной (распознаваемый) образ, нейросеть перестаёт обнаруживать некоторые характерные признаки.

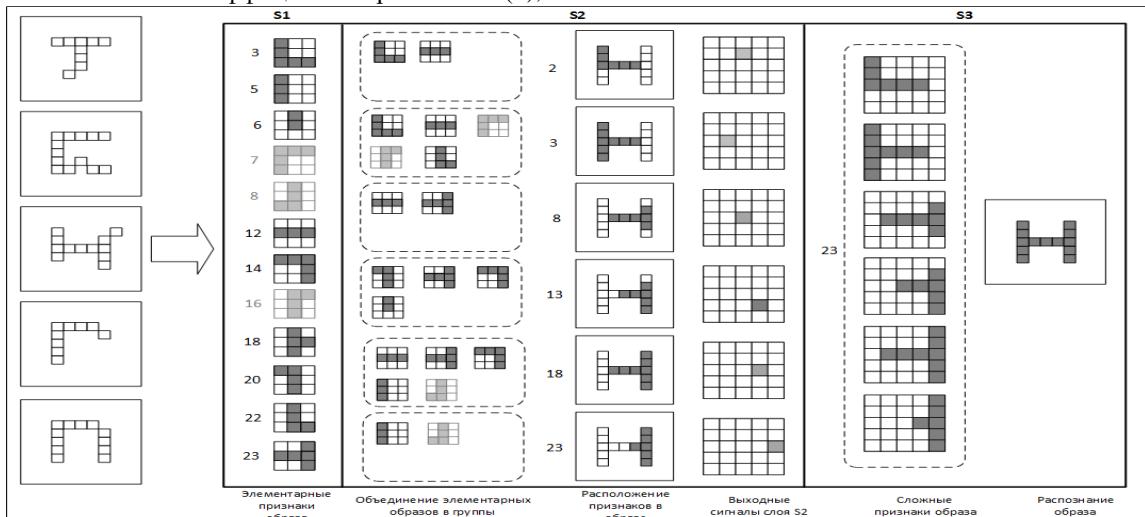


Рисунок 6 – Иллюстрация процесса распознавания символа «Н» с искажениями

Ещё часть нейронов начинает реагировать слабее обычного. Это приводит к заметному уменьшению сигналов на выходах внутренних слоёв.

Также как и в случае со сдвигом, сигналы на втором модуле заметно уменьшились. Однако нейросеть всё равно продолжает правильно

распознавать даже искажённые образы за счёт других характерных признаков.

Выходные сигналы нейронов тех плоскостей, которые отреагировали на символ «Н», во всех слоях неокогнитрона можно посмотреть на рис. 7. По номерам плоскостей



можно определить на какой признак они среагировали (см. рис. 4–6).

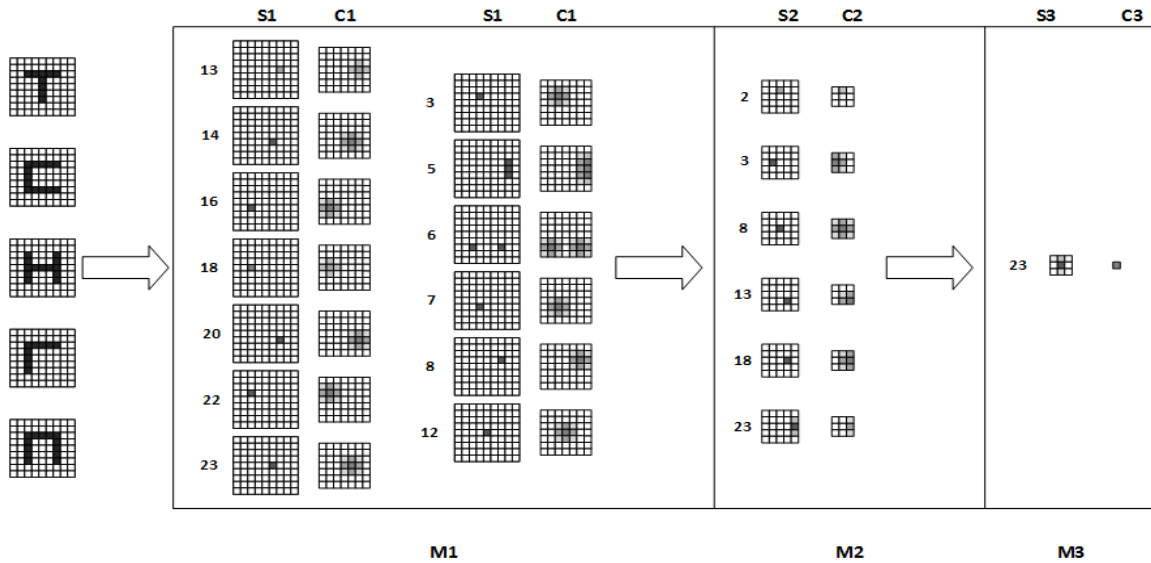


Рисунок 7 – Детальная иллюстрация выходных сигналов плоскостей, среагировавших на символ «Н»

Из этого иллюстративного примера хорошо видно, что в процессе самоорганизации неокогнитрон выделил на первом слое характерные признаки во входных образах. Эти признаки зафиксированы в значениях весовых коэффициентах нейронов соответствующих плоскостей, которые при распознавании будут выступать в роли «экспертов» по обнаружению своих признаков во всех позициях входного изображения и на их основе правильно его классифицировать независимо от его позиции и формы.

**Выводы**

В статье проанализирован принцип настройки многослойной нейронной сети типа неокогнитрон на распознавание различных образов. Для этого была разработана программная модель неокогнитрона с алгоритмом обучения нейронов на основе их самоорганизации. Неокогнитрон, настроенный по алгоритму конкурентного обучения, приобретает способность правильно распознавать образы при различных искажениях в их форме.

На модели проведен ряд экспериментов по изучению взаимовлияния каналов усиления и торможения нейронов сети, а также изучены особенности неконтролируемого обучения S-нейронов.

Для более детального изучения процессов, происходящих при обучении неокогнитрона, а также при распознавании графических образов были проведены ряд экспериментов на простейших буквенных образах. В результате этих

экспериментов было улучшено понимания процессов, происходящих внутри нейросети. Кроме того были получены важные данные по влиянию различных параметров нейросети на её поведение во время обучения и распознавания. На примере графического образа буквы был проиллюстрирован процесс самоорганизации и приведены получившиеся коэффициенты *a* в наглядном виде, что позволило лучше прояснить логику формирования характерных признаков на первом модуле нейросети.

По результатам этих экспериментов можно сделать вывод о том, что полученная модель неокогнитрона позволяет исследовать сложные алгоритмы неконтролируемого обучения и распознавания, а объектно-ориентированная концепция в дальнейшем обеспечит успешное наращивание функциональных и сервисных возможностей программного эмулятора неокогнитрона.

Как известно, при решении задачи распознавания образов приходится сталкиваться с рядом проблем. Самая существенная из них – это выделение и учёт топологии входного изображения. Другой проблемой является вариация представления одного и того же образа, а также его деформация. Эти проблемы можно решать благодаря архитектуре и самоорганизации нейронной сети типа неокогнитрон [4, 5, 7].

В перспективе на основе математического описания можно будет сформулировать и более точно решить задачу оптимизации многочисленных параметров модели неокогнитрона, адекватных природе распознаваемых образов.

**Литература**

1. Kunihiro Fukushima, Sei Miyake Neocognitron: a new algorithm for pattern recognition tolerant of deformations and shifts in position — Great Britain: Pattern Recognition Vol 15, No 6, 1982. — pp 455-469.
2. Уоссермен Ф. Нейрокомпьютерная техника. Теория и практика. — М.: Мир, 1992.
3. Руденко О.Г., Бодянский Е.В. Основы теории искусственных нейронных сетей. Уч. пособие. — Харьков: ТЕЛТЕХ, 2002. — 317 с.
4. 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.
5. Федяев О.И., Махно Ю.С. Распознавание графических образов при наличии искажений с помощью неокогнитронных нейросетей // Тр. 8-й междунар. науч. конф. «Интеллектуальный анализ информации ИАИ-2008». — К.: Просвіта, 2008. - С. 512-521.
6. Труханов Т.И., Федяев О.И. Модель многослойной нейросети с архитектурой неокогнитрона // Інформаційні управляючі системи та комп'ютерний моніторинг / Матеріали IV міжнародної науково-технічної конференції — Донецьк, ДонНТУ — 2013, Том 1, с. 456-462.
7. Труханов Т.И., Федяев О.И. Алгоритм самоорганизации неокогнитрона с целью формирования иерархической системы признаков из распознаваемых образов // Нейросетевые технологии и их применение / Материалы двенадцатой всеукраинской научной конференции с международным участием — Краматорск, ДГМА — 2013. — С. 123-128.

*Статья поступила в редакцию 20.09.2015  
Рекомендована к публикации д-ром техн. наук В.Н. Павлышом*