

К. Л. Т а с с о в

ПРИМЕНЕНИЕ ИСКУССТВЕННОЙ НЕЙРОННОЙ СЕТИ НЕОКОГНИТРОН ДЛЯ РАСПОЗНАВАНИЯ ГОСУДАРСТВЕННЫХ РЕГИСТРАЦИОННЫХ ЗНАКОВ

Представлена реализация искусственной нейронной сети неокогнитрон и проведено сравнение с другими нейронными сетями. Исследования показали высокую устойчивость сети к шумам, сдвигу и повороту в сопоставлении с другими нейронными сетями.

E-mail: tassov@mail.ru

Ключевые слова: нейронная сеть, неокогнитрон, распознавание, персептрон, сеть встречного распространения.

В существующих системах идентификации транспортных средств (СИТС) используются искусственные нейронные сети (ИНС) типа персептрон для распознавания символов государственных регистрационных знаков (ГРЗ) [1]. Изображение подается на вход ИНС персептрон в бинаризованном виде. В идеальных условиях видеофиксации (хорошее освещение, чистый, без повреждений и искажений ГРЗ) при бинаризации персептроном удается распознавать примерно 96 % ГРЗ. Однако в плохих условиях ввиду бинаризации качество распознавания падает до 30 %. В целях повышения качества распознавания предлагается использовать одну из модификаций ИНС — неокогнитрон [2], для которого не требуется бинаризовать изображение.

В процессе работы создана программа для исследования различных ИНС, с помощью которой проведена оценка производительности сетей. В целях апробации методов распознавания ГРЗ на основе неокогнитрона разработан и реализован на языке C++ программно-аппаратный комплекс, показавший принципиальную работоспособность метода.

Программа состоит из четырех модулей: пользовательский интерфейс; система распознавания (RecognizeSystem); процессор кадров (FrameProcessor); распознаватель номеров (NumberRecognizer). Модули функционируют в отдельных программных потоках и взаимодействуют посредством очередей: numberQueue содержит выделенные номера и recognizedQueue — распознанные номера.

В процессор кадров входят класс чтения видеопотока InputAVI и класс детектирования номеров NumberDetector. Методы stop(), pause(), play() класса FrameProcessor позволяют управлять воспроизведением.

Детектор номеров построен на алгоритме поиска белой точки (WhitePointFinder), алгоритме заливки (Filler) и анализаторе полученных регионов (RegionAnalyzer). Анализатор регионов, используя

серию проверок, отбрасывает области, которые не содержат номерной знак.

Распознаватель номеров включает нейрораспознаватель (NeuroRecognizer), в котором реализован интерфейс Recognizer.

Класс автомобильный номер (CarNumber), используя объект Recognizer, может распознавать составные части номера: буквы (recognizeLetter) и цифры (recognizeNumber) области номера и цифры области региона (recognizeRegionNumber).

Распознаватель выполняется в отдельном потоке. Он извлекает нераспознанные номера из очереди (если очередь пуста, поток останавливается до появления первого номера), дает команду на распознавание с помощью нейрораспознавателя и помещает их в выходную очередь.

Структурно нейрораспознаватель состоит из сети типа неокогнитрон, которая обучена на распознавание определенных групп символов: цифр и букв области номера и цифр области региона. При инициализации распознавателя параметры сети типа неокогнитрон загружаются из заранее созданных файлов. Каждый такой файл содержит топологию сети (число плоскостей, нейронов и связей), параметры сети и весовые коэффициенты.

Неокогнитрон — многослойная ИНС. Изображение с транспортным средством подается на входной слой (слой фоторецепторов) U_0 . За ним следует слой клеток U_G , выделяющих контрасты. Он состоит из двух плоскостей: одна служит для выделения внешней границы, другая — для выделения внутренней границы.

С выхода слоя U_G сигнал подается на вход S -клеток первой стадии (слой U_{S1}). Эти клетки обучены с учителем на выделение ребер определенной ориентации из входного изображения. S -клетки промежуточных стадий (слои U_{S2} и U_{S3}) обладают свойством самоорганизации (конкурентное обучение без учителя). S -клетки высшей стадии (слой U_{S4}) обучаются по алгоритму конкурентного обучения с учителем. Стадии соединяются между собой и образуют единую систему.

Неокогнитрон состоит из пяти стадий [3–5]: входной, контрастной обработки, выделения ребер, промежуточной и высшей.

Входная стадия (рис. 1) предназначена для извлечения графической информации из различных источников и передачи ее к следующим стадиям неокогнитрона.

Для общности архитектуры создан класс InputPlane, наследуемый от класса Plane. Для получения входной графической информации InputPlane использует объекты класса InputHelper. В качестве источников могут служить DIB-секция (подается на вход посредством класса DIBInput) или растровое изображение (используется класс InputImage).

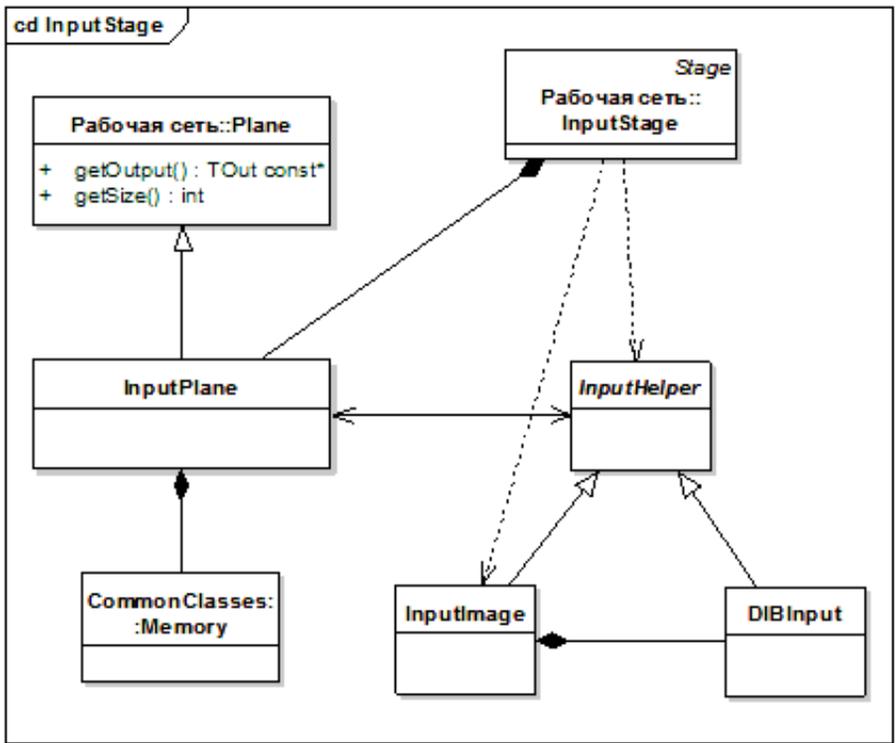


Рис. 1. Входная стадия

На стадии контрастной обработки (рис. 2) на изображении выделяются внешние и внутренние контуры, и эта информация передается к следующей стадии.

Данная стадия состоит из двух плоскостей — ведущей (MContrastPlane) и ведомой (SContrastPlane). Ведущая плоскость вычисляет выходы своих нейронов и конфигурирует выходы нейронов ведомой плоскости. Также в состав стадии входит класс генератора весов (ContrastWeightsGenerator), который создает массивы весовых коэффициентов в момент построения сети.

Стадия выделения ребер предназначена для выделения ребер различной ориентации из входного изображения. Она состоит из S -слоя с заранее заданным числом плоскостей и статического C -слоя (рис. 3). S -слой состоит из массива S -плоскостей, одной V -плоскости и общей для всех плоскостей области связей. S - и V -плоскости содержат весовые коэффициенты и ссылаются на область памяти, в которую будут записаны результаты вычислений. C -слой содержит массив C -плоскостей, массив весовых коэффициентов (общий для всех плоскостей) и область связей.

Промежуточная стадия является самоорганизующейся системой (обучается без учителя, сама определяет необходимое число плоскостей). В сети имеются две такие стадии. Каждая стадия предназначена

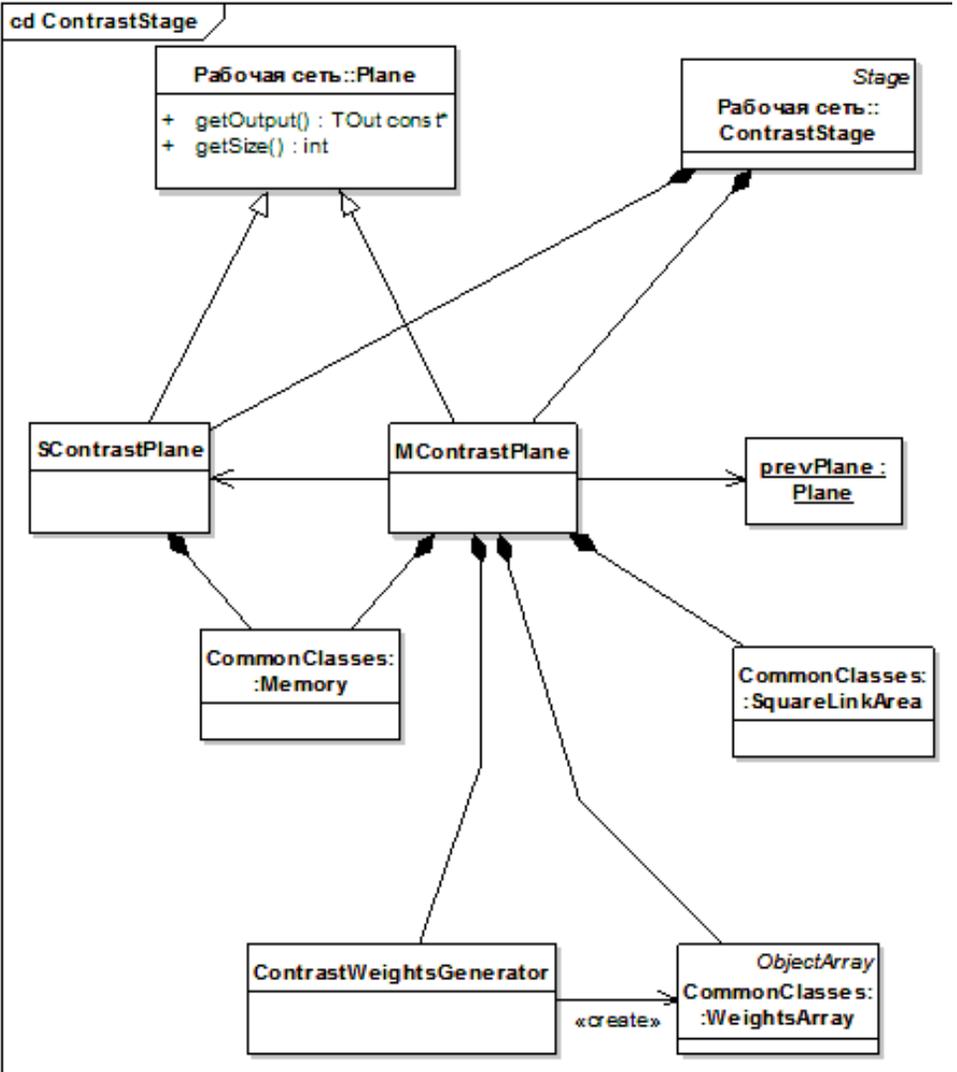


Рис. 2. Стадия контрастной обработки

для выделения элементарных образов из входного изображения (множество образов определяется автоматически во время обучения), они состоят из динамического (число плоскостей переменное) S -слоя и статического C -слоя (рис. 4). C -слой создается после завершения обучения, когда уже известно число плоскостей в S -слое.

Динамический S -слой состоит из массива S -плоскостей и одной V -плоскости. S -плоскости промежуточных слоев проходят процедуру конкурентного обучения без учителя согласно процедуре «победитель забирает все».

Высшая стадия используется для окончательного распознавания входного изображения. Стадия состоит из H -слоя (рис. 5), который в свою очередь содержит массив H -плоскостей и область связей. H -

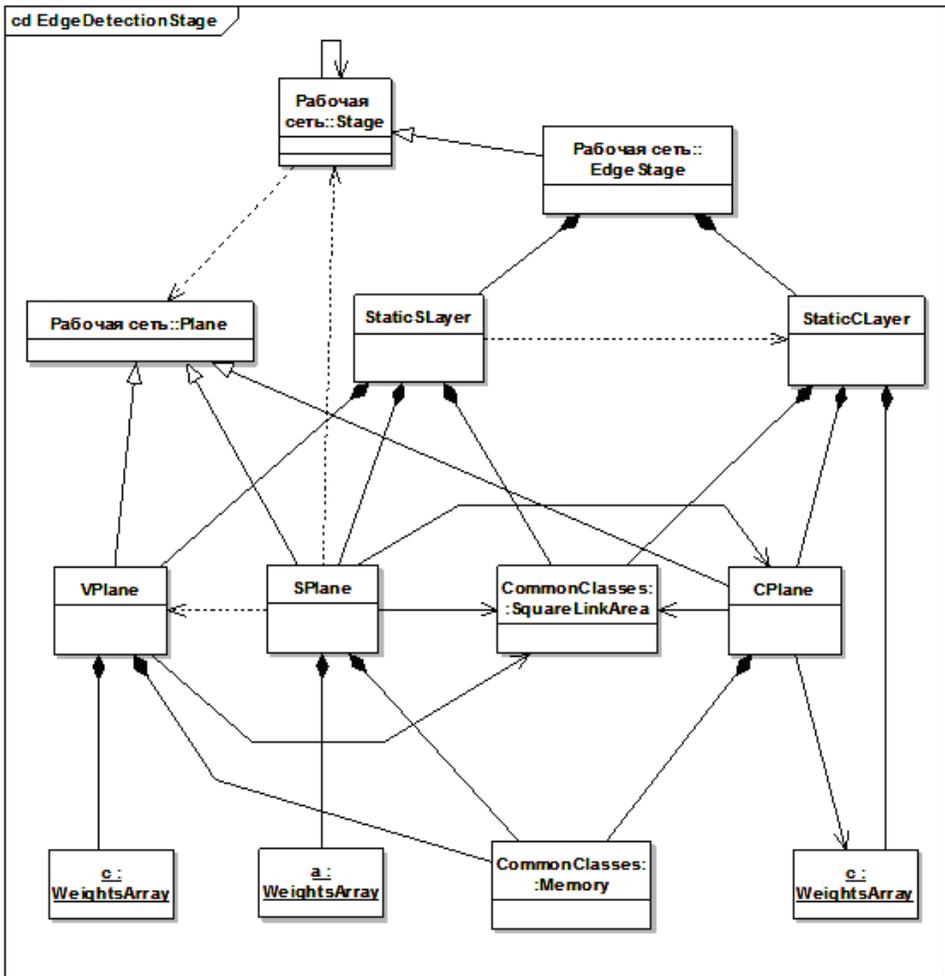


Рис. 3. Стадия выделения ребер

плоскость наследуется от S -плоскости, и к ней добавляется атрибут распознаваемого образа.

Для оценки применимости различных ИНС распознавания ГРЗ разработана программа, структура которой приведена на рис. 6. Данная программа состоит из следующих частей:

- библиотека классов реализованных нейронных сетей;
- модуль загрузки исходных данных, считывающий с диска изображения с символами и формирующий входные векторы;
- модуль классификации отдельного вектора;
- модуль проведения исследований, позволяющий классифицировать набор векторов с возможностью наложения шума, сдвига, поворота, масштабирования и с автоматическим подсчетом результатов классификации;
- пользовательский интерфейс.

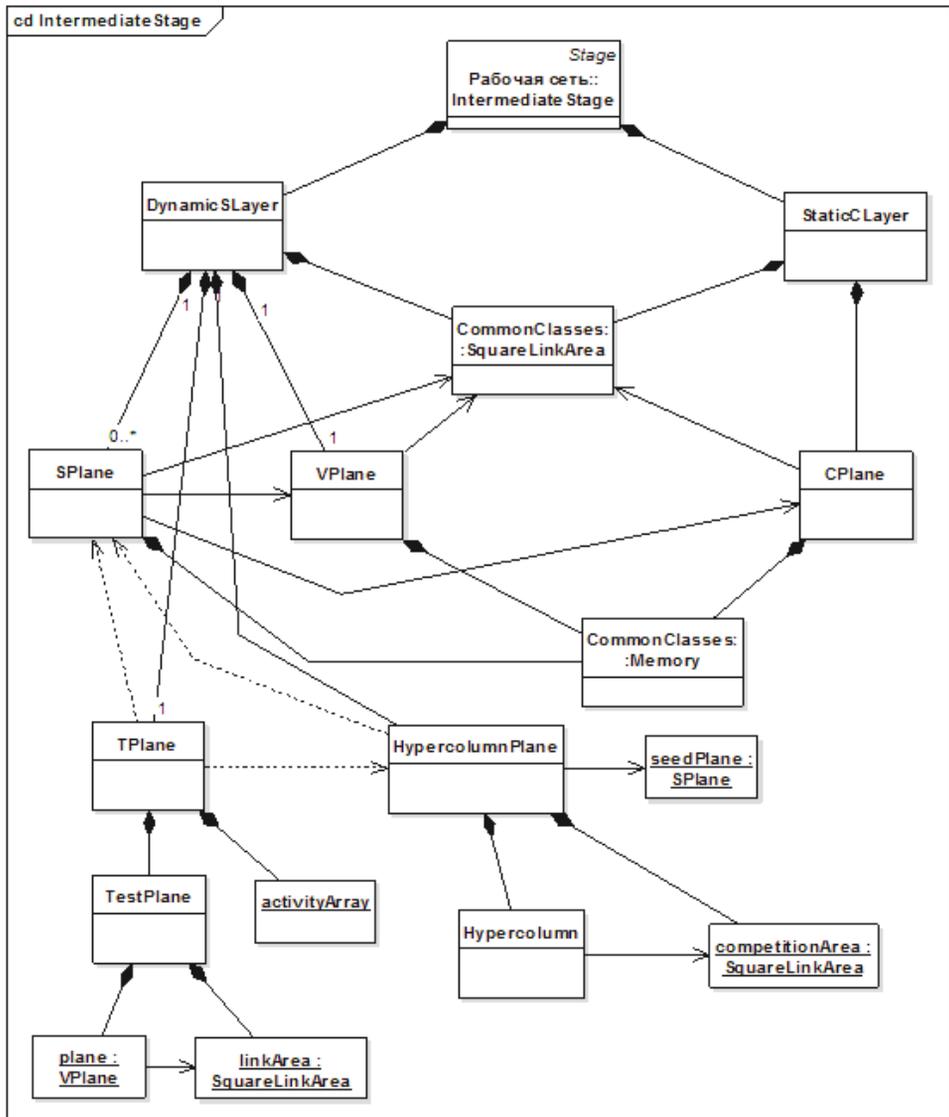


Рис. 4. Промежуточная стадия

Исходными данными для исследования является видеофайл продолжительностью 15 мин (запись проводилась 27.05.2009 г. на Ленинском проспекте с 15:37 до 15:52 при выезде из Москвы). Всего выделено и прошло предварительную обработку 692 номера, символов в них 6228. Для определения возможности использования неокогнитрона в рамках данной работы создана программная модель нейронных сетей.

Для исследования характеристик работы реализованных нейронных сетей были выбраны конфигурации, приведенные в табл. 1. Параметры конфигурации ИНС определяются в соответствии с задачами, стоящими перед сетью.

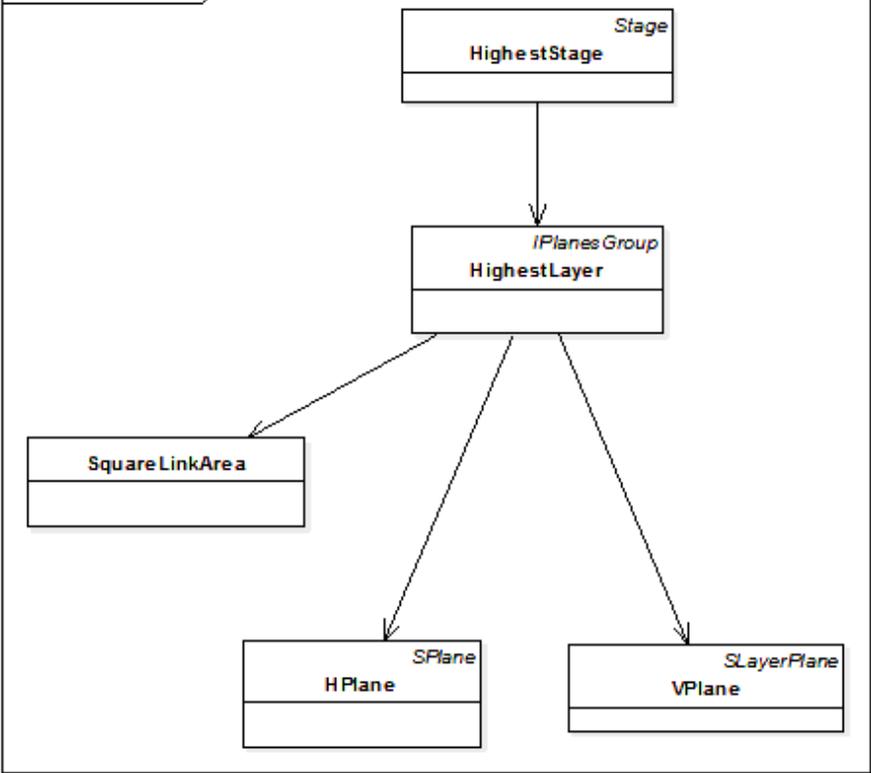


Рис. 5. Высшая стадия

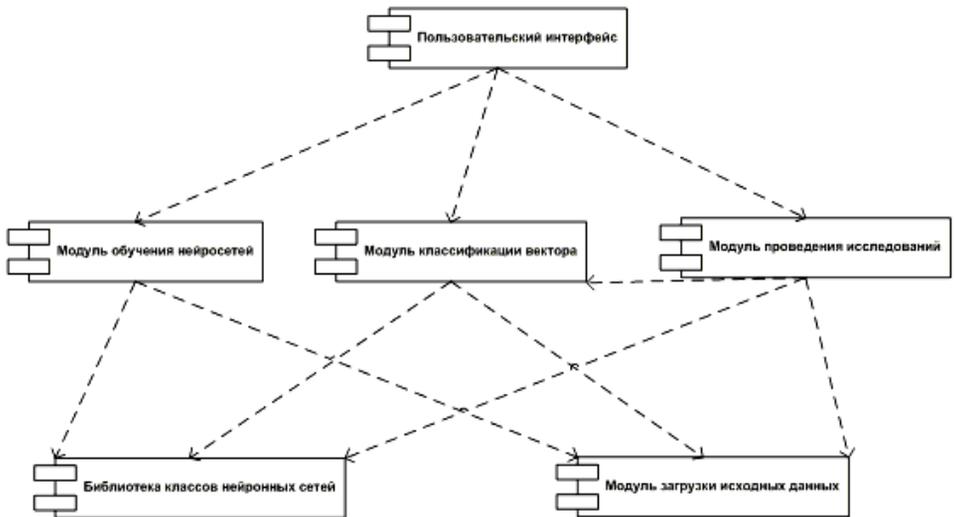


Рис. 6. Схема разработанной программы

Конфигурация исследуемых нейронных сетей

Сеть	Число слоев	Число нейронов в слое	Число весов нейронов
Сеть встречного распространения [4]	2*	По 20**	150*** в слое Кохонена, 20 в слое Гроссберга
Однослойный персептрон, активационная функция — сигмоида [1]	1	20	150
Однослойный персептрон, активационная функция — гиперболический тангенс	1	20	150
Многослойный персептрон, активационная функция — сигмоида	3	По 150 в скрытых, 20 в выходном	150
Многослойный персептрон, активационная функция — гиперболический тангенс	3	По 150 в скрытых, 20 в выходном	150
Вероятностная сеть	2	По 20	150 в первом слое, переменное количество в выходном
Неокогнитрон	9	Общее число 45786	—

Примечание. * Число слоев определяется топологией нейронной сети.

** Каждый ГРЗ содержит 10 цифр и буквы, схожие по начертанию в латинском алфавите и кириллице (А, С, Е, Н, К, М, Р, Т, Х, У). Таким образом, общее число классов, к которым может быть отнесен тот или иной вектор — 20. Исходя из этого определяется по 20 нейронов в выходном слое. Данное число соответствует числу символов, используемых в ГРЗ согласно ГОСТ Р 50557–93.

*** Ширина и высота образца 10×15 символов, исходя из чего число весов нейронов во входном слое Кохонена берут 150.

С помощью подаваемых на вход нейросети векторов описываются изображения, уже прошедшие предварительную обработку и бинаризацию. В результате исследования получены данные о количестве ошибок, допускаемых на тестовой выборке нейросетями, которые в порядке убывания точности классификации приведены ниже:

Точность классификации входных векторов на тестовой выборке

Сеть:

неокогнитрон	0,912
встречного распространения	0,840909
многослойный тангенциальный персептрон	0,840909

многослойный сигмоидальный персептрон	0,795455
однослойный тангенциальный персептрон	0,704545
однослойный сигмоидальный персептрон	0,613636
вероятностная	0,590909

Результаты также подтверждают предположение о том, что для процедуры обратного распространения в качестве активационной функции желательнее использовать гиперболический тангенс, а не сигмоиду.

Замеры времени обучения и классификации проводились на компьютере с процессором Intel Solo, тактовой частотой 1,4 ГГц, 512 МБ ОЗУ. Для сравнения результаты затрачиваемого времени приведены в табл. 2.

Таблица 2

Время обучения и классификации в нейросетях

Сеть	Время обучения на выборке из 500 векторов, с	Время классификации одного вектора, с
Встречного распространения	0,078	0,002062
Однослойный сигмоидальный персептрон	9,532	0,001937
Однослойный тангенциальный персептрон	1,719	0,001953
Многослойный сигмоидальный персептрон	6976,812	0,003797
Многослойный тангенциальный персептрон, конфигурация	1744,265	0,003672
Вероятностная сеть	0,015	0,002031
Неокогнитрон	1093,043	0,015700

Устойчивость работы ИНС зависит от внешних факторов воздействия на изображение ГРЗ. Для определения оптимального типа ИНС (согласно поставленной задаче) необходимо выделить основные факторы. Во всех случаях входными данными для ИНС является изображение ГРЗ, следовательно, на качество изображения влияют такие факторы, как зашумление изображения, его масштабирование, сдвиг и поворот изображения.

На основе разработанных алгоритмов реализации различных типов ИНС рассмотрим влияние данных факторов на устойчивость работы ИНС.

Для изучения устойчивости нейросетей к зашумленности входного изображения используются случайные помехи, искусственно внесен-

ные в изображение. Исследование проводили с теми же данными, которые применяли для обучения сетей, поскольку тестовые изображения уже искажены непредсказуемым образом. Кроме того, вывод об устойчивости сети проще сделать, сравнивая полученные данные с выборкой, на которой сеть показала 100 %-ный результат точности распознавания, т.е. на той выборке, на которой проводилось обучение сети.

При исследовании некоторое фиксированное число пикселей бинаризованных изображений заменяют на противоположные значения и подсчитывают ошибочно классифицированные символы. Число зашумленных пикселей изменяется в заданном диапазоне значений с определенным шагом. С учетом стохастичности вносимых изменений процесс целесообразно повторить несколько раз и вычислить среднюю ошибку. Доля ошибочного распознавания принимает наименьшее значение для сети неокогнитрон: при относительном количестве зашумленных пикселей 37 % доля ошибки составляет 5 %, в то время как для однослойного перцептрона она составляет 30 %, для многослойного — 25 %, для вероятностной сети — 40 %.

Устойчивость к масштабированию также исследовали по измененным изображениям обучающей выборки. Масштаб пошагово изменяли от меньшего по сравнению с начальным изображением размера до превосходящего его. Результаты исследования устойчивости к масштабированию показали лишь ограниченную инвариантность как для сети типа неокогнитрон, так и для других алгоритмов. Так, неокогнитрон сохранял безошибочное распознавание в диапазоне значений $0,87-1,16$ коэффициента масштабирования.

Неокогнитрон также показал более высокую устойчивость к повороту и сдвигу. По сравнению с многослойным перцептроном при сдвиге на 2 пиксела ошибка составила 10 % против 60 %; при повороте на 15—12 % против 40 % в сопоставлении с многослойным перцептроном.

Согласно полученным результатам, для решения поставленной задачи наилучшие параметры качества распознавания по устойчивости к шумам, масштабированию, повороту и сдвигу обеспечивает неокогнитрон.

Проведенные исследования точности классификации показали, что неокогнитрон наиболее устойчив к шумам, масштабированию, повороту и сдвигу. Для данной сети не требуется предварительная бинаризация изображения, благодаря чему можно исключить алгоритм бинаризации при распознавании ГРЗ. Таким образом, применение неокогнитрона в задачах распознавания ГРЗ целесообразно.

СПИСОК ЛИТЕРАТУРЫ

1. У о с с е р м е н Ф. Нейрокомпьютерная техника: Теория и практика. – М.: Мир, 1992.
2. F u k u s h i m a К. Neocognitron: A self-organising neural network for mechanism of pattern recognition unaffected by shift in position // Biological Cybernetics. – 1980. – No. 36. – P. 193–202.
3. F u k u s h i m a К. Neocognitron for handwritten digit recognition // Neurocomputing. – 2003. – No. 51. / ed. by R. Eckmiller & C. von der Malsburg.
4. F u k u s h i m a К. Neural network model for completing occluded contours // Neural Networks. – 2010. – No. 23 (4).
5. К о h o n e n Т. Self-organizing maps. – Berlin, Heidelberg: Springer, 2001.

Статья поступила в редакцию 10.05.2012