

СПИСОК ЛИТЕРАТУРЫ

1. Макушкин Ю.С., Тютюрев Вл.Г. Методы возмущений и эффективные гамильтонианы в молекулярной спектроскопии. Новосибирск: Наука, 1984. 240 с.
2. Тютюрев Вл.Г. Контактные преобразования и динамические переменные // Молекулярная спектроскопия высокого и сверхвысокого разрешения. Новосибирск: Наука, 1976. С. 93-115.
3. Жилинский Б.И., Перевалов В.И., Тютюрев Вл.Г. Метод неприводимых тензорных операторов в теории спектров молекул. Новосибирск: Наука, 1987. 231 с.
4. Лободенко Е.И. и др. Резонансы Кориолиса в молекулах симметрии C_{3v} // Оптика и спектроскопия. 1987. Т. 62. Вып. 1. С. 54-58.
5. Lobodenko, E.I. et al. Reduced Effective Hamiltonian for Coriolis-Interacting ν_n and ν_l Fundamentals of C_{3v} Molecules // J. Mol. Spectrosc. 1987. V. 126. N 1. P. 159-170.
6. Watson, J.K.G. Determination of centrifugal distortion coefficients of asymmetric — top molecules // J. Chem. Phys. — 1967. V. 46. № 5. P. 1935-1949; Higher-degree centrifugal distortion of tetrahedral molecules // J. Mol. Spectrosc. 1975. V. 55. № 3. P. 498-499.
7. Курош А.Г. Курс высшей алгебры. М.: Наука, 1975. 302 с.
8. Арнольд В.И. Геометрия комплексных чисел, кватернионов и спинов. М.: МЦНМО, 2009. 40 с.
9. Макушкин Ю.С., Тютюрев Вл.Г. Колебательно-вращательные взаимодействия в молекулах типа асимметричного волчка // Молекулярная спектроскопия высокого и сверхвысокого разрешения. Новосибирск: Наука, 1976. С. 66-92.
10. Березин А.В., Курочкин А.В., Толкачев Е.А. Кватернионы в релятивистской физике. М.: УРСС, 2003. 200 с.

*Александр Григорьевич ИВАШКО —
зав. кафедрой информационных систем
Института математики и компьютерных наук
Тюменского государственного университета,
доктор технических наук, профессор*

*Александр Владимирович КУГАЕВСКИХ —
гл. специалист отдела автоматизации
филиала ОАО АКБ «ЮГРА» (г. Тюмень)
a-kugaevskikh@yandex.ru*

УДК 004.93'14

ВОЗМОЖНОСТЬ ПРИМЕНЕНИЯ ИСКУССТВЕННЫХ НЕЙРОННЫХ СЕТЕЙ «НЕОКОГНИТРОН» ДЛЯ РАСПОЗНАВАНИЯ ДРЕВНЕЕГИПЕТСКИХ ИЕРОГЛИФОВ POSSIBILITY OF NEOCOGNITRON APPLICATION FOR RECOGNITION OF ANCIENT EGYPTIAN HIEROGLYPHS

АННОТАЦИЯ. В статье рассмотрена возможность применения аппарата искусственных нейронных сетей на примере неоконитрона для задачи распознавания древнеегипетских иероглифов.

SUMMARY. The given article considers the possibility of application of a means of artificial neural networks on the basis of neocognitron used for recognition of ancient Egyptian hieroglyphs.

КЛЮЧЕВЫЕ СЛОВА. Искусственные нейронные сети, неоконитрон, распознавание символов, распознавание образов.

KEY WORDS. Artificial neural networks, neocognitron, character recognition, a pattern recognition.

Хрупкость материальных носителей иероглифических текстов и графическая сложность самих иероглифов ставят перед специалистами проблему эффективного копирования текстов для их дальнейшего изучения. Фотокопии полностью не решают данную проблему ввиду больших размеров файлов, что обуславливает трудность передачи информации. Решением данной проблемы стало создание иероглифических текстовых процессоров и разработка стандарта хранения иероглифических текстов MdC (Manuel de Codage) [1]. В работах [2, 3] была предложена концепция представления символов как элементов шрифта и описана соответствующая архитектура иероглифического процессора. Тем не менее, текстовые процессоры решают проблему переноса текстов лишь частично, так как ввод текстов осуществляется вручную. Решением данной проблемы служит OCR (Optical Character Recognition)-система, состоящая из трех подсистем: сегментации текста, предобработки выделенных символов и подсистемы распознавания символов. В настоящий момент подобной OCR-системы для древнеегипетского языка не существует. Проблема с распознаванием иероглифов состоит в их графической сложности и многообразии самих символов.

Задача распознавания образов ставится следующим образом. Имеется некоторый способ кодирования символов, принадлежащих известному множеству классов $C = \{C_1, \dots, C_{6845}\}$, и множество объектов (обучающее множество), про каждый из которых известно, какому классу он принадлежит. Обучающее множество состоит из правильных начертаний каждого из 6845 иероглифов [4]. Нужно построить алгоритм, который по любому входному объекту, необязательно принадлежащему обучающему множеству, решает, какому классу этот объект принадлежит, и делает это достаточно хорошо [5].

Согласно работе [6] большинство алгоритмов распознавания правильно классифицируют не более 82% латинских символов. В работе [7] описана искусственная нейронная сеть «Неокогнитрон», которая, по словам автора, на латинских символах дает качество распознавания в 98%. В рамках настоящей работы рассматривается возможность применения неокогнитрона в распознавании иероглифов.

Общая иерархическая структура сети представлена на рис. 1. Процесс распознавания проходит следующие стадии: $U_0 \rightarrow U_G \rightarrow U_{S1} \rightarrow U_{C1} \rightarrow U_{S2} \rightarrow U_{C2} \rightarrow U_{S3} \rightarrow U_{C3} \rightarrow U_{S4} \rightarrow U_{C4}$.

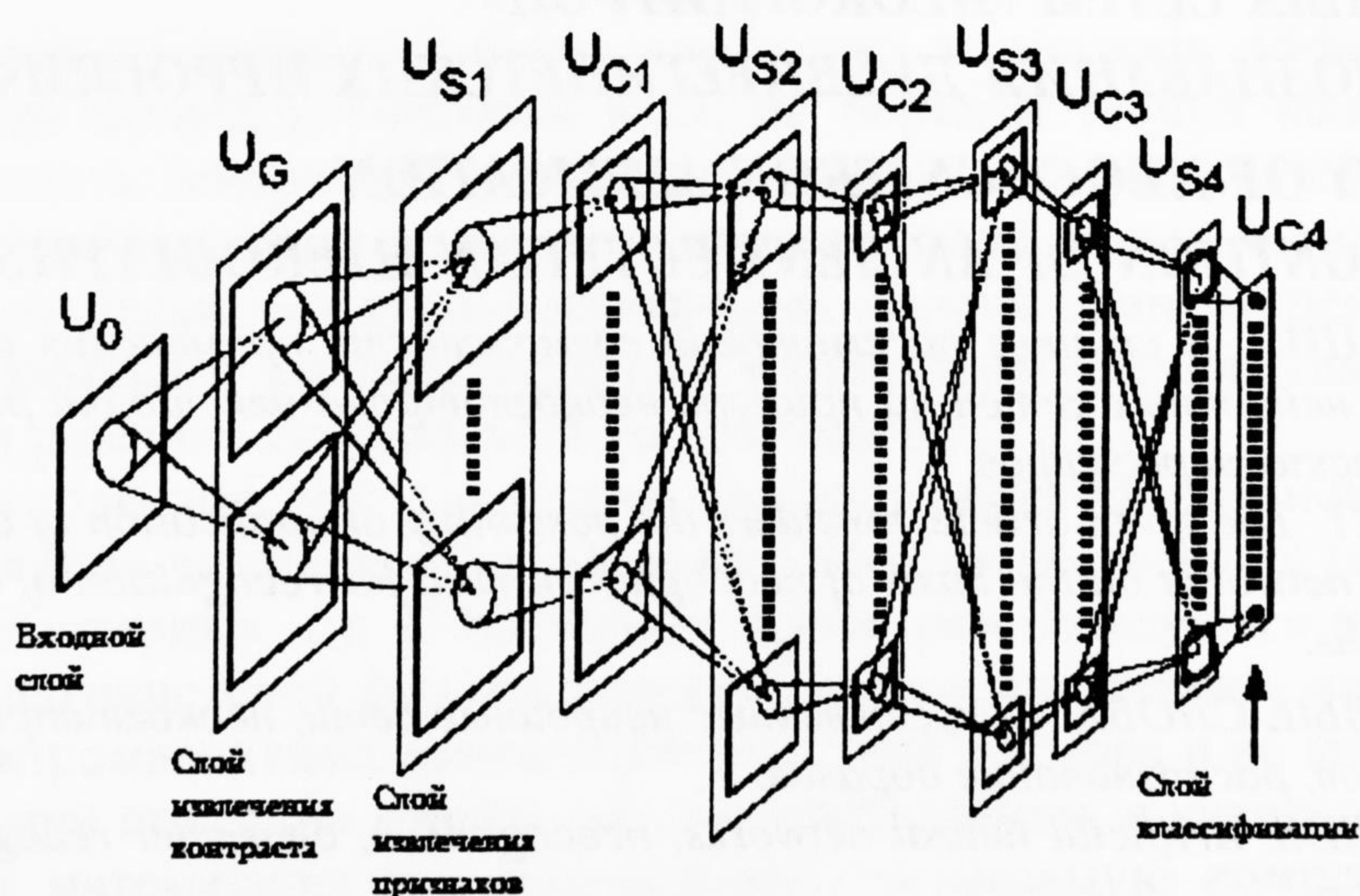


Рис. 1. Структура ИНС «Неокогнитрон»

Первый слой сети (U_0) представляет собой слой фоторецепторных нейронов, которые непосредственно реагируют на входящие образы. На изображение символа, поступающее на вход сети, налагается ряд ограничений. Во-первых, изображение должно быть строго определенного размера (51×51 пиксель), во-вторых, контур символа должен быть минимально возможной толщины (1-3 пикселя), в-третьих, изображение должно быть монохромным. Так как образ монохромен, то слой U_0 представлен в виде двумерного массива, где каждый элемент принимает значение 0 или 1 в зависимости от цвета соответствующего пикселя.

Слой извлекающих контраст ячеек (U_G) состоит из двух плоскостей: первая плоскость содержит ячейки с концентрическими рецептивными полями с оп-центром (возбуждаются при освещении центра рецептивного поля ячейки), другая плоскость состоит из ячеек с концентрическими рецептивными полями с off-центром (возбуждаются при его затемнении).

$$u_G(n_G, k_G) = \varphi \left[-1^k \sum_{|\nu| < A_G} a_G(\nu) u_0(\nu) \right] \quad (1)$$

где ν — индекс связи; $u_0(\nu)$ — выход ячейки U_0 , связанной связью ν ; $\varphi[x] = \max(x, 0)$; $a_G(\xi)$ сила фиксированной связи с ячейкой; k_G — номер плоскости: с оп-центром ($k_G = 2$), с off-центром ($k_G = 1$); A_G — радиус рецептивного поля нейрона, в данной работе он равен 3.3.

Для расчета массива параметров a_G чаще всего применяются вейвлеты [8]. В данной работе применяется вейвлет «сомбреро» [9]. Он обладает весьма специфичной функцией, пик которой приходится на центр рецептивного поля (радиус равен 1.2), а отрицательные значения — на тормозную зону. Размер каждой плоскости составляет 65×65 ячеек.

Каждый S- и C-слой нейронов разделен на подгруппы, называемые плоскостями в соответствии с признаками образа, которые они выделяют, а каждая плоскость нейронов организована в 2-мерный массив ячеек.

S-ячейки слоя U_{S1} соответствуют простым ячейкам в первичной зрительной зоне коры головного мозга и выделяют характерные признаки символа (4-9 на символ). Ячейки последующих S-слоев, кроме U_{S4} , выделяют детали признаков, выделенных слоем U_{S1} .

В результате обучения S-нейрон активизируется только тогда, когда в его рецептивном поле оказывается именно тот признак образа, на который обучали данный нейрон. Каждый C-нейрон получает в качестве входа сигналы, поступающие от группы S-нейронов, которые выделяют одинаковый признак образа, но с небольшими смещениями в положении рецептивного поля. Таким образом, C-нейрон распознает выделяемый элемент образа, даже если он немного сдвинут в своем положении, что обеспечивает инвариантность к сдвигам.

Уровень U_{C4} является уровнем распознавания, реакция которого показывает конечный результат распознавания образов сетью в виде имени класса.

Схематично связи между ячейками отображены на рис. 2.

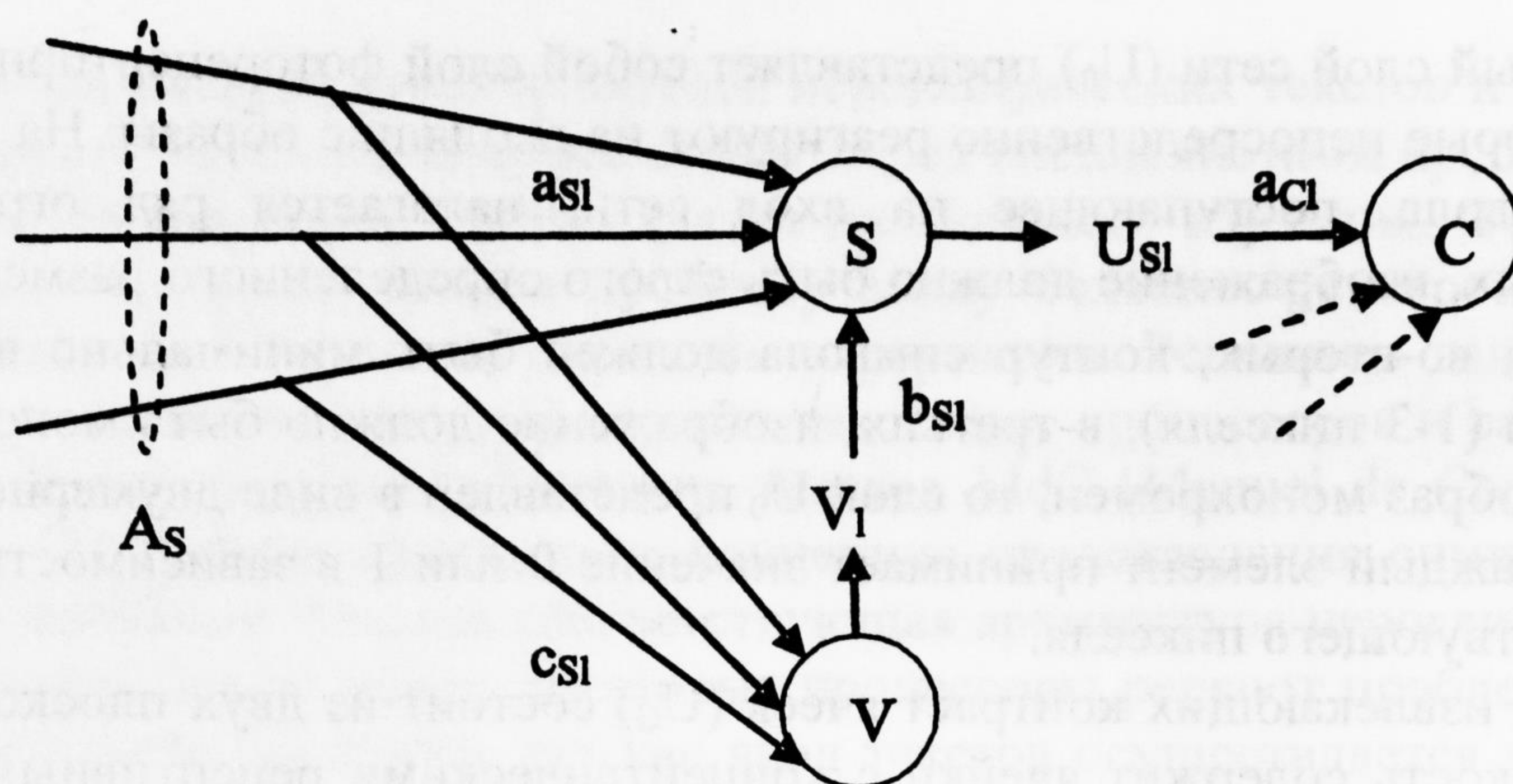


Рис. 2. Схема связей между ячейками в слое

Выход S- и V-ячеек соответственно равен:

$$u_{Sl}(n, k) = \frac{\theta_l}{1 - \theta_l} \varphi \left[\frac{1 + \sum_{k=1}^{K_{Cl-1}} \sum_{|v| < A_{Sl}} a_{Sl}(v, \kappa, k) u_{Cl-1}(v, \kappa)}{1 + \theta_l b_{Sl}(k) v_l(n)} \right] - 1 \quad (2)$$

$$v_l(n) = \sqrt{\sum_{\kappa=1}^{K_{Cl-1}} \sum_{|v| < A_{Sl}} C_{Sl}(v) \{u_{Cl-1}(v, \kappa)\}^2} \quad (3)$$

где κ — номер плоскости текущего слоя; k — номер плоскости предыдущего слоя; l — номер текущего слоя; n — позиция центра рецептивного поля ячейки; v — индекс связи; $a_{Sl}(v, \kappa, k) (\geq 0)$ — сила переменной возбуждающей связи, идущей от С-ячейки $u_{Cl-1}(v, \kappa)$ предыдущей стадии; A_{Sl} — радиус рецептивного поля ячейки, то есть, размер пространственного сдвига входящего образа относительно центра рецептивного поля; $b_l(k) (\geq 0)$ — сила переменной тормозящей связи, идущей от V-ячейки; $c_{Sl}(v)$ — сила постоянной возбуждающей связи с V-ячейкой; $\theta_l (< 1)$ — порог чувствительности S-ячейки.

Следует отметить, что все ячейки плоскости совместно используют то же самое множество входных связей.

В (2) и (3) для $l = 1$, $u_{Cl-1}(n; k)$ заменяется на $u_G(n_G, k_G)$.

Ответ S-ячейки уровня U_{Sl} пространственно размыт в последующих С-ячейках уровня U_{Cl} . Математически ответом С-ячейки U_{Cl} , исключая самую высокую стадию U_{C4} , является

$$U_{Cl}(n, k) = \psi \left[\sum_{|v| < A_{Cl}} a_{Cl}(v) u_{Sl-1}(v, \kappa) \right] \quad 4$$

где $\psi [x] = \frac{\varphi [x]}{1 + \varphi [x]}$; $a_{Cl}(v)$ — сила постоянных возбуждающих связей, сходящихся от группы S-ячеек, которые распространяются в пределах радиуса A_{Cl} .

По результатам моделирования и серии экспериментов получились следующие параметры слоев $U_{S1} - U_{C4}$:

1. Размер плоскости слоя U_{S1} : 62×62 ячейки, количество плоскостей 47915, $A_{S1} = 4$, $\theta_1 = 0,55$.

2. Размер плоскости слоя U_{C1} : 34×34 ячейки, количество плоскостей 47915, $A_{C1}=9$.

3. Размер плоскости слоя U_{S2} : 34×34 ячейки, количество плоскостей 95850, $A_{S2}=4$, $\theta_2^R = 0,51$, $\theta_2^L = 0,66$.

4. Размер плоскости слоя U_{C2} : 14×14 ячеек, количество плоскостей 95850, $A_{C2}=7$.

5. Размер плоскости слоя U_{S3} : 14×14 ячеек, количество плоскостей 143725, $A_{S3}=4$, $\theta_3^R = 0,58$, $\theta_3^L = 0,67$.

6. Размер плоскости слоя U_{C3} : 10×10 ячеек, количество плоскостей 143725, $A_{C3}=5$.

7. Размер плоскости слоя U_{S4} : 62×62 ячейки, количество плоскостей 142785, $A_{S4}=5$, $\theta_4 = 0,30$.

8. Размер плоскости слоя U_{C4} : 1×1 ячейку, количество плоскостей 6845, $A_{C4}=3$.

Обучение сети выполняется от более низких стадий к более высоким. Хотя метод выбора начальных параметров ячеек во время обучения немного отличается в разных слоях, правило усиления изменяющихся связей $a_l(v, k)$ и $b_l(k)$ одинаково для всех слоев.

Сначала выбирается простой узел, веса синапсов которого должны быть увеличены, он рассматривается как представитель всех узлов в плоскости, вызывая увеличение их связей на том же самом образе. Это делается каждый раз, когда представляется новый обучающий образ и для каждой плоскости. В слое U_{S1} , обучаемом с помощью контролируемого обучения, ячейки-представители выбираются «учителем». Ячейка, центр рецептивного поля которой совпадает с местоположением признака, становится представителем. В слоях U_{S2} и U_{S3} представители выбираются из тех S-ячеек, которые дают максимальный выход. Слой U_{S4} обучается таким же образом, как и слои U_{S2} и U_{S3} , за исключением того, что здесь также запоминается имя соответствующего класса.

Для представителя S-ячейки усиливаются только входные соединения, через которые приходят не равные нулю сигналы. Все другие S-ячейки в S-плоскости, от которых выбран представитель, имеют такое же усиление своих входных соединений, как и их представитель.

Пусть ячейка $u_{Sl}(\hat{n}, \hat{k})$ была выбрана в качестве начальной, следовательно, изменяющиеся связи $a_l(v, k, \hat{k})$ с этой ячейкой, как и со всеми S-ячейками в этой плоскости, увеличиваются следующим образом:

$$\Delta a_l(v, k, \hat{k}) = q_l c_{Sl}(v) u_{Cl-1}(v, k) \quad (6)$$

где q_l — положительная константа ($=100$), определяющая скорость обучения.

$$b_{Sl}(\hat{k}) = \sqrt{\sum_{\kappa=1}^{K_{Cl-1}} \sum_{|v| < A_{Sl}} \frac{\{a_{Sl}(v, \kappa, \hat{k})\}^2}{c_{Sl}(v)}} \quad (7)$$

При использовании сети «Неокогнитрон» возникают две проблемы: большой объем используемой оперативной памяти и огромный объем вычислений. В качестве мер по оптимизации использования памяти были применены следующие: использование динамических массивов и, по возможности, типов памяти с минимальным объемом, хранение и использование только непустых массивов данных, удаление из памяти неиспользуемых данных. С уменьшением данных для расчета проблема огромного объема вычислений частично снимается.

Тестовая выборка для проведения экспериментов по качеству распознавания формировалась из символов обучающей выборки с повреждениями образа в каждой из четвертей. Тестирование проводилось путем последовательного предъявления на вход сети каждого символа из тестовой выборки, с последующим сравнением имени класса, выданного сетью, с настоящим именем.

На обучающей выборке из 6845 символов сеть показала качество распознавания в 99%, на тестовой выборке из 34225 символов сеть показала результат в 86%.

Проведенные эксперименты доказывают возможность применения ИНС «Неокогнитрон» для распознавания иероглифов. Дальнейшее увеличение качества распознавания возможно только в случае применения лингвистического анализа, в случае если семантика древнеегипетского языка будет исследована в должном объеме.

СПИСОК ЛИТЕРАТУРЫ

1. Baurman, J., Grimal, N. Manuel de Codage des textes hiéroglyphiques en vue de leur saisie sur ordinateur. Paris, 1988.
2. Кугаевских А.В. // Реализация иероглифического текстового процессора как инструмент автоматизации деятельности египтолога: Тез. докл. Дистанционные образовательные технологии: опыт применения и перспективы развития. Тюмень: ТюмГУ, 2008. С. 186-191.
3. Кугаевских А.В. Свидетельство об официальной регистрации программы для ЭВМ № 2007614348. Иероглифический текстовый процессор Thoth, 12.10.2007.
4. Berg, H. van den, Grimal, N., Hallof, G., Hallof, J. Hieroglyphica. 2nd ed. Paris & Utrecht, 2000.
5. Вархаген К., Дейн Р., Грун Ф. и др. Распознавание образов: состояние и перспективы. Пер. с англ. М.: Радио и связь, 1985.
6. Хайкин С. Нейронные сети: полный курс, 2-е издание. Пер. с англ. М.: ИД Вильямс, 2008.
7. Fukushima, K. Neocognitron for handwritten digit recognition // Neurocomputing. Vol. 51. 2003. P. 161-180.
8. Jonathan, A. Receptive Field Functions for Face Recognition // Cognitive Science Research Papers. Vol. 391. 1995.
9. Добеши И. Десять лекций по вейвлетам. Ижевск: НИЦ «Регулярная и хаотическая динамика», 2001.