



## КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

УДК 004.932

### ПРИМЕНЕНИЕ МЕТОДА ЭКСТРЕМАЛЬНОГО ОБУЧЕНИЯ НЕЙРОННОЙ СЕТИ ДЛЯ КЛАССИФИКАЦИИ ОБЛАСТЕЙ ИЗОБРАЖЕНИЯ

**Д. А. ЮДИН**  
**В. З. МАЕРГУТ**

*Белгородский  
государственный  
технологический  
университет  
им. В.Г. Шухова*

*e-mail:  
yuddim@yandex.ru  
valerymag@nm.ru*

При создании систем технического зрения для промышленных приложений, работающих в режиме реального времени, важно применение эффективных методов распознавания изображений. В статье рассмотрен способ классификации областей изображения, основанный на обучении нейронной сети прямого распространения с помощью метода экстремального обучения (ELM, Extreme Learning Machine). Определено количество нейронов скрытого слоя нейронной сети, обеспечивающее максимальную точность классификации на тестовой выборке и приемлемую точность на обучающей выборке. Произведено сравнение с методом опорных векторов (SVM), вероятностной нейронной сетью (PNN), методом обратного распространения ошибки для двухслойной нейронной сети и радиально-базисной сетью, которое показало преимущество классификации с применением метода экстремального обучения как по быстродействию, так и по точности.

Ключевые слова: метод экстремального обучения, нейронные сети, классификация, области изображения, распознавание.

При создании систем технического зрения для промышленных приложений, работающих в режиме реального времени, важно применение эффективных методов распознавания изображений. Задача распознавания изображений может, в частности, заключаться в их классификации изображений на основе определенных требований, причем изображения, относящиеся к одному классу образов, должны обладать относительно высокой степенью близости [1]. При классификации изображение или область изображения может рассматриваться как совокупность вычисляемых признаков, которые требуется отобразить на множество классов (образов). Существует ряд задач распознавания, где изображение описывается как структурная модель, на элементы которой накладываются определенные связи [2].

Задача классификации эффективно решается с применением аппарата искусственных нейронных сетей [3], при этом важен выбор наиболее подходящей архитектуры сети и метода обучения.

В настоящей статье рассматривается вопрос классификации малых областей изображения для обеспечения корректной сегментации изображения, выполняемой в реальном масштабе времени.

Изображение размером  $U \times V$  разбивается на малые области размером  $M \times M$ , каждой из которых необходимо сопоставить один из нескольких классов, присутствующих на изображении. В частном случае, на рис. 1 представлена фотография процесса обжига во вращающейся цементной печи, на которой необходимо выделить четыре типа (класса) областей: «факел», «материал», «футеровка», «корпус и

горелка печи» [4]. В данной задаче будут рассматриваться именно изображения процесса обжига в печах, характеризующихся отсутствием четко выраженных контуров и высокой зашумленностью.

В общем виде постановка задачи выглядит как  $f: P \rightarrow T$  – требуется решить задачу классификации, то есть поставить в соответствие вектору признаков  $P$  номер класса  $T$  области изображения.

Вектор признаков  $P$ , по которому требуется осуществить классификацию, для данной задачи имеет четыре составляющие  $P = [P^1, P^2, P^3, P^4]$ , которые представляют собой текстурные характеристики областей изображения [4]:

$$1) \text{ Автокорреляция} - P^1 = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} (ij) P_{ij}^d,$$

где  $K$  – количество градаций яркости изображения,  $P_{i,j}^d$  – матрица смежности,  $d$  – расстояние между точками при подсчете матрицы смежности.

$$2) \text{ Тень} - P^2 = \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} (i+j-2\mu)^3 P_{ij}^d,$$

где  $\mu$  – математическое ожидание элементов матрицы  $P_{ij}^d$ .

$$3) \text{ Интенсивность} - P^3 = \frac{1}{M^2} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} f(i, j),$$

где  $f(i, j)$  – яркость пикселя в точке  $(i, j)$  области изображения размером  $M \times M$ .

$$4) \text{ Вариация} - P^4 = \sum_{g=0}^{K-1} \left( g - \frac{1}{M^2} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} f(i, j) \right) C(g),$$

где  $C(g)$  – количество пикселей с яркостью  $g$ .



Рис. 1. Пример исходного изображения, требующего классификации областей

В рассматриваемой задаче номер класса  $T$  области изображения может принимать значения от 1 до 4-х: номеру 1 соответствует область изображения типа «факел», номеру 2 – область «материал», номеру 3 – область «футеровка», номеру 4 – область «корпус и горелка печи». В данном случае удобно номер класса  $T$  представить в виде вектора  $Y$ , имеющего четыре составляющих, каждая из которых соответствует одному из классов и представляют собой вещественное число из



диапазона  $[0, 1]$ . При этом номер класса  $T$  вычисляется как номер составляющей вектора  $Y$ , имеющей максимальное значение

$$T = \underset{i}{\operatorname{arg\,max}}(Y^i). \tag{1}$$

В качестве классификатора в данной работе применяется трехслойная нейронная сеть с одним скрытым слоем (SLFN), обучаемая с помощью современного метода экстремального обучения [5] (ELM, Extreme Learning Machine).

Метод экстремального обучения (ELM) позволяет без итерационной процедуры обучить трехслойную нейронную сеть с одним скрытым слоем, в данном случае, имеющим сигмоидальную функцию активации  $G$  и выходным слоем, выдающим взвешенную сумму своих входов [5].

Структура классификатора для общего случая представлена на рис. 2.

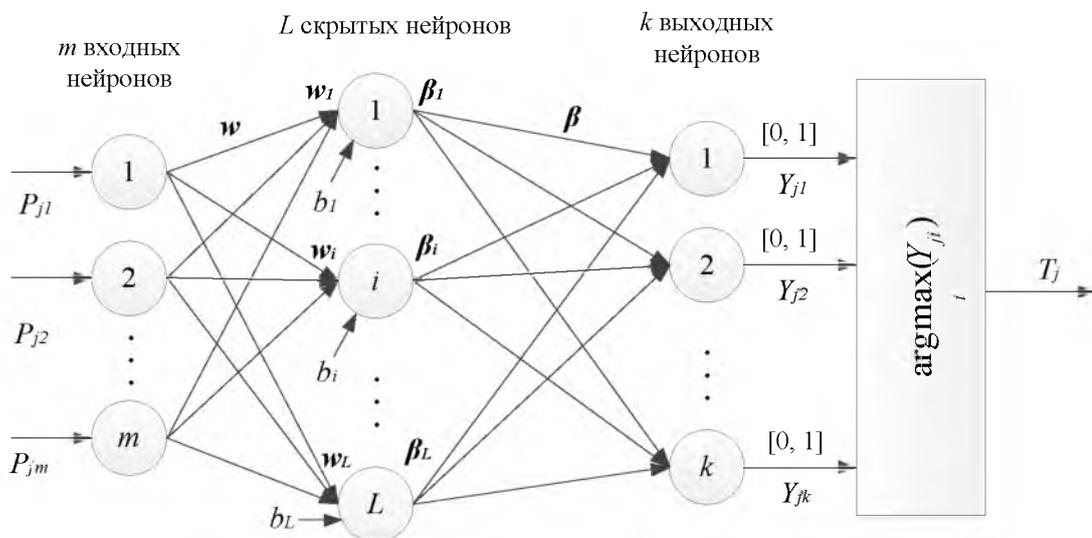


Рис. 2. Структура нейронной сети, обучаемой методом экстремального обучения

Обучение классификатора происходит за счет подачи  $N$  заранее подготовленных примеров, для каждого из которых задана пара векторов  $(P_j, Y_j), j = 1, \dots, N$ , где  $P_j$  –  $j$ -й входной вектор признаков, содержащий  $m = 4$  составляющих, описанных выше, а  $Y_j$  –  $j$ -й выходной вектор классов с  $k = 4$  составляющими.

Входные векторы признаков образуют матрицу входных значений  $P$  классификатора, имеющую размерность  $N \times m$ ,  $m$  – количество входных нейронов:

$$P = \begin{bmatrix} P_1 \\ \dots \\ P_N \end{bmatrix} = \begin{bmatrix} P_{11} & \dots & P_{1m} \\ \dots & \dots & \dots \\ P_{N1} & \dots & P_{Nm} \end{bmatrix}_{N \times m}.$$

Выходные векторы классов образуют матрицу выходных значений  $Y$  классификатора, имеющую размерность  $N \times k$ ,  $k$  – количество выходных нейронов:

$$Y = \begin{bmatrix} Y_1 \\ \dots \\ Y_N \end{bmatrix} = \begin{bmatrix} Y_{11} & \dots & Y_{1k} \\ \dots & \dots & \dots \\ Y_{N1} & \dots & Y_{Nk} \end{bmatrix}_{N \times k}.$$

Матрица входных весов нейронов скрытого слоя  $w$  представлена следующим образом:

$$w = \begin{bmatrix} w_1 \\ \dots \\ w_L \end{bmatrix} = \begin{bmatrix} w_{11} & \dots & w_{1m} \\ \dots & \dots & \dots \\ w_{L1} & \dots & w_{Lm} \end{bmatrix}_{L \times m},$$



где  $w_i$  – вектор входных весов  $i$ -го нейрона скрытого слоя,  $i = 1, \dots, L$ ,  $L$  – количество нейронов скрытого слоя.

В случае применения метода экстремального обучения входные веса  $w_{ij}$  и порог  $b_i$  для каждого  $i$ -го нейрона скрытого слоя задаются с помощью генератора случайных величин:

$$w_{ij} = \text{random}(-1 \dots 1),$$

$$b_i = \text{random}(0 \dots 1).$$

Матрица выходных весов нейронов скрытого слоя  $\beta$ , имеющая размерность  $L \times k$ :

$$\beta = \begin{bmatrix} \beta_1 \\ \dots \\ \beta_L \end{bmatrix} = \begin{bmatrix} \beta_{11} & \dots & \beta_{1k} \\ \dots & \dots & \dots \\ \beta_{L1} & \dots & \beta_{Lk} \end{bmatrix}_{L \times k},$$

где  $\beta_i$  – вектор выходных весов  $i$ -го нейрона скрытого слоя,  $i = 1, \dots, L$ .

Выходной вектор  $Y_j$  вычисляется как

$$Y_j = \sum_{i=1}^L \beta_i G(w_i P_j^T + b_i), \quad j = 1, 2, \dots, N, \quad (2)$$

или в матричном виде

$$Y = H\beta, \quad (3)$$

где  $H$  – матрица выходных значений нейронов скрытого слоя, имеющая размерность  $N \times L$ ,  $N$  – количество примеров,  $L$  – количество нейронов скрытого слоя:

$$H = \begin{bmatrix} G(w_1 P_1 + b_1) & \dots & G(w_L P_1 + b_L) \\ \dots & \dots & \dots \\ G(w_1 P_N + b_1) & \dots & G(w_L P_N + b_L) \end{bmatrix}_{N \times L},$$

где функция  $G$  является сигмоидальной функцией:

$$G(w_i P_j + b_i) = \frac{1}{1 + \exp(w_i P_j + b_i)}.$$

Обучение сети заключается в вычислении матрицы выходных весов нейронов скрытого слоя  $\beta$  по формуле:

$$\beta = H^+ Y, \quad (4)$$

где  $H^+ = (H^T H)^{-1} H^T$  – псевдообратная матрица (матрица Мура-Пенроуза) для случая невырожденной матрицы  $H$ .

Из теории матриц [6] известно, что выражение (4) является наилучшим приближенным решением уравнения (3) по методу наименьших квадратов, что позволяет избежать локальных минимумов при обучении сети.

Когда нейронная сеть обучена, т.е. по формуле (4) найдена матрица выходных весов нейронов скрытого слоя  $\beta$ , при подаче на вход сети вектора признаков  $P$ , не содержащегося в обучающей выборке, выходной вектор вычисляется на основе (2). Затем по формуле (1) вычисляется номер класса  $T$ , к которому относится входной вектор.

При появлении новых обучающих примеров и при возникновении необходимости переобучить сеть снова формируется входная матрица  $P$ , выходная матрица  $Y$  и по формуле (4) пересчитывается матрица выходных весов нейронов скрытого слоя  $\beta$ . Необходимо отметить, что при большом количестве обучающих примеров, требуется вычисление псевдообратной матрицы большого размера, однако эту операцию позволяют эффективно производить существующие библиотеки алгоритмов линейной алгебры, например LAPACK [7]. В зависимости от задачи можно также ограничивать количество обучающих примеров.

Точность классификации находится как процент правильно классифицированных примеров:

$$Q = \left(1 - \frac{N^*}{N}\right) \cdot 100\%, \quad (5)$$

где  $N^*$  – количество неправильно классифицированных примеров (для которых неверно вычислен номер класса  $T$ ),  $N$  – общее число примеров.

За оценку точности вычисления выходной матрицы можно также принять среднеквадратическое отклонение полученной выходной матрицы  $Y^*$  от эталонной выходной матрицы  $Y$ , задаваемой обучающей выборкой из  $N$  примеров:

$$\sigma = \sqrt{\frac{1}{N} \cdot (Y - Y^*)^2}. \quad (6)$$

Рассмотрим для примера работу классификатора основанного на методе экстремального обучения на обучающей выборке, входные векторы признаков  $P_j$  которой имеют две составляющие  $\{P^1, P^2\}$ . Данные векторы признаков необходимо отнести к одному из двух классов  $T = \{1, 2\}$ . При этом рассматривалась нейронная сеть с двумя входными нейронами, 30-ю нейронами скрытого слоя с сигмоидальной функцией активации и 2-мя выходными нейронами с линейной функцией активации.

На рис. 3 – рис. 5 представлены результаты работы классификатора на разных обучающих выборках. Примеры отмеченные «крестом» относятся к классу 1, отмеченные «кружком» – к классу 2.

Для линейно разделимой обучающей выборки, представленной на рис. 3а, точность работы классификатора, вычисленная по формуле (5) составила 100%, величина среднеквадратического отклонения, найденная в соответствии с (6), равна 0,0322.

Для обучающей выборки, представленной на рис. 4а, требующей нелинейного разделения на классы, точность работы классификатора составила 98,41%, среднеквадратическое отклонение – 0,3518.

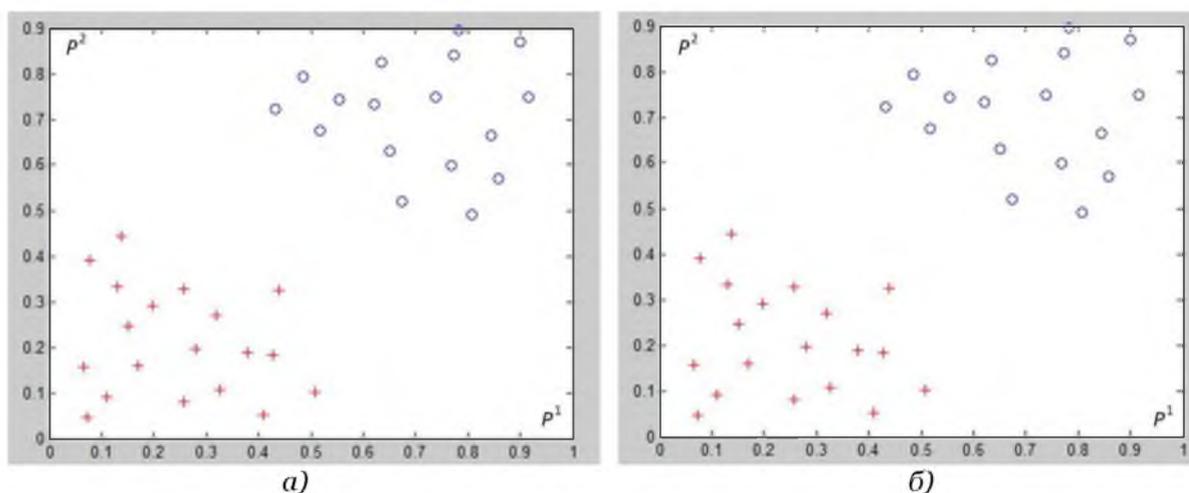


Рис. 3. Результат работы классификатора на линейно разделимой обучающей выборке:  
а – обучающая выборка, б – результат классификации с применением метода экстремального обучения

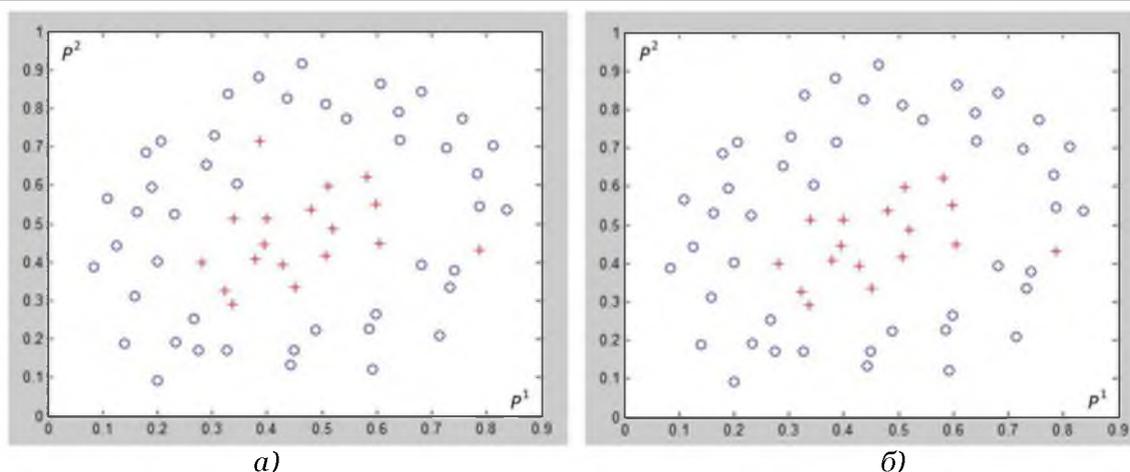


Рис. 4. Результат работы классификатора на линейно неразделимой обучающей выборке: а – обучающая выборка, б – результат классификации с применением метода экстремального обучения

Для случая повышенной зашумленности обучающей выборки (рис. 5) точность классификации снижается и составляет 91,43%, среднеквадратическое отклонение – 0,5731.

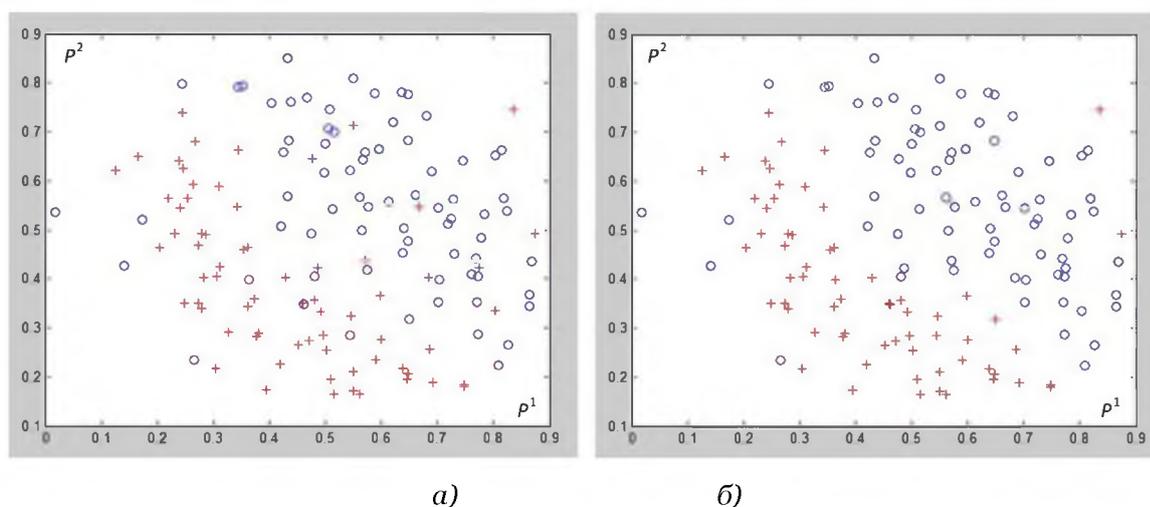


Рис. 5. Результат работы классификатора на обучающей выборке с повышенной зашумленностью: а – обучающая выборка, б – результат классификации с применением метода экстремального обучения

При решении задачи классификации областей изображения входным изображением является фотография процесса обжига размером  $800 \times 600$  пикселей, оно разбивается на области размером  $20 \times 20$  пикселей, таким образом, имеется 1200 областей изображения. Объем обучающей выборки составляет 600 областей изображения, объем тестовой выборки – также 600 областей.

Скрытый слой имеет 30 нейронов с сигмоидальной функцией активации. Такое количество нейронов обеспечивает для данной структуры сети и метода обучения максимальную точность классификации на тестовой выборке и приемлемую точность на обучающей выборке – соответственно 90,5% и 92,33%, что показано на рис. 6.

Работа сети, обучаемая методом экстремального обучения (ELM), с помощью пакета MATLAB сопоставлена с функционированием нейронных сетей, обучаемых другими широко применяемыми способами: методом опорных векторов (SVM) [8], вероятностной нейронной сетью (PNN), методом обратного распространения ошибки (BP) для двухслойной нейронной сети и радиально-базисной сетью [9] (таблица).



Необходимо отметить, что обучение нейронной сети на основе метода экстремального обучения превосходит другие способы, как по быстродействию, так и по точности классификации и на обучающей, и на тестовой выборке.

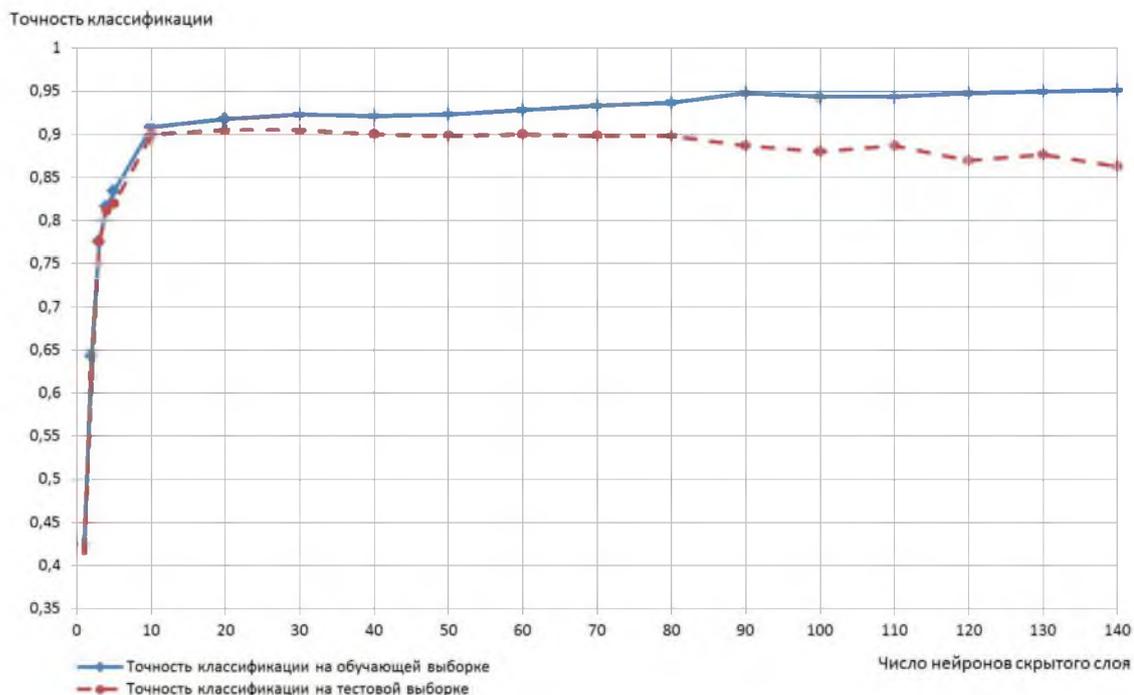


Рис. 6. Точность классификации при разном количестве нейронов скрытого слоя в нейронной сети, обучаемой с помощью метода экстремального обучения

Таблица

**Сравнение методов классификации областей изображения**

№ пп	Метод	Точность классификации на обучающей выборке, %	Точность классификации на тестовой выборке, %	Время обучения, мс
1	Метод экстремального обучения (ELM)	92,33	90,5	7,8
2	Метод опорных векторов (SVM)	89,83	88,5	15,6
3	Вероятностная нейронная сеть (PNN)	90,5	89,67	31,2
4	Метод обратного распространения ошибки (BP)	89,67	88,33	483,6
5	Радиально-базисная сеть	91,33	89,17	499,2

Время обучения получено с применением пакета MATLAB и указано для компьютера с четырехъядерным процессором Intel Core i5-2500S с частотой 2,7 ГГц, ОЗУ 4 Гб, с установленной 64-разрядной операционной системой Windows 7.

Результат классификации областей изображения процесса обжига с применением методом экстремального обучения нейронной сети приведен на рис. 7, для наглядности выбраны только три признака (текстурные характеристики), рассмотренных выше  $\{P^1, P^2, P^4\}$ .

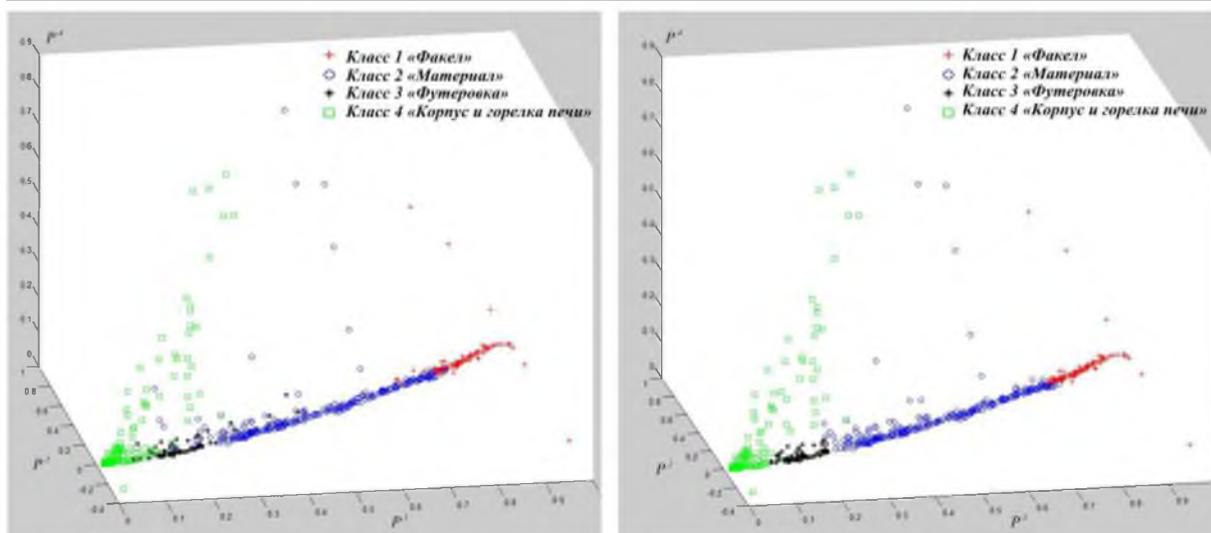


Рис. 7. Результат работы классификатора: а – обучающая выборка, б – результат классификации с применением метода экстремального обучения нейронной сети

Полученная классификация областей изображения приводит к сегментации изображения процесса обжига, представленной на рис. 8.

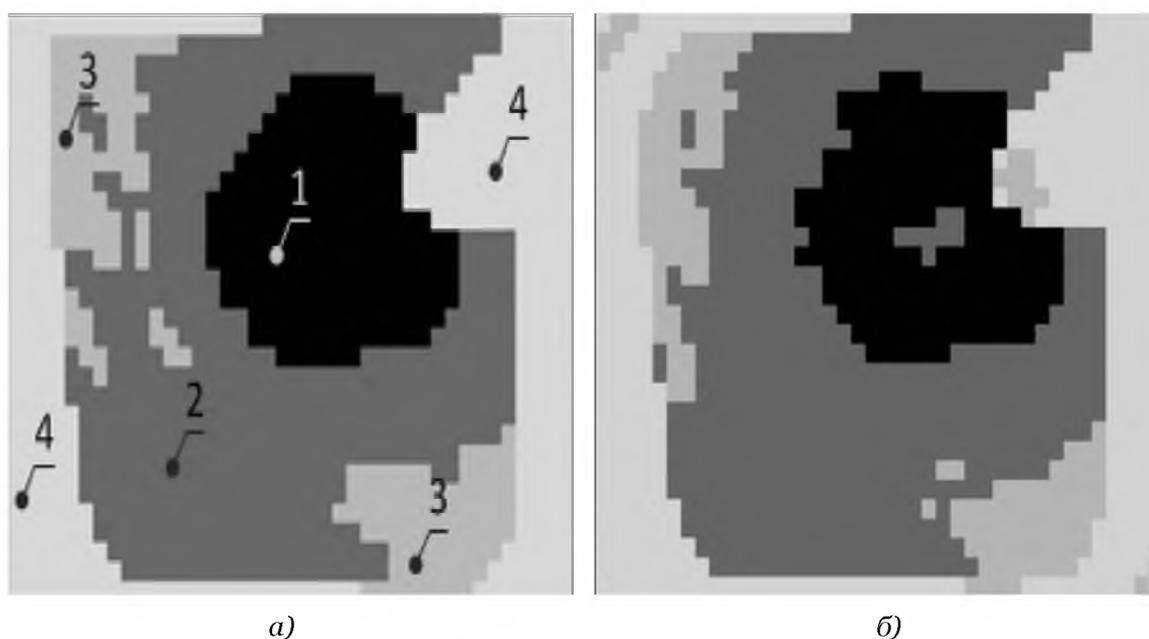


Рис. 8. Результат классификации областей изображения: а – эталонная классификация областей изображения размером 20×20 пикселей на классы: 1- факел, 2 – материал, 3 – футеровка, 4 – корпус и горелка печи, б – результат классификации с применением метода экстремального обучения

Таким образом, рассматриваемый классификатор на основе метода экстремального обучения обеспечивает приемлемое качество разбиения областей изображения на классы и высокое быстродействие, что делает его привлекательным для решения задач классификации в составе систем технического зрения, работающих в режиме реального времени. Кроме того рассмотренный метод может быть применен для распознавания изображений в системах технического зрения мобильных роботов, что позволяет повысить живучесть мобильного робота за счет распознавания и оценки визуальной информации об окружающей среде [10].



Работа выполнена в рамках гранта № А-27/12 Программы стратегического развития БГТУ им. В.Г. Шухова на 2012-2016 г.г. (№ 2011-ПР-146) и гранта РФФИ № 12-07-97526-р\_центр\_а.

### Литература

1. Методы компьютерной обработки изображений / Под ред. В.А. Сойфера. – 2-е изд., испр. – М.: ФИЗМАТЛИТ, 2003. – 784 с.
2. Визильтер Ю.В. Желтов С.Ю., Бондаренко А.В., Осоков М.В., Моржин А.В. Обработка и анализ изображений в задачах машинного зрения: Курс лекций и практических занятий. – М.: Физматкнига, 2010. – 672 с.
3. Haykin S. Neural Networks and Learning Machines. – 3rd ed. – Prentice Hall, 2009. – 906 p.
4. Юдин Д.А., Магергут В.З. Применение текстурного анализа для сегментации видеоизображений процесса обжига // Сб. трудов XXV Междунар. науч. конф. ММТТ-25: в 10 т. Т. 4, Харьков: Национ. техн. ун-т «ХПИ», 2012 – С. 59-63.
5. G.-B. Huang, et al., Extreme learning machine: Theory and applications. Neurocomputing, vol. 70, pp. 489-501, 2006.
6. Гантмахер Ф. Р. Теория матриц. – 5-е изд. – М.: ФИЗМАТЛИТ, 2004. – 560 с.
7. LAPACK – Linear Algebra PACKage – 2012. – Режим доступа: <http://www.netlib.org/lapack/>
8. С.-С. Chang, С.-J. Lin. LIBSVM – A Library for Support Vector Machines –Режим доступа: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
9. Медведев В.С., Потемкин В.Г. Нейронные сети. MATLAB 6 / Под общ. ред. к.т.н. В.Г. Потемкина. – М.: ДИАЛОГ-МИФИ, 2002. – 496 с.
10. Рубанов В.Г. Системный подход к проектированию управляемых мобильных логистических средств, обладающих свойством живучести // Научные ведомости БелГУ. Серия: История, Политология, Экономика, Информатика. – 2011. – №1 (96). – Вып. 17/1. – С. 176-187.

## APPLICATION OF EXTREME LEARNING MACHINES FOR CLASSIFICATION OF IMAGE AREAS

**D. A. YUDIN**  
**V. Z. MAGERGUT**

*Belgorod State  
Technological University  
named after V.G. Shukhov*

*e-mail:  
yuddim@yandex.ru  
valerymag@nm.ru*

In creation machine vision systems for industrial applications, working in real time, it is important to use effective methods of image recognition. This article describes a method of classification of an image areas, based on a learning feedforward neural network with extreme learning machine (ELM). Determined the number of neurons in the hidden layer of the neural network that provides the highest classification accuracy on a test set and an acceptable accuracy for the training set. Made a comparison of the method with support vector machines (SVM), probabilistic neural network (PNN), error back-propagation for two-layer neural network and radial-basic network. It shows the advantage of the method of classification using the extreme learning machine as on computational speed, and accuracy.

Keywords: extreme learning machines, neural networks, classification, image area, recognition.