

Д. Н. Кангин, А. В. Куров

ОПТИМИЗАЦИЯ ПАРАМЕТРОВ НЕЙРОННОЙ СЕТИ НЕОКОГНИТРОН ДЛЯ РАСПОЗНАВАНИЯ СИМВОЛОВ ГОСУДАРСТВЕННЫХ РЕГИСТРАЦИОННЫХ ЗНАКОВ

Исследована нейронная сеть неоконгитрон Фукусимы. Описаны методы определения пороговых значений, измерений слоев, параметров ядра функции Гаусса, используемого для вычисления выходных значений слоев. Предложены методы оптимизации параметров. В результате исследования достигнут коэффициент распознавания тестовой выборки 97%.

E-mail: dkangin@gmail.com, avkur7@mail.ru.

Ключевые слова: нейронные сети, неоконгитрон, оптимизация, машинное обучение, государственные регистрационные знаки транспортных средств.

В настоящее время в различных областях науки и техники актуальны задачи распознавания образов.

Согласно [1] задача распознавания образов заключается в разбиении многомерного (в общем случае) пространства на области, соответствующие заданным категориям, или классам.

Существуют различные методы распознавания образов: искусственные нейронные сети (далее нейронные сети) [1], методы с использованием фурье-дескрипторов [2, 3], топологические дескрипторы [4, 5], байесовы классификаторы, деревья решений и множество других. Однако несмотря на достигнутые в данной области результаты, поиск методов решения задачи распознавания образов продолжается вследствие повышения производительности ЭВМ, возрастающих требований к результатам распознавания, а также появления новых приложений теории распознавания образов.

Частным случаем задачи распознавания образов является задача распознавания изображений. Она может быть решена как общими, так и специализированными методами с учетом требований к анализу и обработке изображений.

Одна из специализированных нейронных сетей — сеть неоконгитрон, применяемая для распознавания изображений. В данной работе исследуется нейронная сеть неоконгитрон на примере распознавания государственных регистрационных знаков (ГРЗ) транспортных средств (ТС). Обоснование выбора данной нейронной сети приводится далее.

Важными характеристиками методов распознавания изображений служат условия соблюдения инвариантности результатов распознавания относительно:

- линейных преобразований системы координат (переноса, поворота);

- перспективного преобразования;

- наличия шума на изображении.

В практических приложениях требования инвариантности могут быть ослаблены. Так, во многих случаях при распознавании изображений необходима инвариантность относительно поворота лишь для заданного диапазона углов отклонения.

Изначально выполнение условия инвариантности к наличию шума на изображение ограничено реализуемо и зависит от характеристик типа шума (случайный, аддитивный, мультипликативный, периодический и т.д.) и его вероятностных характеристик (математического ожидания и дисперсии).

Помимо этого желательно избежать потери данных в результате предварительной обработки информации. Так, для применения многослойного персептрона требуется передача бинаризованных данных, что влечет потерю части данных при бинаризации.

Задача распознавания ГРЗ в свою очередь состоит из нескольких подзадач:

- 1) выделение ТС из входного изображения;

- 2) выделение ГРЗ из изображения, содержащего ТС;

- 3) сегментация и разбиение на символы выделенного изображения ГРЗ;

- 4) распознавание типа и символов ГРЗ.

Подзадачи 1, 2, 4 являются задачами распознавания изображений. Необходимо определить возможность решения поставленных задач с помощью нейронной сети неоконгитрон, а также установить параметры сети, позволяющие успешно решить задачу распознавания.

Под задачей выделения области на изображении (соответствующей ТС и ГРЗ в данной работе) может быть сформулирована для изображений произвольной размерности.

Пусть задано изображение $f : \mathbb{X}^n \rightarrow \mathbb{Y}, \mathbb{X}^n \in \mathbb{R}^n, \mathbb{Y} \in \mathbb{R}$, а также подмножество области определения (область на изображении) $\mathbb{Z}^n \in \mathbb{X}^n$. Необходимо найти функцию $g(x) : \mathbb{X}^n \rightarrow \mathbb{B}$, которая для любого изображения $f(x)$ принимает значение 1, если $x \in \mathbb{Z}^n$, и 0 в противном случае.

Задача распознавания изображений формулируется следующим образом. Пусть дано множество всех изображений $F = f : \mathbb{X}^n \rightarrow \mathbb{Y}, \mathbb{X}^n \in \mathbb{R}^n, \mathbb{Y} \in \mathbb{R}$, а также набор классов $K \in \mathbb{N}$. Кроме того, для некоторого подмножества $G \subset F$ определена функция $g(x) : G \rightarrow K$. Необходимо найти функцию $f(x) : F \rightarrow \mathbb{N}$, совпадающую с функцией $g(x)$ на множестве G и удовлетворяющую ряду дополнительных ограничений.

Методы с использованием нейронных сетей типа неокогнитрон. Сети типа неокогнитрон [6–8] являются продолжением существовавшей ранее сети когнитрон [5]. В отличие от персептрона применение неокогнитрона ограничено задачами анализа и распознавания изображений.

Изначально неокогнитрон был представлен как модель нейронной сети зрительной системы, имеющая иерархическую многослойную структуру согласно классической гипотезе Хубеля и Вайзеля о зрительной системе приматов.

Неокогнитрон состоит из слоев S -элементов (простых элементов) и C -элементов (составных элементов), расположенных поочередно и образующих иерархию. Иными словами, несколько стадий нейронной сети, каждая из которых состоит из S - и C -слоя, соединены в каскад.

В то время как S -элементы с переменными связями реализуют выделение признаков, C -элементы, входные связи которых фиксированы и не подлежат оптимизации, обеспечивают инвариантность относительно позиции шаблона внутри его чувствительной области. Для высшей стадии нейронной сети C -элементы реализуют функцию распознавания, значение которой представляет собой результат распознавания изображения.

Структура неокогнитрона представлена на рис. 1. Входные данные поступают на слой ввода (слой фоторецепторов) U_0 .

Слой выделения контраста U_G состоит из двух плоскостей элементов, одна из которых содержит концентрические чувствительные области элементов, а другая — периферийные (нецентральные) чув-

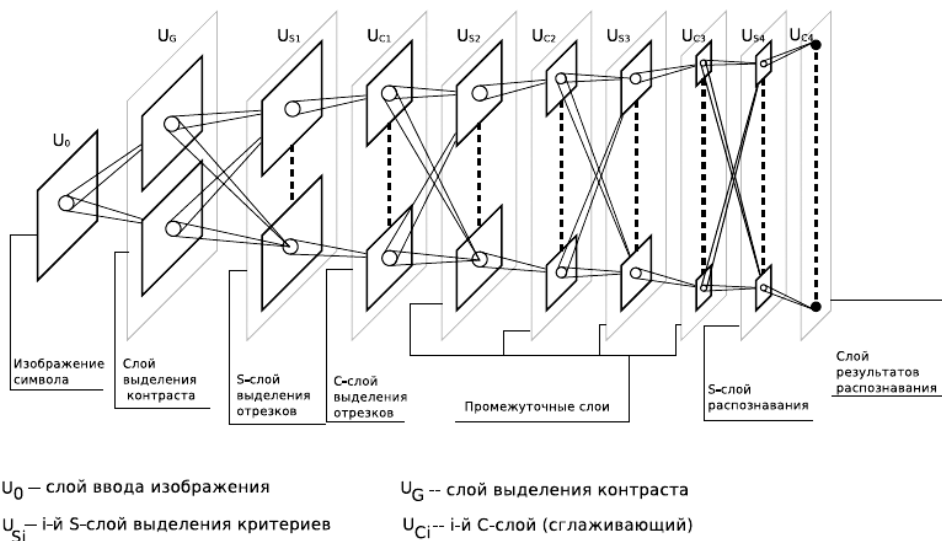


Рис. 1. Структура неокогнитрона

ствительные области элементов. Первая плоскость извлекает разности яркостей исходного изображения, а вторая — инвертированного.

Выход слоя U_G подается на слой S -элементов первого модуля (U_{S1}). S -элементы слоя U_{S1} предназначены для выделения отрезков. Выбор коэффициентов входных связей проводится с помощью алгоритма обучения без учителя.

Данная модель имеет четыре стадии, состоящие из S - и C -элементов. Выходное значение, полученное на слое U_{Sl} , подается на слой U_{Cl} . Входные связи слоев U_{S2} и U_{S3} подлежат оптимизации с использованием алгоритма обучения без учителя, а слоя U_{S4} — с помощью алгоритма конкурентного обучения с учителем.

Слой выделения контраста. Пусть выходные значения слоя U_0 заданы функцией $u_0(n)$, где $n = (n_x, n_y)$ — положение элемента на слое. Выходное значение слоя u_G с центром чувствительной области $n = (n_x, n_y)$ задается выражением

$$u_G(n, k) = \phi[(-1)^k \sum_{|\nu| < A_G} a_G(\nu) u_0(n + \nu)], \quad (1)$$

где $k = 1, 2$; $\phi(x) = \max(x, 0)$; $a_G(\xi)$ — функция весов, задающая зависимость веса от расстояния до центра чувствительной области. Параметр A_G представляет собой радиус пространственного распределения входных связей элемента слоя.

Входные связи к каждому элементу слоя U_G удовлетворяют условию

$$\sum_{|\nu| < A_G} a_G(\nu) = 0, \quad (2)$$

которое означает, что выходное значение слоя U_G не зависит от среднего арифметического входных значений и для любого набора значений, таких, что

$$u_0(n + \nu) = C = \text{const}, \quad |\nu| < A_G, \quad U_G(n, k) = 0.$$

Слой выделения отрезков. Слой U_{S1} , называемый слоем S -элементов первого модуля, является слоем выделения отрезков. Он состоит из 16 плоскостей элементов, каждая из которых содержит элементы для одного предпочтительного угла наклона отрезков. Для каждой из плоскостей элементов предпочтительные углы наклона выбраны с шагом $22,5^\circ$. Для S -элементов устанавливается пороговое значение, определяющее выбор признаков изображения, которое должно быть достаточно низким, чтобы выходные значения были близки при небольших отклонениях угла поворота.

S -элементы данного слоя могут быть обучены с использованием алгоритмов обучения с учителем. Для обучения плоскости элементов

необходимо представить обучающие образы, представляющие собой изображения отрезков с различными углами наклона, на входной слой сети, а также данные об угле наклона отрезка на каждом из изображений. Для элемента, соответствующего положению каждой из точек и лежащего в плоскости, которая соответствует указанному углу, происходит изменение связей с целью увеличения выходного значения (данный процесс называется усилением связей).

Выходные значения S -элементов. Пусть $u_{Sl}(n, k)$ и $u_{Cl}(n, k)$ — выходные значения S - и C -элементов k -й плоскости элементов l -го модуля, где n представляет расположение центра чувствительной области элементов. Слой U_{Sl} содержит не только S -элементы, но и V -элементы, выходные значения которых обозначаются как $v_l(n)$. Связи с C -элементами называются возбуждающими, а с V -элементами — тормозящими. Выходные значения S - и V -элементов задаются выражениями

$$u_{Sl}(n, k) = \frac{\theta_l}{1 - \theta_l} \phi \left[\frac{1 + \sum_{\kappa=1}^{K_{Cl-1}} \sum_{|\nu| < A_{Sl}} a_{Sl}(\nu, \kappa, k) u_{Cl-1}(n + \nu, \kappa)}{1 + \theta_l b_{Sl}(k) v_l(n)} - 1 \right]; \quad (3)$$

$$v_l(n) = \sqrt{\sum_{\kappa=1}^{K_{Cl-1}} \sum_{|\nu| < A_{Sl}} c_{Sl}(\nu) [u_{Cl-1}(n + \nu, \kappa)]^2}, \quad (4)$$

здесь параметр $a_{Sl}(\nu, \kappa, k)$ — весовой коэффициент связи с C -элементом предыдущей стадии, выходное значение которого равно $u_{Cl-1}(n + \nu, \kappa)$. Следует отметить, что все элементы плоскости разделяют общий набор входных связей, так как $a_l(\nu, \kappa, k)$ независима от n . A_{Sl} определяет радиус области суммирования параметра ν , иначе говоря, размер пространственного распределения входных связей к отдельно взятому S -элементу. Параметр $b_l(k) \geq 0$ представляет собой весовой коэффициент переменной тормозящей связи, идущей от V -элемента. Параметр $c_{Sl}(\nu)$ представляет собой весовой коэффициент фиксированных связей к V -элементу и является монотонно убывающей функцией $|\nu|$. Положительная константа θ_l представляет собой пороговое значение S -элемента, определяющее выбор признаков изображений.

Несмотря на то, что шаг элементов плоскости выбран равным для S - и C -элементов, расположения S - и C -элементов согласовывать не обязательно. В некоторых модулях они могут быть смещены относительно друг друга на половину шага, как будет показано далее. В таком случае, если S -элементы расположены в узлах целочисленной сетки n , узлы сетки ν , указывающие расположение C -элементов, соответствуют центрам сетки n .

Входные связи S -элементов. Обучение нейронной сети реализуется в направлении от входного слоя к слою распознавания (от нижних модулей к верхним). Для обучения всех слоев, кроме U_{Sl} , используется один набор данных.

Хотя метод выбора начальных элементов, рассмотренный позднее, значительно отличается от слоя к слою, правило усиления переменных связей $a_l(\nu, \kappa, k)$ и $b_l(k)$ одинаково для всех связей. Усиление связей проводится в зависимости от выходного значения элементов.

Пусть для обучения выбран начальный элемент $u_{Sl}(\hat{n}, \hat{k})$. Тогда веса переменных связей $a_l(\nu, \kappa, \hat{k})$ к начальному элементу и всем элементам плоскости, которой он принадлежит, изменятся на величину

$$\Delta a_{Sl}(\nu, \kappa, \hat{k}) = q_l c_{Sl}(\nu) u_{Cl-1}(\hat{n} + \nu, \kappa), \quad (5)$$

где q_l — положительная константа, определяющая скорость обучения.

Тормозящие связи элемента определяются выражением

$$b_{Sl}(\hat{k}) = \sqrt{\sum_{\kappa=1}^{K_{Cl-1}} \sum_{|\nu| < A_{Sl}} \frac{\{a_{Sl}(\nu, \kappa, \hat{k})\}^2}{c_{Sl}(\nu)}}. \quad (6)$$

Рассмотрим выражения для выходных значений S -элементов.

Пусть x — вектор входных значений S -элемента $X(\nu)$, равных выходным значениям предшествующих C -элементов $u_{Cl-1}(n + \nu, \kappa)$ при $|\nu| < A_{Sl}$.

Также пусть $x^{(i)}$ является i -м обучающим вектором S -элемента. Более точно, $x^{(i)}$ — обучающий вектор для i -го начального элемента плоскости, содержащей данный элемент. Вектор X , называемый опорным вектором, обозначим как сумму всех обучающих векторов для плоскости элементов

$$X = \sum_i x^{(i)}.$$

Для двух векторов X , x и весов $c_{Sl}(\nu)$ можно определить операцию взвешенного скалярного произведения:

$$(X, x) = \sum_{|\nu| < A_{Sl}} c_{Sl}(\nu) X(\nu) x(\nu),$$

где $c_{Sl}(\nu)$ — вес фиксированной возбуждающей связи V -элемента (4), (5). Под нормой вектора здесь и далее понимается евклидова норма $\|x\| = \sqrt{(x, x)}$. Вычитая из (4), (5) выражения (6) и (3) и используя введенные выше обозначения, получаем:

$$u_{Sl}(n, k) = \frac{\alpha}{1 - \theta_l} \phi(s - \theta_l),$$

где

$$s = \frac{(X, x)}{\|X\| \cdot \|x\|}$$

— мера близости между опорным вектором X и заданным вектором x , представляющая косинус угла между двумя векторами. Переменная α , определяемая выражением

$$\alpha = \frac{\theta_l b_{Sl}(k) v_l n}{1 + \theta_l b_{Sl}(k) v_l(n)},$$

принимает значение $\alpha \approx 1$, если $v(n) = \|x\| \neq 0$ по окончании обучения при условии достаточно большого значения $b_{Sl}(k)$.

Если $\alpha \approx 1$, то получим:

$$u_{Sl}(n, k) \approx \frac{\phi(s - \theta_l)}{1 - \theta_l}.$$

Выходное значение S -элемента принимает максимальное значение 1, когда входной вектор равен опорному, и 0, если мера близости не превышает порогового значения элемента θ_l .

Область многомерного пространства признаков, удовлетворяющая условию $s > \theta_l$, является областью толерантности определения признака S -элементом, а порог θ_l определяет диаметр области толерантности.

Если порог принимает достаточно низкое значение, радиус области толерантности становится большим, и S -элементы имеют ненулевое выходное значение даже в том случае, если вектор признаков существенно удален от опорного вектора. Тогда на один признак отвечают несколько S -элементов слоя. Эмпирически показано, что такая ситуация может повлиять на коэффициент верного распознавания неоконченного.

Связи C -элементов. Выходное значение элементов слоя U_{Sl} подвергается размытию в последующем слое U_{Cl} . Выходное значение C -элемента слоя U_{Cl} , за исключением слоя U_{C4} , определено выражением

$$u_{Cl}(n, k) = \psi \left[\sum_{|\nu| < A_{Cl}} a_{Cl}(\nu) u_{Sl}(n + \nu, k) \right],$$

где $\psi[x] = \phi[x]/(1 + \psi[x])$. Параметр $a_{Cl}(\nu)$ представляет веса фиксированных возбуждающих связей, исходящих из группы S -элементов, распространенных внутри радиуса A_{Cl} .

Алгоритм конкурентного обучения. S -элементы промежуточных модулей U_{S2} и U_{S3} подлежат обучению без учителя методом самоорганизации. Начальные элементы выбираются с помощью алгоритма «победитель забирает все». Согласно этому алгоритму, каждый раз, когда

обучающие образы подаются на входной слой, каждый S -элемент соревнуется с остальными элементами окрестности, называемой окрестностью соревнования, которая имеет форму гиперстолбца в наборе плоскостей элементов. Элемент $U_{Sl}(n, k)$ выбирается в качестве начального тогда и только тогда, когда его выходное значение превосходит выходное значение любого из элементов окрестности соревнования. Из уравнения (5) видно, что каждая входная связь начального элемента увеличивается пропорционально выходному значению элемента, из которого ведет данная связь. Поскольку существуют общие связи для каждой из плоскостей элементов, то каждый из элементов плоскости должен иметь то же самое множество входных связей, что и начальный элемент. Способность нейронных сетей к устойчивому распознаванию образов обусловлена избирательностью элементов к признакам, управляемой пороговым значением.

Для S -элементов модуля распознавания (U_{S4}) реализуется конкурентное обучение с учителем. Алгоритм обучения аналогичен применяемому для слоев U_{S2} и U_{S3} , но для обучения также необходимы метки классов образов. Когда нейронная сеть проходит обучение на различных деформированных образах с использованием конкурентного обучения, для одного класса может быть сгенерировано несколько различных плоскостей элементов. Поэтому каждой добавляемой плоскости слоя U_{S4} следует ставить в соответствие метку класса, которому принадлежит образ.

Метка S -элемента с максимальным выходным значением определяет окончательный результат распознавания. Каждый раз, когда в процессе обучения создается новая плоскость элементов, образуются возбуждающие связи ото всех S -элементов к C -элементу данного класса. После обучения, в процессе распознавания, также проходит соревнование между S -элементами, и лишь максимальное выходное значение может быть передано на слой U_{C4} .

В процессе распознавания для слоя U_{S4} пороги θ_4^R выбираются настолько низкими, чтобы для большинства входных значений имелся ненулевой выход от нескольких S -элементов. Следовательно, процесс нахождения S -элемента с наибольшим выходным значением эквивалентен процессу нахождения ближайшего опорного вектора в многомерном пространстве признаков.

Оптимизация параметров неокогнитрона. *Оценка параметров распределения Гаусса.* Фиксированные связи (a_{Cl} , c_{Sl}) вычисляются с помощью функций Гаусса. Начальное значение переменных связей a_{Sl} , b_{Sl} также задают функцией Гаусса.

Параметры ядра функции Гаусса могут быть рассчитаны различными методами. Предлагаемый метод основан на определении корреляции между пикселями векторов входных значений. Рассмотрим

интенсивность пикселей векторов входных значений как случайную величину $\xi(0, 0)$. Также предположим, что соседний пиксель, смещенный на (x, y) относительно данного, является случайной величиной $\xi(x, y)$, коррелирующей с $\xi(0, 0)$. Таким образом, можем вычислить коэффициент корреляции по Пирсону между $\xi(0, 0)$ и $\xi(x, y)$:

$$\rho_{\xi(0,0),\xi(x,y)} = \frac{\sum_{i=1}^n (\xi_i(0,0) - \mu_{\xi(0,0)})(\xi_i(x,y) - \mu_{\xi(x,y)})}{\sqrt{\sum_{i=1}^n (\xi_i(0,0) - \mu_{\xi(0,0)})^2}} \times$$

$$\times \frac{1}{\sqrt{\sum_{i=1}^n (\xi_i(x,y) - \mu_{\xi(x,y)})^2}},$$

где индекс суммирования проходит все пиксели входного изображения. Далее результирующая функция аппроксимируется функцией Гаусса с помощью метода наименьших квадратов. Затем размерность матрицы связей вычисляется согласно правилу трех сигм.

Оптимизация пороговой функции. Выбор порогового значения θ_l является существенным для обучения и распознавания. Для этого максимизируют выходное значение u_{Sl} :

$$u_{Sl}(\vec{n}, \vec{k}) \rightarrow \max.$$

Согласно теореме Ферма, необходимо решить уравнение

$$\frac{\partial u_{Sl}}{\partial t} = 0.$$

Пусть

$$\alpha = 1 + \sum_{\kappa=1}^{K_{C_{l-1}}} \sum_{|\vec{v}| < A_{Sl}} a_{Sl}(\vec{v}, \kappa, k) u_{C_{l-1}}(\vec{n} + \vec{v}, \kappa);$$

$$\beta = b_{Sl}(k) v_l(\vec{n}).$$

Тогда

$$u_{Sl}(\vec{n}, k) = \frac{\theta_l}{1 - \theta_l} \phi \left(\frac{\alpha}{1 + \theta_l \beta} - 1 \right).$$

Если $1 + \theta_l \beta > 0$, то

$$u_{Sl}(\vec{n}, k) = \frac{\theta_l \alpha}{(1 - \theta_l)(1 + \theta_l \beta)} - \frac{\theta_l}{1 - \theta_l}. \quad (7)$$

Дифференцируя выражение (7)

$$\frac{\partial u_{sl}(\vec{n}, k)}{\partial \theta_l} = \frac{\beta(\alpha - \beta)\theta_l^2 - 2\beta\theta_l + \alpha - 1}{(1 - \theta_l)^2(1 + \theta_l\beta)^2},$$

получаем пороговое значение

$$\theta_l = \frac{1 - \sqrt{1 - \frac{(\alpha - 1)(\alpha - \beta)}{\beta}}}{\alpha - \beta}.$$

Выбор размерностей слоев неокогнитрона. Рассматриваемые варианты сети неокогнитрон позволяют проводить независимую настройку размерностей (ширины и высоты) слоев неокогнитрона. Далее приведены несколько эвристических формул для настройки размерностей слоев неокогнитрона, обобщающих параметризацию неокогнитрона Фукусимы.

Размерности входного слоя U_0 выбираются исходя из размерности входных данных. Размерности слоя U_G определяются согласно формулам

$$U_{GX} = U_{0X} + A_{GX} - 1;$$

$$U_{GY} = U_{0Y} + A_{GY} - 1.$$

Такая параметризация измерений U_G позволяет утверждать, что все ненулевые элементы матрицы свертки входной матрицы и матрицы связей будут учтены в выходной матрице.

Размерности слоев U_{C1} , U_{C2} и U_{C3} установлены согласно формулам

$$U_{CX_n} = 2\left(\left\lfloor \frac{\lfloor U_{SX_{n-1}}/2 \rfloor - \lfloor AC_{X_n}/2 \rfloor + 1}{m_x} \right\rfloor + \left\lfloor \frac{AC_{X_n}}{m_x} \right\rfloor\right) + U_{SX_{n-1}} \bmod 2 - 1, \quad (8)$$

$$U_{CY_n} = 2\left(\left\lfloor \frac{\lfloor U_{SY_{n-1}}/2 \rfloor - \lfloor AC_{Y_n}/2 \rfloor + 1}{m_y} \right\rfloor + \left\lfloor \frac{AC_{Y_n}}{m_y} \right\rfloor\right) + U_{SY_{n-1}} \bmod 2 - 1. \quad (9)$$

Здесь m_x и m_y обозначают коэффициенты снижения размерности матрицы выходных значений слоя относительно матрицы входных значений. Формулы (8) и (9) реализуют сокращение размерности. Также они обеспечивают, чтобы четность размерностей слоя изменялась от слоя к слою при смещении на половину пиксела, которое реализуется в неокогнитроне.

Размерности слоев U_{S2} и U_{S3} определены так

$$U_{SX_2} = U_{CX_1} + 1; \quad (10)$$

$$U_{SY_2} = U_{CY_1} + 1. \quad (11)$$

Согласно формулам (10) и (11) осуществляется смещение связей на половину элемента, для чего требуется изменение четности между слоями.

Слой U_{S1} имеет размерности, получаемые согласно формулам:

$$U_{SX_1} = U_{GX} + 1 - \lfloor \frac{A_{SX_1}}{2} \rfloor - (\lfloor \frac{A_{SX_1}}{2} \rfloor \bmod 2); \quad (12)$$

$$U_{SY_1} = U_{GY} + 1 - \lfloor \frac{A_{SY_1}}{2} \rfloor - (\lfloor \frac{A_{SY_1}}{2} \rfloor \bmod 2). \quad (13)$$

В формулах (12), (13) учитывается смещение связей на половину элемента.

Размерности слоя U_{S4} вычисляются согласно формулам

$$U_{SX_4} = U_{CX_3} + A_{SX_4};$$

$$U_{SY_4} = U_{CY_3} + A_{SY_4}.$$

Размерности слоя U_{C4} постоянны:

$$U_{CX_4} = U_{CY_4} = 1.$$

Определение размерности экспериментальной выборки. Размерность экспериментальной выборки установлена согласно центральной предельной теореме (ЦПТ), которая устанавливает, что сумма n независимых событий, имеющих одинаковое статистическое распределение, сходится к нормальному распределению при $n \rightarrow \infty$. ЦПТ можно считать справедливой, тогда

$$\forall u_0^i : E [u_0^i] = \mu, D [u_0^i] = \sigma^2 \Rightarrow E \left[\sum_{i=0}^n u_0^i \right] = \mu, D \left[\sum_{i=0}^n u_0^i \right] = \sigma^2/n.$$

Далее можно определить размерность выборки.

В результате обработки символов тестовой и обучающей выборки получен ряд результатов, показывающий возможность применения методов, описанных выше. Тестовая и обучающая выборка содержат по 5 000 символов каждая. Анализируемые объекты представляют собой изображения различных размеров до 14×17 пикселей, содержащие собой буквы и цифры ГРЗ Российской Федерации.

Доля верно распознанных событий может различаться в зависимости от способа формирования тестовой выборки. На тестовой выборке результат составляет 97 % верно распознанных символов, на обучающей — 100 %. Примеры обучающей и тестовой выборки приведены на рис. 2.

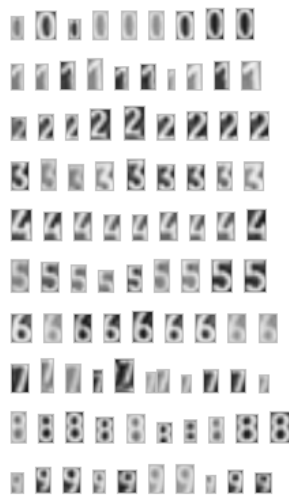


Рис. 2. Примеры обучающей и тестовой выборки цифр ГРЗ

СПИСОК ЛИТЕРАТУРЫ

1. F u k u n a g a K. Introduction to Statistical Pattern Recognition. Second ed. – London: Academic Press, 1990.
2. B r i g g s W. L., Van Emden Henson The DFT: an owner’s manual for the discrete Fourier transform. Society of industrial and applied mathematics. PA, USA. 1995. P. 143–179.
3. K i m u r a F., S h r i d d h a r M. Handwritten Numerical Recognition Based on Multiple Algorithms // Pattern Recognition. – 1991. – Vol. 24, No. 10. – P. 969–983.
4. Persistence Barcodes for Shapes / G. Carlsson, A. Zomorodian, A. Collins, and L. Guibas // Int. J. of Shape Modeling. 2005. 11 P. 149–187.
5. D u t t a A., C h a u d h u r y S. Bengali alpha-numeric character Recognition Using Curvature Features // Pattern Recognition 1993. 26. P. 1707–1720
6. F u k u s h i m a K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position // Biological Cybernetics. 1980.– 36(4). – P. 93–202.
7. F u k u s h i m a K. Neocognitron for handwritten digit recognition // Neurocomputing / R. Eckmiller & C. Von der Malsburg, eds. 2003. 51. 161–180.
8. F u k u s h i m a K. Increasing robustness against background noise: Visual pattern recognition by a neocognitron // Neural networks. – 2011. – Vol. 24 (7). – P. 767–778.
9. F u k u s h i m a K. Cognitron: A self-organizing multilayer neural network model // NHK Technical Monograph 1981. No. 30. – P. 1–25.

Статья поступила в редакцию 10.05.2012