

УДК 528.854

НЕЙРОАНАЛИТИЧЕСКИЙ МЕТОД КЛАССИФИКАЦИИ ОБРАЗОВ

д-р техн. наук, проф. А.С. ЯРМОЛЕНКО

(Новгородский государственный университет им. Ярослава Мудрого, Россия);

О.А. КУЦАЕВА

(Белорусская государственная сельскохозяйственная академия, Горки)

Представлено теоретическое обоснование использования нейронных сетей при распознавании образов. В результате исследований создан нейроаналитический алгоритм, позволяющий безошибочно осуществлять классификацию образов при отсутствии шумов входных сигналов. Но алгоритм распознавания образов должен быть эффективным и при наличии шумов. Зашумленные входные сигналы классифицируются существующими нейросетевыми алгоритмами неудовлетворительно. Эксперименты же с незначительными шумами в 0,2 пиксела подтверждают лишь незначительное повышение точности. Проверка эффективности разработанного нейроаналитического алгоритма привела к выводам, что необходимо разработать и исследовать робастный алгоритм классификации, который будет устойчив как к незначительным, так и к весьма ощутимым шумам.

Введение. К настоящему времени актуальной является задача оптимизации существующих методов классификации образов, разработка и реализация новых подходов к анализу данных дистанционного зондирования Земли. Это вызывается ростом значимости пространственных данных и необходимостью повышения точности результатов их обработки.

В отечественной геодезической, фотограмметрической литературе практически отсутствует разработка методов цифровой классификации образов. Использование же фирменных программ, в основном зарубежных, не раскрывает их алгоритмов, что не позволяет их дальнейшую разработку с целью повышения точности классификации. Кроме того, существующие нейросетевые алгоритмы даже при отсутствии шумов входных сигналов могут приводить к неверным значениям выходных сигналов (например, сети Хопфилда в [3, с. 114]), что, в частности, отражено в [1]. Даже незначительно зашумленные входные сигналы классифицируются существующими нейросетевыми алгоритмами неудовлетворительно [4, с. 182; 5, с. 202]; эксперименты же с ограниченными шумами в 0,2 пиксела подтверждают пренебрегаемо малое повышение точности лишь на 4 %, кроме того, они несущественны, так как не учитывают шумы в ± 1 пиксел. В связи с этим возникает проблема как совершенствования алгоритмов классификации образов, так и повышения их устойчивости к значительным шумам входных сигналов.

Теоретическое обоснование использования нейронных сетей при распознавании образов.

В общем случае процесс распознавания по аналогии [1] представим следующей аналитической зависимостью:

$$X = h \cdot l, \quad (1)$$

где X – вектор выходных сигналов; h – вектор-строка весов $h = (h_1, h_2, \dots, h_n)$; l – вектор входных сигналов.

Будем считать, что вектор l отягощен шумами (ошибками составляющих) и его точность характеризуется следующей ковариационной матрицей

$$K_l = \begin{pmatrix} \sigma_1^2 & K_{12} & \dots & K_{1n} \\ K_{21} & \sigma_2^2 & \dots & K_{2n} \\ \dots & \dots & \dots & \dots \\ K_{n1} & K_{n2} & \dots & \sigma_n^2 \end{pmatrix}. \quad (2)$$

Тогда корреляционная матрица вектора X будет иметь вид [9]:

$$K_x = h \cdot K_l \cdot h^T. \quad (3)$$

Можно принять входные сигналы равноточными с корреляционной матрицей:

$$K_l = \sigma^2 \cdot E, \quad (4)$$

где E – единичная матрица; σ – стандарт шума (ошибки положения пиксела изображения).

Тогда с учетом (4) будем иметь:

$$K_x = \sigma^2 \cdot h \cdot E \cdot h^T = \sigma^2 \cdot h \cdot h^T. \quad (5)$$

При обучении с учителем на вектор весов налагаются следующие условия:

$$\left. \begin{aligned} l_{0_1} \cdot h^T - X_{0_1} &= 0; \\ l_{0_2} \cdot h^T - X_{0_2} &= 0; \\ l_{0_3} \cdot h^T - X_{0_3} &= 0; \\ \dots\dots\dots\dots\dots\dots \\ l_{0_r} \cdot h^T - X_{0_r} &= 0, \end{aligned} \right\} \quad (6)$$

где r – число уравнений; X_{0_i} – код обучающего объекта-учителя; l_{0_i} – вектор обучающих сигналов, который можно представить следующим образом:

$$l_{0_i} = (l_{0_{i1}} \quad l_{0_{i2}} \quad \dots \quad l_{0_{in}}). \quad (7)$$

Уравнение (6) представим в матричном виде:

$$B \cdot h^T - X_0 = 0, \quad (8)$$

где

$$B = \begin{pmatrix} l_{0_{1,1}} & l_{0_{1,2}} & \dots & l_{0_{1,n}} \\ l_{0_{2,1}} & l_{0_{2,2}} & \dots & l_{0_{2,n}} \\ l_{0_{3,1}} & l_{0_{3,2}} & \dots & l_{0_{3,n}} \\ \dots & \dots & \dots & \dots \\ l_{0_{r,1}} & l_{0_{r,2}} & \dots & l_{0_{r,n}} \end{pmatrix}; \quad (9)$$

$$X_0 = \begin{pmatrix} X_{0_1} \\ X_{0_2} \\ \dots \\ X_{0_r} \end{pmatrix}. \quad (10)$$

При условиях (8) найдем вектор весов h^T на основе минимума функционала Лагранжа:

$$\phi = h \cdot h^T + 2K^T (B \cdot h^T - X_0) = \min. \quad (11)$$

Решение в соответствии с (11) будем называть решением методом наименьших квадратов (МНК). Очевидно, что при этом

$$\frac{\partial \phi}{\partial h^T} = 2h + 2K^T \cdot B = 0. \quad (12)$$

Выразим из этого уравнения h^T :

$$h = -K^T \cdot B; \quad (13)$$

$$h^T = -B^T \cdot K. \quad (14)$$

С учетом (8)

$$-B \cdot B^T \cdot K - X_0 = 0$$

и

$$K = -(B \cdot B^T)^{-1} \cdot X_0. \quad (15)$$

Соответственно на основании (13) и (1) найдем

$$X = X_0^T \cdot (B \cdot B^T)^{-1} \cdot B \cdot l. \quad (16)$$

Выражением (16) определяется распознавание образов с учителем.

Применение алгоритма при распознавании образов

Пусть распознаются фигуры, соответствующие цифрам от 0 до 9 (рис. 1).

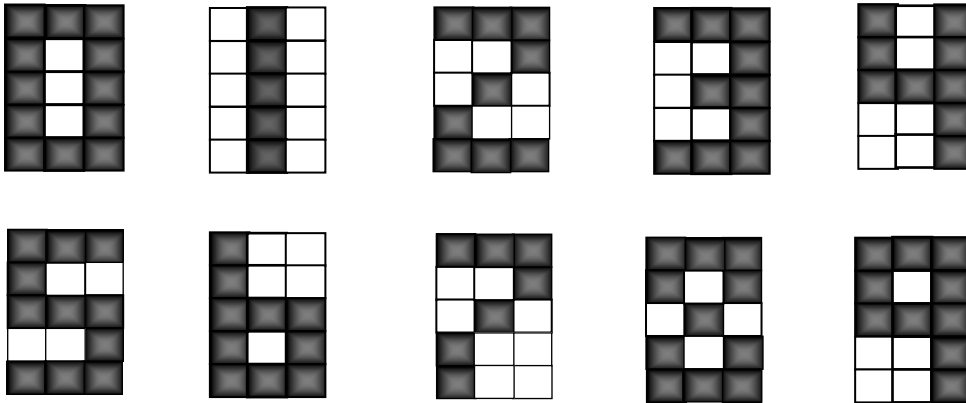


Рис. 1. Растровое изображение чисел от нуля до девяти

Пиксели черного цвета соответствует единичному значению бита, а белого – нулевому. Тогда матрица B (9) будет иметь следующий вид:

$$B = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Тогда

$$B \cdot B^T = \begin{pmatrix} 12 & 2 & 8 & 9 & 8 & 10 & 9 & 6 & 10 & 9 \\ 2 & 5 & 3 & 3 & 1 & 3 & 2 & 2 & 3 & 2 \\ 8 & 3 & 9 & 8 & 5 & 7 & 6 & 6 & 9 & 6 \\ 9 & 3 & 8 & 10 & 7 & 9 & 7 & 7 & 9 & 8 \\ 8 & 1 & 5 & 7 & 9 & 8 & 7 & 7 & 7 & 9 \\ 10 & 3 & 7 & 9 & 8 & 11 & 9 & 9 & 9 & 9 \\ 9 & 2 & 6 & 7 & 7 & 9 & 10 & 10 & 8 & 7 \\ 6 & 2 & 7 & 6 & 4 & 5 & 4 & 4 & 7 & 5 \\ 10 & 3 & 9 & 9 & 7 & 9 & 8 & 8 & 11 & 8 \\ 9 & 2 & 6 & 8 & 9 & 9 & 7 & 7 & 8 & 10 \end{pmatrix}.$$

Обратная ей матрица будет

$$(B \cdot B^T)^{-1} = \begin{pmatrix} 0,70 & 0,20 & -0,16 & -0,04 & 0,38 & -0,07 & -0,26 & 0,10 & -0,21 & -0,53 \\ 0,20 & 0,37 & -0,16 & -0,04 & 0,38 & -0,07 & -0,09 & 0,10 & -0,04 & -0,36 \\ -0,16 & -0,16 & 1,47 & -0,45 & -0,08 & 0,10 & -0,02 & -0,68 & -0,53 & 0,41 \\ -0,04 & -0,04 & -0,45 & 0,79 & -0,24 & -0,42 & 0,22 & 0,07 & -0,01 & 0,09 \\ 0,38 & 0,38 & -0,08 & -0,24 & 2,16 & 0,52 & -0,74 & 0,25 & -0,02 & -2,18 \\ -0,07 & -0,07 & 0,10 & -0,42 & 0,52 & 1,15 & -0,56 & 0,14 & -0,01 & -0,82 \\ -0,26 & -0,09 & -0,02 & 0,22 & -0,74 & -0,56 & 0,75 & 0,01 & -0,13 & 0,83 \\ 0,10 & 0,10 & -0,68 & 0,07 & 0,25 & 0,14 & 0,01 & 0,81 & -0,08 & -0,45 \\ -0,21 & -0,04 & -0,53 & -0,01 & -0,02 & -0,01 & -0,13 & -0,08 & 0,97 & -0,09 \\ -0,53 & -0,36 & 0,41 & 0,09 & 2,18 & -0,82 & 0,83 & -0,45 & -0,09 & 2,74 \end{pmatrix}.$$

Принимая $X_0^T = (0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9)$, найдем вектор коррелят (15):

$$K = \begin{pmatrix} -6,34507 \\ -2,34507 \\ -3,81690 \\ 0,92958 \\ -11,74648 \\ -3,14085 \\ 5,73944 \\ 1,69014 \\ 4,33099 \\ 13,69718 \end{pmatrix}$$

и вектор h в соответствии с (13):

$$h = (1,34 \ 5,00 \ -4,40 \ 2,54 \ -2,35 \ -1,26 \ -1,80 \ 5,34 \ -0,87 \ 1,60 \ -2,35 \ 3,47 \ -0,61 \ -4,65 \ -0,35).$$

Контрольное произведение $h \cdot B^T = (0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9)$ подтверждает абсолютную правильность алгоритма распознавания приведенного множества фигур.

При сравнении с исходными данными матрица выходных сигналов абсолютно совпадает с кодами, заданными по условию. Это справедливо при отсутствии шумов (ошибок) в исходных сигналах. Максимальные ошибки, возможные в исходных сигналах, составляют значения +1 или -1. Алгоритм распознавания образов должен быть эффективным и при наличии шумов. Проверим его эффективность на примере. Для этого введем в каждый образ как минимум по одной такой ошибке, построим соответствующую матрицу $B_{ош}$ и потребуем, чтобы и для зашумленного вектора удовлетворялись условия (6). В таком случае условные уравнения имеют вид

$$\begin{pmatrix} B \\ B_{ош} \end{pmatrix} h^T - \begin{pmatrix} X_0 \\ X_0 \end{pmatrix} = 0, \quad (17)$$

где

$$B_{ош} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}. \quad (18)$$

Здесь увеличенные нули и единицы являются ошибочными значениями. В месте нуля ошибка равна -1, а в месте единицы ошибка равна +1.

Решение системы (17) по алгоритму (11) – (16) при матрице (18) невозможно, так как определитель получаемой при этом матрицы нормальных уравнений может оказаться близким или равным нулю. В данном примере он составляет величину $2,27E - 74$.

$$N = \begin{pmatrix} B \\ B_{ош} \end{pmatrix} (B^T B_{ош}^T). \quad (19)$$

В таком случае систему можно решить методом регуляризации по Тихонову [2], представляя матрицу нормальных уравнений в виде регуляризирующего оператора

$$N + \lambda E, \quad (20)$$

где λ – параметр регуляризации; E – единичная матрица.

Обоснование (20) базируется на двойственной задаче условной оптимизации. Прямая задача в данном случае записывается так:

$$\Phi = h^T h - {}^2 K^T \begin{pmatrix} B \\ B_{ou} \end{pmatrix} h^T - \begin{pmatrix} X_0 \\ X_0 \end{pmatrix} = \min. \quad (21)$$

Полагая

$$\frac{\partial \Phi}{\partial h^T} = 2h - 2K^T \begin{pmatrix} B \\ B_{ou} \end{pmatrix} = 0. \quad (22)$$

Найдем

$$h = K^T \begin{pmatrix} B \\ B_{ou} \end{pmatrix}. \quad (23)$$

Подставляя (23) в (21), получаем формулировку двойственной задачи:

$$\Phi = K^T \begin{pmatrix} B \\ B_{ou} \end{pmatrix} (B^T B_{ou}^T) K - {}^2 K^T \begin{pmatrix} B \\ B_{ou} \end{pmatrix} (B^T B_{ou}^T) K + {}^2 K^T \begin{pmatrix} X_0 \\ X_0 \end{pmatrix}, \quad (24)$$

или

$$\Phi = -K^T \begin{pmatrix} B \\ B_{ou} \end{pmatrix} (B^T B_{ou}^T) K + {}^2 K^T \begin{pmatrix} X_0 \\ X_0 \end{pmatrix}, \quad (25)$$

которая сводится к максимизации целевой функции (25) по вектору коррелят K или к минимизации целевой функции:

$$\Phi = K^T \begin{pmatrix} B \\ B_{ou} \end{pmatrix} (B^T B_{ou}^T) K - {}^2 K^T \begin{pmatrix} X_0 \\ X_0 \end{pmatrix}. \quad (26)$$

Для получения гладкого решения, удовлетворяющего условиям непрерывности при наличии вырожденных матриц нормальных уравнений, выражение (25) дополняется стабилизатором $K^T K$ (26) и записывается в виде функционала Тихонова [2, с. 358].

$$\Phi = K^T \begin{pmatrix} B \\ B_{ou} \end{pmatrix} (B^T B_{ou}^T) K - {}^2 K^T \begin{pmatrix} X_0 \\ X_0 \end{pmatrix} + \lambda K^T K = \min. \quad (27)$$

Дифференцирование (27) по K с учетом обозначения (19) приводит к уравнению

$$\frac{\partial \Phi}{\partial K} = K^T N - (X_0' \ X_0') + \lambda K^T = 0 \quad (28)$$

и, соответственно, к нормальным уравнениям в матричном виде:

$$K^T (N + \lambda E) = (X_0' \ X_0') \quad (29)$$

с решением

$$K = (N + \lambda E)^{-1} \begin{pmatrix} X_0' \\ X_0' \end{pmatrix}. \quad (30)$$

С учетом (30) и (23) окончательное решение будет

$$h^T = (B^T B_{ou}^T) (N + \lambda E)^{-1} \begin{pmatrix} X_0' \\ X_0' \end{pmatrix}. \quad (31)$$

На нашем примере покажем, что можно подобрать такой параметр регуляризации λ , при котором система нормальных уравнений имеет единственное решение и достигается минимум функции

$$\psi = V^T V, \quad (32)$$

где

$$V = \begin{pmatrix} B \\ B_{ou} \end{pmatrix} h^T - \begin{pmatrix} X_0 \\ X_0 \end{pmatrix}. \quad (33)$$

Отметим, что система уравнений (32) должна быть переопределенной и в ней не должно быть линейно зависимых или повторяющихся столбцов, т.е. столбцы матрицы $\begin{pmatrix} B \\ B_{ou} \end{pmatrix}$ не должны быть взаимно коллинеарными (должна отсутствовать их мультиколлинеарность).

Минимизация функционала (31) по вектору h^T приводит к следующей системе нормальных уравнений:

$$Nh^T + (B^T B_{ou}^T) \begin{pmatrix} X_0 \\ X_0 \end{pmatrix} = 0, \quad (34)$$

где

$$N = (B^T B_{ou}^T) \begin{pmatrix} B_0 \\ B_{ou} \end{pmatrix}. \quad (35)$$

И соответствующее решение будет следующим:

$$h^T = -N^{-1} (B^T B_{ou}^T) \begin{pmatrix} X_0 \\ X_0 \end{pmatrix}. \quad (36)$$

Контроль решения осуществляется по формуле

$$Y = \begin{pmatrix} B_0 \\ B_{ou} \end{pmatrix} h^T. \quad (37)$$

Оценка точности распознавания образов при наличии шумов

№ п/п	Решение по формуле (31) при $\lambda = 0,00001$	Решение по формуле (36)	Контроль по формуле (37). Выходные сигналы*	Ближайший образ	Истинные значения выходных сигналов	Истинные ошибки распознавания	
						по ближайшему образу	по контрольному значению
1	2	3	4	5	6	7	8
1	8,64649E-05	4,26326E-14	0,183516	0	0	0	0,1835
2	2,187336238	2,187379266	2,185448	2	1	+1	1,1854
3	-0,633575857	-0,633612363	1,683194	2	2	0	-0,317
4	0,918282325	0,918222795	3,000000	3	3	0	7E-14
5	-1,869287186	-1,869285254	5,685769	6	4	+2	1,6858
6	-1,050860897	-1,050869285	5,316806	5	5	0	0,3168
7	0,347706542	0,347714102	4,948487	5	6	-1	-1,052
8	5,267800165	5,267868641	7,684482	8	7	+1	0,6845
9	-1,969711682	-1,969735995	7,073406	7	8	-1	-0,927
10	1,185422749	1,185447521	7,873149	8	9	-1	-1,127
11	0,934982894	0,934964585	0,183516	0	0	0	0,1835
12	4,47185624	4,471989697	-0,001930	0	1	-1	-1,002
13	0,728237373	0,728267869	2,316806	2	2	0	0,3168
14	-4,335403062	-4,335479717	3,918223	4	3	+1	0,9182
15	-1,665802982	-1,665808113	3,816484	4	4	0	-0,184
16			4,265937	4	5	-1	-0,734
17			6,918223	7	6	+1	0,9182
18			6,499034	6	7	-1	-0,501
19			8,008371	8	8	0	0,0084
20			8,808113	9	9	0	-0,192
СКО						0,84	0,77

Исходя из представленной выше таблицы можно сделать следующие *выводы*:

- наличие в образе хотя бы одного ошибочного пиксела даже в процессе обучения ведет к ошибкам его распознавания в 100 процентах случаях;
- округление вычисляемых значений выходных сигналов с точностью 0,5 позволяет получать ближайшие образы (графа 5), которые в 11 случаях из 20 не совпадают с желаемым откликом, т.е. с истинным значением образа;
- наличие значительных шумов, равных ± 1 пикселу, одинаково часто искажает классификацию как идеальных, так и зашумленных сигналов;
- в случае отсутствия шумов классификация по разработанному алгоритму осуществляется идеально, что является явным преимуществом его перед алгоритмом Хопфилда [3, с. 114] и существенно улучшает алгоритм, предложенный в [1];
- зашумленные входные сигналы классифицируются существующими нейросетевыми алгоритмами неудовлетворительно [4, с. 182; 5, с. 202]; эксперименты же с незначительными шумами в 0,2 пиксела подтверждают незначительное повышение точности лишь на 4 %, кроме того, они не существенны, так как базируются на пренебрегаемо малых ошибках сигналов.

Заключение. В результате выполненной работы создан нейроаналитический алгоритм, позволяющий безошибочно осуществлять классификацию образов при отсутствии шумов входных сигналов. Однако исследования показали, что при наличии шумов возникает проблема достоверной классификации образов. В связи с этим возникает необходимость разработки алгоритмов максимально устойчивых к значительным шумам, т.е. робастных алгоритмов.

ЛИТЕРАТУРА

1. Ярмоленко, А.С. Применение теории нейронных сетей в геоинформатике / А.С. Ярмоленко // Изв. вузов. Геодезия и аэрофотосъемка. – 2008. – № 2. – С. 33 – 44.
2. Хайкин, С. Нейронные сети: полный курс / С. Хайкин. – 2-е изд. испр.; пер. с англ. – М.: ООО «И.Д. Вильямс», 2006. – 1104 с.
3. Каллан, Р. Основные концепции нейронных сетей / Р. Калан; пер. с англ. – М.: Издат. дом «Вильямс», 2001. – 288 с.
4. Осовский, С. Нейронные сети для обработки информации / С. Осовский; пер. с польск. И.Д. Рудинского. – М.: Финансы и статистика, 2002. – 344 с.
5. Медведев, В.С. Нейронные сети. MATLAB 6 / В.С. Медведев, В.Г. Потемкин; под общ. ред. В.Г. Потемкина. – М.: ДИАЛОГ-МИФИ, 2002. – 496 с. – (Пакеты прикладных программ; Кн. 4).

Поступила 14.12.2012

NEUROANALITICAL METHOD OF IMAGE CLASSIFICATION

A. YARMOLENKO, A. KUTSAEVA

In this article theoretical justification of use of neural networks at recognition of images is considered. As a result of research the neuroanalytical algorithm allowing unmistakably to carry out classification of images in the absence of noise of entrance signals is created. But the algorithm of recognition of images has to be effective and in the presence of noise. Noisy entrance signals are classified by existing neuronetwork algorithms unsatisfactorily. Experiments with insignificant noise in 0,2 pixels confirm only slight increase of accuracy. Check of efficiency of the developed neuroanalytical algorithm led to conclusions that it is necessary to develop and research robustn algorithm of classification which will be steady both to insignificant and against very considerable noise.