

СПЕЦИАЛИЗИРОВАННОЕ ВЫЧИСЛИТЕЛЬНОЕ УСТРОЙСТВО НА ОСНОВЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ ОБРАБОТКИ ПЛОСКИХ ИЗОБРАЖЕНИЙ

д.т.н., проф. В.Д. Дмитриенко, Р.Д. Расрас, А.М. Сырой

В статье рассматриваются вопросы построения специализированного вычислительного устройства, предназначенного для обработки плоских полутоновых изображений на основе нейронных сетей.

При обработке данных физического эксперимента часто возникают задачи классификации в условиях помех изображений объектов $A_k(x, y)$, $k = \overline{1, m_0}$ при их произвольном расположении в поле зрения $D(x, y)$ и различных масштабных изображений m_{jr} , $j = \overline{1, m}$, $r = \overline{1, R_j}$, где m – число различных отдельных распознаваемых объектов, представленных в разных масштабах (или классов объектов); R_j – число разных масштабов изображений j -го объекта (или класса объектов). Будем рассматривать в поле зрения $D(x, y)$ изображения $S(x, y)$, содержащие фон $A_\Phi(x, y)$ и $m_0 = \sum_{j=1}^m \sum_{r=1}^{R_j} j_r$ объектов. Пусть $D_\Phi \subset D$ – область фона,

$D_k \subset D$, $k = \overline{1, m_0}$ – область расположения k -го объекта. При этом области изображений объектов не пересекаются, т.е.

$$D_i \cap D_n = \emptyset \text{ при } i \neq n \quad (1)$$

и удовлетворяют условию

$$D_\Phi \cup D_1 \cup D_2 \cup \dots \cup D_{m_0} = D. \quad (2)$$

Совокупность изображений фона $S_\Phi(x, y)$ и отдельных объектов $S_k(x, y)$, $k = \overline{1, m_0}$ образуют изображение поля зрения

$$S(x, y) = S_\Phi(x, y) \cup S_1(x, y) \cup \dots \cup S_{m_0}(x, y),$$

$$S_i \cap S_n = \emptyset, \text{ если } i \neq n. \quad (3)$$

Выполнение условий (1 - 3) означает, что:

$$S_\Phi(x, y) = D_\Phi(x, y) \text{ при } (x, y) \in D_\Phi;$$

$$S_k(x, y) = D_k(x, y) \text{ при } (x, y) \in D_k, k = \overline{1, m_0}.$$

При решении задачи классификации на изображении $S(x,y)$ множества из m_0 объектов на заранее неизвестное число m различных классов вначале производится процедура сегментации изображения $S(x,y)$, т.е. разбиение его на изображение фона $S_{\Phi}(x,y)$ и изображения отдельных объектов $S_k(x,y)$, $k = \overline{1, m_0}$. В математическом плане это означает определение предиката вида

$$P(x,y) = \begin{cases} k, & \text{если } (x,y) \in D_k; \\ 0, & \text{если } (x,y) \in D_{\Phi}. \end{cases} \quad (4)$$

В результате определения предиката (4) исходное изображение (3) декомпозируется на m_0 отдельных изображений $S_k(x,y)$, $k = \overline{1, m_0}$, часть из которых отличается друг от друга только масштабом. При различном масштабе произвольно расположенных в поле зрения изображений объектов для определения принадлежности их к определенным классам (или объектам) целесообразно воспользоваться переходом к сокращенному пространству признаков, содержащему периметры P_k и площади $S_k^{пл}$ объектов, поскольку известно [1], что соотношение

$$\rho = \frac{P}{\sqrt{S^{пл}}}, \quad (5)$$

где P – периметр; $S^{пл}$ – площадь многоугольника, инвариантно к изменению масштаба любого многоугольника. Поскольку в работе [1] указанное утверждение приведено без доказательства, то покажем его справедливость для произвольных прямоугольников. Действительно, для прямоугольника со сторонами a и b из соотношения (5) имеем

$$\rho_{пр} = \frac{2(a+b)}{\sqrt{ab}}. \quad (6)$$

При изменении масштаба изображения в k ($0 < k < \infty$) раз стороны прямоугольника также изменятся в k раз: $a_1 = ka$, $b_1 = kb$.

С учетом этого имеем

$$\rho_1 = \frac{2(a_1 + b_1)}{\sqrt{a_1 b_1}} = \frac{2(ka + kb)}{\sqrt{k^2 ab}} = \rho_{пр}.$$

Таким образом, соотношение (5) действительно инвариантно к изменению масштаба прямоугольников. Теперь покажем, что инвариантность выражения (5) к изменению масштаба выполняется для любых плоских фигур, ограниченных замкнутыми кривыми одинаковой формы.

Любую из таких фигур (рис. 1) можно представить как сумму прямо-

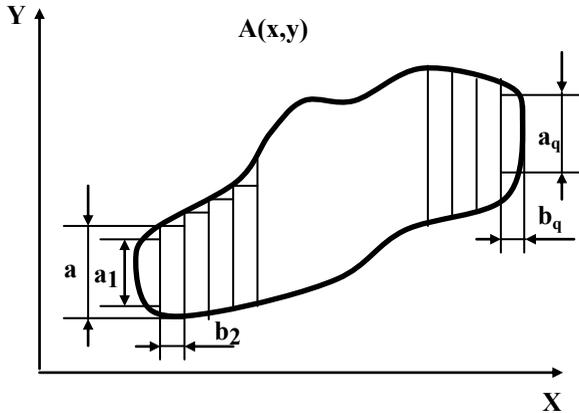


Рис.1. Пример плоской фигуры, ограниченной замкнутыми кривыми одинаковой формы

угольников, площадь каждого из которых стремится к нулю, а суммарная площадь всех прямоугольников неограниченно близко приближается к площади рассматриваемой фигуры $A(x,y)$. Для фигуры A соотношение (5) можно записать следующим образом:

$$\rho_A = \frac{P_A}{\sqrt{S_A^{н.л}}} = \frac{a_1 + a_q + 2 \sum_{i=1}^q b_i + \sum_{i=1}^{q-1} \Delta_i}{\sqrt{\sum_{i=1}^q a_i b_i}},$$

где a_1, a_q - длины сторон первого и последнего прямоугольников, параллельные оси ординат; $b_i (i = \overline{1, q})$ - стороны прямоугольников, параллельные оси абсцисс; $\Delta_i = (a_{i+1} - a_i)$ - разность длин сторон соседних прямоугольников.

При изменении масштаба изображения фигуры $A(x,y)$ в k раз имеем:

$$\rho_k = \frac{P_k}{\sqrt{S_k^{н.л}}} = \frac{ka_1 + ka_q + 2 \sum_{i=1}^q kb_i + \sum_{i=1}^{q-1} k\Delta_i}{\sqrt{\sum_{i=1}^q k^2 a_i b_i}};$$

$$P_k = \frac{a_1 + a_q + 2 \sum_{i=1}^q b_i + \sum_{i=1}^{q-1} \Delta_i}{\sqrt{\sum_{i=1}^q a_i b_i}} = \rho,$$

где P_k и $S_k^{пл}$ - соответственно периметр и площадь фигуры, масштаб изображения которой изменился в k раз.

Таким образом, соотношение (5) является инвариантом для любых плоских фигур, ограниченных замкнутой кривой и отличающихся масштабом.

Для подсчета площади k -го объекта, находящегося в поле зрения $D(x,y)$, достаточно с помощью предиката (4) подсчитать число N_k пикселей этого изображения и умножить его на площадь S_p одного пикселя

$$S_k^{пл} = N_k \cdot S_p.$$

Периметр любого k -го изображения складывается из длин сторон отдельных пикселей, которые будем считать квадратными с длиной стороны d . Любой пиксель изображения может быть расположен одним из пяти способов (рис. 2):

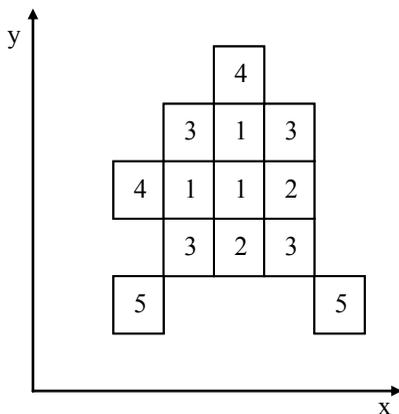


Рис. 2. Расположение пикселей изображения

– пиксели, обозначенные на рис.2 цифрой 1, имеют в изображении объекта 4 соседние (сверху, снизу, слева и справа); вклад этих пикселей в длину периметра – нулевой;

– пиксели, обозначенные на рис.2 цифрой 2, имеют по три соседних; их вклад в длину периметра – длина d одной стороны пикселя;

– пиксели, обозначенные цифрой 3, имеют по два соседних; их вклад в длину периметра составляет $2d$;

– пиксели, обозначенные цифрой 4, имеют только один соседний пиксель и увеличивают длину периметра на $3d$;

– пиксели, обозначенные цифрой 5, не имеют соседних и увеличивают длину периметра на $4d$.

Из анализа рис.2 и приведенных рассуждений следует, что длина периметра любого изображения определяется соотношением

$$P = N_1 d + 2N_2 d + 3N_3 d + 4N_0 d, \quad (7)$$

где $N_r, r = \overline{1,3}$ – число возбужденных нейронов, имеющих в изображении r соседних возбужденных элементов; N_0 – число нейронов, не имеющих в пространстве соседних элементов.

Классификацию изображений с помощью соотношения вида (5) рационально выполнять с помощью технологии нейронных сетей [2]. Одна из возможных архитектур специализированного вычислительного устройства (СВУ) на основе нейронных сетей приведена на рис.3. Устройство имеет поле зрения $D(x,y)$ из $n_1 \times n_2$ чувствительных бинарных x -нейронов, где n_1 – число строк поля и n_2 – число столбцов. Каждый x -нейрон соответствует одному пикселю поля зрения $D(x,y)$.

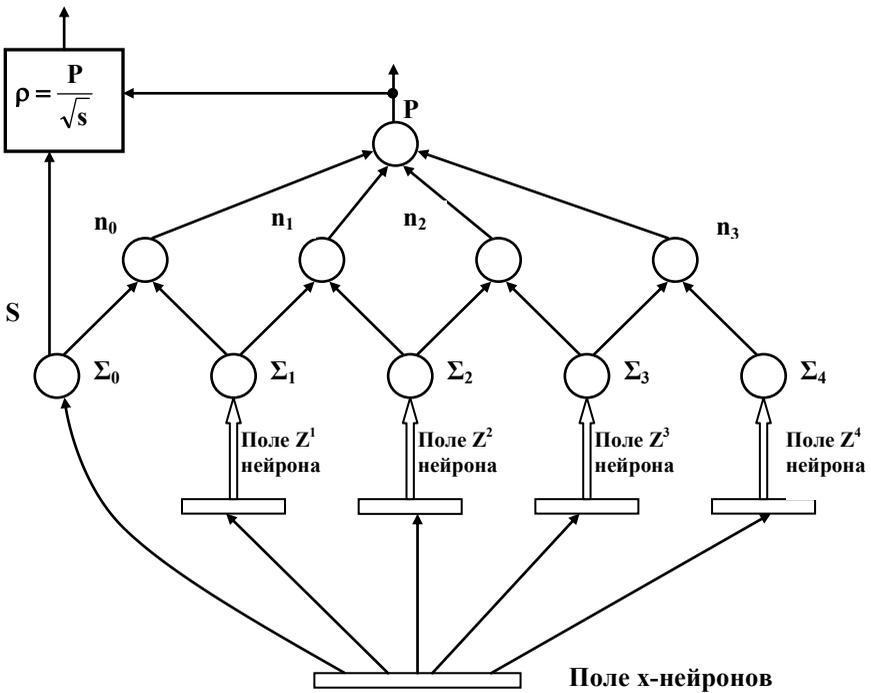


Рис. 3. Пример СВУ на основе нейронных сетей

Функционирование чувствительных x -нейронов описывается выражением

$$U_{\text{вых}} = \begin{cases} 1, & \text{если } U_{\text{вх}} \geq \theta_n; \\ 0, & \text{если } U_{\text{вх}} \leq \theta_n. \end{cases}$$

где $U_{\text{вых}}$, $U_{\text{вх}}$ – соответственно выходной и входной сигнал x - элемента; θ_n - порог чувствительности x - элемента.

Бинарные сигналы с выходов x -элементов поступают на входы нейронов четырех полей $Z^1 - Z^4$, обрабатывающих информацию о входном изображении. На входы элемента Z_{ij}^k ($k = \overline{1,4}$, $i = \overline{1, n_1}$, $j = \overline{1, n_2}$) k - го поля Z^k поступают сигналы от нейрона x_{ij} и его ближайших соседей, т.е. нейронов

$$x_{i-1,j}(i-1 \geq 1), x_{i+1,j}(i+1 \leq n_1), x_{i,j-1}(j-1 \geq 1), x_{i,j+1}(j+1 \leq n_2).$$

Если для нейронов $x_{i-1,j}$, $x_{i+1,j}$, $x_{i,j-1}$, $x_{i,j+1}$ нарушается соответствующее им неравенство, то это означает, что нейрон x_{ij} является граничным и число его ближайших соседей уменьшается до трех или двух. Выходные сигналы элементов Z^k -полей определяются выражением

$$U_{\text{вых}} Z_{ij}^k = \begin{cases} 1, & \text{если } U_{\text{вх}} Z_{ij}^k \geq \theta^k; \\ 0, & \text{если } U_{\text{вх}} Z_{ij}^k < \theta^k, \end{cases} \quad (8)$$

где $U_{\text{вх}} Z_{ij}^k$ – входной сигнал нейрона Z_{ij}^k , $i = \overline{1, n_1}$, $j = \overline{1, n_2}$;

$$U_{\text{вх}} Z_{ij}^k = U_{\text{вых}} x_{ij} + \varepsilon \sum_{p1q \in M} U_{\text{вых}} x_{p1q};$$

ε – константа, $\varepsilon < 0.25$;

M – множество индексов, определяющих в изображении объекта соседние элементы нейрону x_{ij} ;

θ^k – порог нейрона поля Z^k , $\theta^k = 1 + \varepsilon k$, $k = 1, 4$.

Анализ выражения (8) показывает, что бинарные нейроны Z_{ij}^1 поля Z^1 переходят в возбужденное состояние, если возбужден нейрон x_{ij} и он имеет от одного до четырех соседних нейронов, имеющих единичные выходные сигналы. Аналогично, нейроны Z_{ij}^k поля Z^k ($k = 2, 3, 4$) переходят в возбужденное состояние, если возбужден нейрон x_{ij} и он имеет k соседних элементов, имеющих единичные выходные сигналы.

Нейроны \sum_p , $p = \overline{0, 4}$ имеют линейные функции активации, т.е. их выходные сигналы определяются соотношением

$$U_{\text{вых}} \Sigma_p = k_1 \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} U_{\text{вых}} Z_{ij}^p, \quad p = \overline{0,4}, \quad (9)$$

где k_1 – положительная константа; $U_{\text{вых}} Z_{ij}^p$ – выходной сигнал нейрона

Z_{ij}^p p -го поля, при этом принято, что $Z_{ij}^0 \equiv x_{ij}$.

Таким образом, на выходе нейрона Σ_0 будет сигнал, пропорциональный общему числу возбужденных x - нейронов и площади предъявленного изображения. Сигнал на выходе нейрона Σ_q ($q = \overline{1,4}$) пропорционален числу элементов изображения, имеющих от q до четырех возбужденных состояний. Используя выражение (9), нетрудно получить сигналы, пропорциональные числу N_0 нейронов, не имеющих соседних элементов в изображении, и сигналы, пропорциональные числам N_r ($r = \overline{1,3}$) нейронов, имеющих в изображении соответственно от одного до трех возбужденных элементов изображения:

$$\begin{aligned} N_0 &= k_1 (U_{\text{вых}} \Sigma_0 - U_{\text{вых}} \Sigma_1); \\ N_1 &= k_1 (U_{\text{вых}} \Sigma_1 - U_{\text{вых}} \Sigma_2); \\ N_2 &= k_1 (U_{\text{вых}} \Sigma_2 - U_{\text{вых}} \Sigma_3); \\ N_3 &= k_1 (U_{\text{вых}} \Sigma_3 - U_{\text{вых}} \Sigma_4). \end{aligned} \quad (10)$$

Правые части соотношения (10) в специализированном вычислительном устройстве выполняются с помощью нейронов n_q ($q = \overline{0,3}$), имеющих линейные функции активации.

Очевидно, что, имея на выходах нейронов n_q ($q = \overline{0,3}$) сигналы, пропорциональные числам N_r ($r = \overline{1,3}$), нетрудно с помощью нейрона Σ_p (рис. 3), имеющего линейную функцию активации и выполняющего суммирование выходных сигналов нейронов n_q в соответствии с соотношением (7), получить сигнал, пропорциональный длине периметра рассматриваемого изображения. Имея сигналы, пропорциональные площади и периметру изображения, нетрудно, с помощью специализированного блока или АЛУ получить и сигнал, пропорциональный правой части выражения (5). С помощью величины этого сигнала и выполняется классификация предъявляемых изображений.

Замечание. Поскольку нейроны $n_q(q = \overline{0,3})$ имеют нелинейные функции активации, то их наличие в структуре устройства не обязательно, так как их функции может выполнить нейрон \sum_p . Они приведены в структуре нейронной сети только с целью упрощения описания ее алгоритма функционирования.



Рис. 4. Пример обрабатываемого изображения

Эксперименты с изображениями вида, приведенного на рис.4, подтвердили работоспособность предлагаемых алгоритмов обработки изображений, положенных в основу специализированного вычислительного устройства на основе предложенной вычислительной сети.

ЛИТЕРАТУРА

1. Федер Е. Фракталы. – М.: Мир, 1991. – 254 с.
2. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – М.: Горячая линия –Телеком, 2001. – 382 с.

Поступила в редколлегию 15.10.2001
