

А. Н. Виноградов, М. Д. Недев, А. А. Талалаев,  
В. П. Фраленко, В. М. Хачумов

## **Распознавание графических образов на основе искусственных нейронных сетей**

**Аннотация.** Исследуется эффективность искусственных нейронных сетей (ИНС) на задачах распознавания графических образов. Показано, что существует порядка трех классов изображений, доставляемых средствами технического зрения, которые наилучшим образом обрабатываются на определенных типах ИНС. Рассмотрены тестовые задачи распознавания зашумленных цифр, фотографий лиц и изображений летательных аппаратов, полученных их космоса, с помощью ИНС. Выполнена экспериментальная программная реализация однослойного персептрона ИНС на отечественной суперкомпьютерной платформе «СКИФ».

### **1. Введение**

В последние десятилетия получила развитие новая прикладная область информатики, занимающаяся вопросами распознавания образов на основе ИНС. Разработано достаточно большое количество ИНС как специализированных, так и универсальных, однако фактически отсутствуют практические рекомендации по распознаванию графических образов. Ввиду наличия нескольких методов настройки, различных классов изображений и разнообразия постановок задач решение указанной проблемы является достаточно сложным. Систематическая работа в этом направлении не проводилась.

Решение задач на ИНС требует значительного времени, как на этапе обучения, так и на стадии распознавания, когда требуется обработка большого числа образцов или изображений большого формата. Отметим, что в России проводится большая работа по созданию отечественных суперкомпьютеров класса «СКИФ», но практически отсутствуют прикладные системы распознавания на основе ИНС. В настоящей работе распознавание на основе однослойного персептрона проводится в реальном масштабе времени с использованием модифицированных алгоритмов ИНС, с учетом новейших технологий программирования и аппаратных средств суперкомпьютера кластерного

типа. Выполненные исследования не претендуют на полноту и являются начальным вкладом в решение задачи систематизации свойств ИНС, ориентированных на решение задач распознавания графических образов.

## **2. Типы обрабатываемых изображений и соответствующие им ИНС**

Все множество различных цифровых снимков, полученных средствами технического зрения и отображаемых в памяти компьютера, может быть разбито на классы в соответствии с представлением отдельной точки (пикселя). За основу приняты следующие типы изображений:

1. Снимки, сделанные из космоса в различных частотных спектрах. Каждый такой спектр дает полутоновое изображение (градации серого), а совокупность спектров - цветное изображение. Всего используется порядка шести спектров. Каждая точка изображения имеет координаты и характеризуется шестью признаками, которые можно использовать для создания классов. Одна из задач связана, например, с раскраской карты. Каждая краска – это отдельный класс соответствующий, например, облакам, воде, зданию и т.д. Обучение ведется с помощью учителя, который выделяет область (облако точек) и назначает класс. Распознавание можно выполнить с помощью расстояния Махаланобиса. Другой способ классификации – использование двухслойной ИНС прямого распространения с сигмоидальной активационной функцией и настройкой генетическим методом. Альтернатива – вероятностная нейронная сеть с активационной функцией Гаусса.

2. Растровые полутоновые изображения, полученные с помощью телевизионных, видео – и фотокамер. Задача – обнаружение локальных объектов (самолетов, танков, зданий и т.д.). Создается сканирующее окно (круглое, прямоугольное, др. формы), захватывающее локальный объект, который следует обнаружить. Метод обнаружения – сравнение эталонного кода с кодом исследуемого образца. Данную технологию предлагается реализовать с помощью вероятностных нейронных сетей, персептронов и сетей общего назначения. Требуемая предобработка – яркостная нормализация изображения.

3. Растровые бинарные изображения. Используются бинарные изображения, (контуры домов и самолетов, бинарное изображение лиц, местности и т.д.), полученные в результате первичной обработки

снимков. Для получения изображений такого типа требуется предварительная обработка – фильтрация, пороговая (яркостная) обработка, выделение контуров и др. Обработка направлена на получение описания приемлемого для работы ИНС Хемминга и Хопфилда.

### 3. Тестирование ИНС на задачах распознавания графических образов

**3.1. Однослойный персептрон.** Будем работать с полутоновыми изображениями (класс ??), имеющими 256 градаций серого цвета. Изображения разворачиваются и подаются на вход ИНС в виде вектора. Целью настройки является обучение ИНС распознаванию множества эталонных векторов. Для этого необходимо скорректировать весовые коэффициенты так, чтобы получить при заданном входе заданный выход.

Персептрон представляет собой однослойную нейронную сеть, в которой количество нейронов равно числу классов. На вход нейронной сети подаётся слово  $X(x_1, \dots, x_n)$ , описывающее образ для обработки. На выходе получаем слово  $Y(y_1, \dots, y_m)$  в соответствии с необходимым преобразованием  $X \rightarrow Y$ . Номер  $j$ , для которого выход  $y_j$  максимален, соответствует номеру класса входного образца. Значение выхода  $y_j$  лежат в диапазоне от 0 до +1, что позволяет оценивать уверенность сети в выносимом решении. Активационная функция имеет вид линейного скачка:

$$f(s) = \begin{cases} 0 & , \text{ если } s \leq -1 \\ E(s) & , \text{ если } -1 < s < +1, \text{ где } E(s) = (1 + s)/2 \\ 1 & , \text{ если } s \geq +1 \end{cases}$$

Метод обучения Видроу–Хоффа известен еще как дельта–правило (delta-rule). Этот метод ставит своей целью минимизацию функции ошибки  $E(k)$  в пространстве весовых коэффициентов:

$$E = \sum_{k=1}^P E(k) = \frac{1}{2} * \sum_{k=1}^P (d_k + y_k)^2,$$

где  $E$  – суммарная среднеквадратичная ошибка,  $P$  – количество обработанных примеров,  $E(k)$  – ошибка для  $k$ -го примера,  $y_k$  – реальный выход ИНС для  $k$ -го примера,  $d_k$  – желаемый (идеальный) выход ИНС для  $k$ -го примера.

В соответствии с методом Видроу–Хоффа весовые коэффициенты в процессе настройки сети изменяются по правилу

$$\omega_j(t+1) = \omega_j(t) * \eta(d_k + y_k) * x_j^k$$

Алгоритм обучения методом Видроу – Хоффа строится следующим образом:

п.1. задать скорость обучения  $-1 < \eta < +1$ , задать минимальную ошибку сети  $E_{min}$  (весовые коэффициенты ИНС инициализируются малыми случайными значениями);

п.2. подать на вход ИНС очередной эталон, рассчитать выход ИНС, скорректировать веса;

п.3. пока  $E > E_{min}$  переход на п.2., иначе конец.

Существует проблема выбора оптимального значения  $\eta$ , когда скорость настройки не меняется в процессе настройки. В случае выбора слишком большой скорости обучения решение может не сойтись, в случае малого  $\eta$  возможна длительная работа сети. В настоящей работе оптимальная скорость настройки подбирается автоматически.

Для оценки эффективности работы ИНС были использованы различные тестовые изображения. В серии экспериментов были исследовано качество распознавания зашумленных и искаженных цифр, представленных полутоновыми изображениями размерности  $54 \times 75$ . Количество нейронов в выходном слое было выбрано равным 10 по числу распознаваемых цифр. Результаты тестовых испытаний ИНС отражены в табл. ??.

В результате настройки все тестовые изображения были распознаны без ошибок.

**3.2. Нейронная сеть Хемминга.** Данная сеть работает только с изображениями класса ??. ИНС Хемминга состоит из двух слоев. Первый и второй слои имеют по  $m$  нейронов, где  $m$  – число образцов. Нейроны первого слоя имеют по  $n$  синапсов, соединенных с входами сети (образующими фиктивный нулевой слой). Нейроны второго слоя связаны между собой отрицательными обратными синоптическими связями. Единственный синапс с положительной обратной связью для каждого нейрона соединен с его же аксоном. Работа сети состоит в нахождении расстояния Хемминга от тестируемого образа до всех образцов. Сеть должна выбрать образец с минимальным расстоянием Хемминга до неизвестного входного сигнала, в результате чего будет активизирован только один выход сети, соответствующий

этому образцу. Активационная функция  $f(s)$  имеет вид линейного скачка (рис. ??).

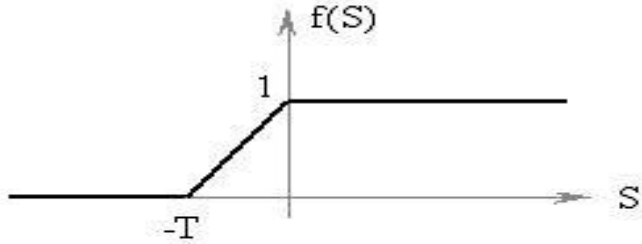


Рис. 1. Функция активации

На стадии инициализации весовым коэффициентам первого слоя и порогу активационной функции присваиваются следующие значения:

$$W_{ik} = \frac{x_i^k}{2} \quad i = 0, \dots, n-1, \quad k = 0, \dots, m-1$$

смещение:

$$T_k = \frac{n}{2} \quad k = 0, \dots, m-1$$

Здесь  $x_i^k$  —  $i$ -ый элемент  $k$ -ого образца. Таким образом, процесс настройки сети практически отсутствует.

Весовые коэффициенты тормозящих синапсов во втором слое берут равными некоторой величине  $0 < \epsilon < 1/m$ . Синапс нейрона, связанный с его же аксоном имеет вес  $+1$ .

Рассмотрим алгоритм функционирования ( $p$  — номер итерации):

1. На входы сети подается неизвестный вектор  $X = \{x_i : i = 0 \dots n-1\}$ , исходя из которого рассчитываются состояния нейронов первого слоя (верхний индекс в скобках указывает номер слоя):

$$y_j^{(1)} = s_j^{(1)} = \sum_{i=0}^{n-1} w_{ij} x_i + T_j, \quad j = 0, \dots, m-1.$$

После этого полученными значениями инициализируются значения аксонов второго слоя:

$$y_j^{(2)} = y_j^{(1)}, \quad j = 0, \dots, m-1.$$

2. Вычислить новые состояния нейронов второго слоя:

$$s_j^{(2)}(p+1) = y_j(p) - \epsilon \sum_{k=0}^{m-1} y_k^{(2)}(p), \quad k \neq j, \quad j = 0, \dots, m-1$$

и значения их аксонов:

$$y_j^{(2)}(p+1) = f \left[ s_j^{(2)}(p+1) \right], \quad j = 0, \dots, m-1$$

3. Проверить, изменились ли выходы нейронов второго слоя за последнюю итерацию. Если да – перейти к шагу 2, иначе – конец.

В соответствии с представлениями Хемминга второй слой ИНС служит только для доведения победившего выхода до значения 1, а всех остальных выходов до 0. В программной реализации ИНС это условие применения второго слоя становится неактуальным. В этой связи в принятой модели второй слой не был использован.

К достоинствам сети Хемминга можно отнести: быстрое обучение, малые затраты памяти (как следствие возможность распознавания больших изображений), легкость масштабирования сети по числу эталонов.

Недостатки ИНС: сеть работает только с бинарными изображениями, сеть предназначена для распознавания статичных изображений, имеет место замедление процесса распознавания из-за наличия обратных связей во втором слое.

В качестве входных данных использовались фотографии размером 200x200 (табл. ?? и ??). Для работы сети Хемминга цветное изображение преобразовывалось в черно-белое, причем специальная обработка не проводилась. В графе «результат распознавания», в скобках, дано значение выхода нейрона первого слоя, что может служить некоторой оценкой качества распознавания.

Как видно из теста, ошибка при распознавании возникает при внесении значительных изменений в исходное изображение (11 и 12 строки таблицы).

**3.3. Сравнительное распознавание графических изображений.** Эксперимент проводился на серии космических снимков размером 140\*185, полученных из космоса, часть из которых представлена в табл. ??.

Как и ожидалось, наилучший результат был получен с помощью однослойного персептрона. Сеть Хемминга давала большее число ошибок на тестовой серии. В целом же обе сети достаточно уверенно распознавали предложенные графические образы. Исключение составили некоторые образы, подвергнутые сильным искажениям.

#### **4. Эксперименты по распараллеливанию ИНС и задач на кластерном устройстве**

Существует следующая, достаточно общая, классификация вариантов параллельной декомпозиции нейросетевых вычислений:

- 1) параллелизм уровня сессии обучения;
- 2) параллелизм уровня учебного примера;
- 3) параллелизм уровня слоёв;
- 4) параллелизм уровня узлов (нейронов);
- 5) параллелизм уровня весов (связей);

Параллелизм уровня сессии обучения подразумевает одновременное обучение одной и той же сети на одинаковых обучающих примерах, однако, с различными параметрами обучения. Параллелизм уровня учебного примера называют также обучением группами (batch learning): множество обучающих примеров разбивается на определенное количество обучающих подмножеств, каждое из которых используется для обучения своего экземпляра нейронной сети данного типа. В процессе обучения накапливаются изменения весов исходной сети для каждого обучающего подмножества. В конце обучения накопленные изменения весов объединяются и вычисляются новые значения весов. Параллелизм уровня слоёв полезен, когда сеть имеет достаточное количество слоёв. Вычисления ведутся конвейерно: каждый слой обрабатывает в данный момент времени свой отдельный пример и, таким образом, увеличивается скорость прохождения данных одного примера через всю сеть, так как сетью обрабатывается несколько примеров одновременно. Параллелизм уровня обрабатываемых узлов (нейронов) означает, что каждый узел является отдельной вычислительной единицей, работающей независимо от других узлов. Параллельно в этом случае выполняется, например, суммирование взвешенных входов каждого нейрона, а также другая необходимая обработка. Параллелизм уровня весов (связей) подразумевает еще более глубокое распараллеливание, обеспечивая одновременные вычисления «проводимости» каждой взвешенной связи в отдельности.

Параллелизм уровня узлов и параллелизм уровня весов применим ко всем моделям нейронных сетей.

Наиболее общими и универсальными в применении являются способы распараллеливания нейросетевых вычислений 4 и 5. Однако, представляется очевидным, что в рамках кластерной архитектуры данные подходы могут оказаться неэффективными в виду слишком малых гранул параллелизма соответствующих моделей вычислений. Под гранулой параллелизма понимается среднее число операций на каждый введенный в гранулу или выведенный из гранулы байт. Наиболее перспективными для реализации на кластерном вычислительном устройстве, по-видимому, являются варианты 1 и 2, а также в некоторой степени (зависит от решаемой задачи) вариант 3.

Следует отметить, что при кластерной организации вычислений, подразумевающей весьма существенный объем суммарной оперативной памяти, распределённой по отдельным узлам, появляется благоприятная возможность для параллельной обработки довольно больших массивов данных, что может быть востребовано при работе с графическими изображениями.

Метод 3 при благоприятных условиях (хорошая реализация, большое количество слёв, и т.д.) может быть применен для ускоренной обработки потока изображений (например, видео-съемки).

Известно, что наиболее интенсивная вычислительная нагрузка приходится на фазу обучения нейронных сетей (например сетей прямого распространения по алгоритму обратного распространения ошибки). Наиболее ресурсоемкими операциями в нейровычислениях являются матричные операции (скалярное умножение, умножение матрицы на вектор и т.п.). Существуют хорошо отработанные методы распараллеливания таких вычислений на распределенных вычислительных системах.

Принципы информационного обмена как нельзя лучше накладываются на структуру распределенных кластеров, где, с одной стороны, происходит быстрая обработка на одном процессоре и, с другой, относительно более медленная передача информации через сеть, связывающую кластер, при этом не требуется синхронизации работы различных участков нейросети.

Для параллельной реализации был использован суперкомпьютер «СКИФ», включающий: управляющую ЭВМ; вычислительные узлы кластерного уровня; системную сеть кластера (SCI), объединяющую вычислительные узлы; вспомогательную сеть (семейства Ethernet, с



поддержкой TCP/IP), объединяющую управляющую ЭВМ и вычислительные узлы. Кластер работает под операционной системой Linux (RedHat 6.2), на которой предварительно установлено программное средство инсталляции Scal и библиотека программ MPI (ScaMPI 2.1), реализующие прикладной интерфейс передачи сообщений высокого уровня по стандарту MPI 1.2.

Были проведены исследования эффективности реализации настройки (обучения) однослойного персептрона на суперкомпьютере (рис. ??).

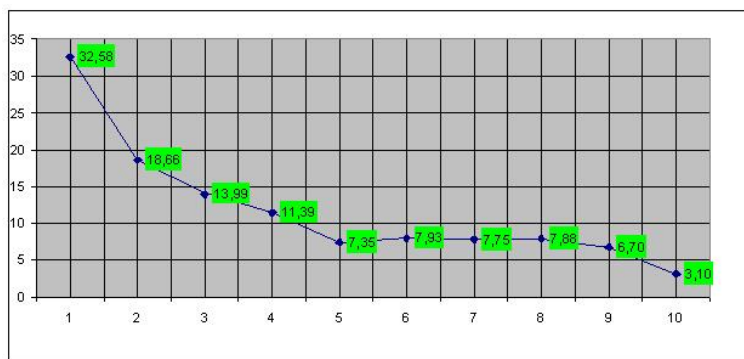


Рис. 2. График зависимости ср. времени настройки сети от числа процессоров

Из графика (рис. ??) видно, что оптимальным выглядит запуск программы на числе узлов, равном 5, дальнейшее увеличение производительности вычислений обеспечивается при числе узлов 10.

## 5. Заключение

В настоящей работе рассмотрены некоторые типы ИНС для распознавания образов. Показано, что сети достаточно уверенно распознают цифры, черно-белые и полутонные снимки. Все исследования проводились над изображениями, не прошедшими первичную обработку, за исключением операции бинаризации. Дальнейший цикл исследований предполагает проведение предварительной обработки изображений для повышения эффективности распознавания каждой из рассмотренных ИНС. Рассмотрены подходы к распараллеливанию

и проведены эксперименты по параллельной реализации однослойного персептрона. Получен положительный эффект по ускорению вычислений - увеличение скорости настройки ИНС с увеличением числа процессоров. Возможные области применения ИНС: обработка космических снимков, задачи мониторинга, диагностики и прогнозирования состояния сложных технических систем.

Настоящая работа выполняется при поддержке РФФИ (проект № 06-07-89083, проект № 05-01-00527).

### Список литературы

- [1] Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. — Москва: Горячая линия - Телеком, 2001. ↑
- [2] Галушкин А.И. Нейрокомпьютеры в системах обработки изображений. Кн.7. — Москва: Радиотехника, 2003. ↑
- [3] Хайкин С. Нейронные сети: полный курс. — Москва: Вильямс, 2006. ↑
- [4] Каллан Р. Основные концепции нейронных сетей. — Москва: Вильямс, 2001. ↑

ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР ИСКУССТВЕННОГО ИНТЕЛЛЕКТА ИПС РАН

A. N. Winogradov, M. D. Nedev, A. A. Talalaev, V. P. Fralenko, V. M. Khachumov. *Recognition of images using neural networks.* (in russian.)

ABSTRACT. The authors discuss a problem of neural networks's efficiency in image recognition. The paper includes description of three classes of images which processed by the neural nets in best way. The authors use as examples noised pictures of numbers, human's faces, and planes. The paper also shows practical parallel realization of a single-layered perceptron. The Russian-Belarusian collaborative project «SKIF» served as a basis for this experiment.







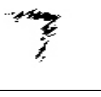


Изображение	Распознано как	Вероятность
	0	76%
	9	59%
	1	55%
	2	48%
	3	91%
	4	52%
	7	58%
	8	52%
	6	56%

ТАБЛИЦА 1. Результаты тестовых испытаний одно-  
слойного персептрона









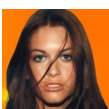
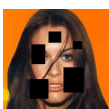






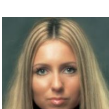
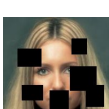


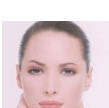
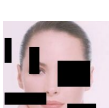


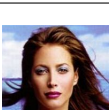



№	Эталон	Тестовое изображение	Эталон (черно-белый)	Тестовое изображение (черно-белое)	Результат распознавания
0					0 (0.90)
1					1 (0.70)
2					2 (0.88)
3					3 (0.88)
4					4 (0.75)
5					5 (0.47)
6					6 (0.79)

Таблица 2. Результаты тестовых испытаний ИНС Хемминга


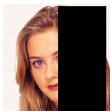






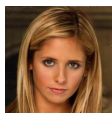







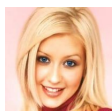
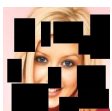



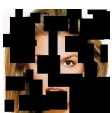


№	Эталон	Тестовое изображение	Эталон (черно-белый)	Тестовое изображение (черно-белое)	Результат распознавания
7					7 (0.45)
8					8 (0.57)
9					9 (0.70)
10					10 (0.45)
11					12 (0.24)
12					3 (0.34)

ТАБЛИЦА 3. Результаты тестовых испытаний ИНС Хемминга (продолжение)



















Изображе- ние цветное	Изобра- жение бинарное	Распознано как	Однослой- ный пер- септрон	Сеть Хеммин- га
			0.64	0.90
			0.84	0.95
			0.57	0.91
			0.88	0.95
			0.66	0.93
			0.48	0.84 Ошибка!

Таблица 4. Сравнительные результаты распознавания графических изображений