

Использование комбинированной структуры искусственной нейронной сети для распознавания образов

77-30569/409197

03, март 2012

Стругайло В. В.

УДК 004.932

Россия, Московский автомобильно-дорожный государственный технический университет
strugailov@mail.ru

Одним из актуальных и перспективных направлений для решения задач распознавания образов является применение искусственных нейронных сетей (ИНС).

В данной статье рассматриваются модели и сравниваются результаты распознавания образов на изображениях с помощью классических полносвязных трехслойных ИНС, основанных на алгоритме с обратным распространением ошибок. Первая модель представлена в виде отдельной ИНС с пятью выходными нейронами. Вторая модель представлена в виде комбинированной структуры, в которой каждая ИНС имеет по одному выходному нейрону и предназначена для распознавания одного класса объектов, но конечный результат распознавания в рамках модели определяется по порогу от выходов ИНС. С помощью рассматриваемых моделей распознаются по пять классов объектов представленных круглой, овальной, каплевидной, серповидной и нерегулярной геометрической формой.

Форма представления изображений

Изображения, используемые в обучающих примерах, были получены с помощью фотокамеры с различной разрешающей способностью. При фиксированной размерности изображений, содержащиеся в них образы, отличаются по размерам. В используемых при обучении и тестировании примерах каждое отдельное изображение содержит только один образ.

Используемые изображения являются растровыми и были представлены в полутоновой цветовой модели в диапазоне значений яркостей от 0 до 255. В таблице 1 приведены размерности используемых в обучении и тестировании изображений.

Таблица 1 – Размерность изображений

Размерность матрицы изображения	Число элементов
64x64	4096
32x32	1024
16x16	256
8x8	64

Элементы изображений построчно преобразовывались в линейные вектора имеющие вид

$$X = \{x_1, x_2, \dots, x_i \dots x_L\},$$

где: $L = N \times N$, x_i – i -ый элемент вектора линейно представленного изображения.

В работе применяется нормирование значений элементов входных векторов X , подаваемых на ИНС по следующей формуле

$$a_i = \frac{x_i}{\sqrt{\sum_{i=1}^L x_i^2}},$$

где: a_i – нормированный элемент входного вектора; x_i – i -ый элемент входного вектора; L – число элементов вектора.

Нормированный вектор имеет вид

$$A = \{a_1, a_2, \dots, a_L\}.$$

Используемые изображения представлены в различных диапазонах яркостей. Такое представление связано с тем, что ИНС вполне могут запоминать такие характеристики как фон, вместо свойств образов действительно представляющих интерес.

Так как, ИНС вполне могут запоминать такие характеристики как фон, вместо свойств образов действительно представляющих интерес. Для снижения влияния такого фактора использовались изображения с различным диапазоном яркостей как внутри одного класса, так и по всем примерам.

Структура ИНС

На рисунке 1 приведена структура отдельной ИНС с пятью выходами, распознающая соответствующее число классов объектов. В модели, основанной комбинированной структуре, используются ИНС с такой же структурой, но с одним выходным нейроном.

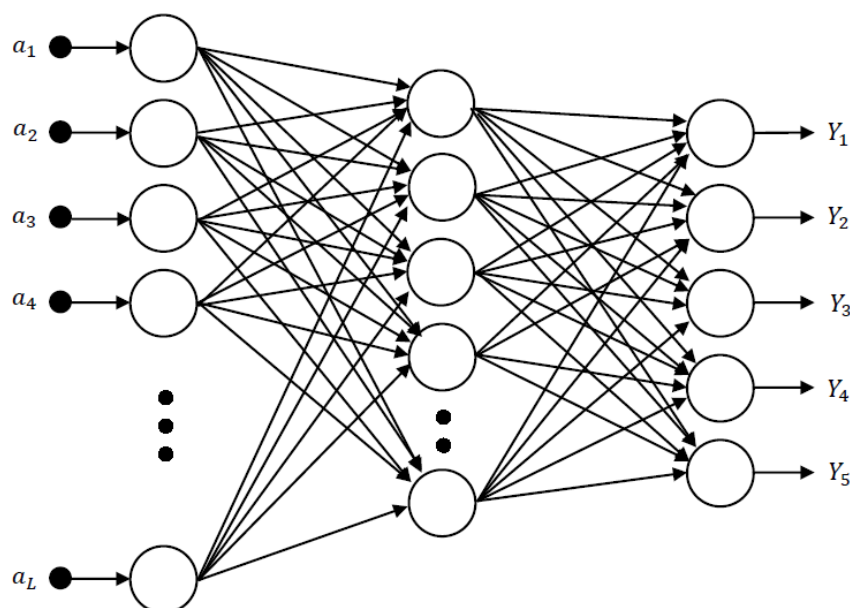


Рисунок 1 – Структура полносвязной нейронной сети

Модель на основе комбинированной структуры ИНС

Разработанная модель на основе комбинированной структуры для классификации образов показана на рисунке 2. В данной модели входной образ анализируется комбинацией ИНС. Каждая ИНС распознает один класс изображений. Из результатов работы ИНС выбирается максимальное значение $Argmax(y_i)$, которое определяет класс объекта.

ИНС осуществляют преобразование входных данных, представляющих образ, которое можно представить следующим образом

$$f_i(\varphi(a)) = y_i ,$$

где $\varphi(a)$ - нелинейное преобразование входных данных ИНС определенной архитектуры, i – номер ИНС в комбинированной структуре.

Множество ИНС, задействованных в комбинированной структуре, можно представить в виде

$$F \ni \{f_1(\varphi(a)), f_2(\varphi(a)), \dots, f_K(\varphi(a))\} ,$$

где: K – число используемых ИНС.

В зависимости от решаемых задач, используемые в комбинированной структуре ИНС, могут иметь различные архитектуры. Таким образом, в общем случае комбинированная структура имеет вид:

$$F \ni \{f_i(\varphi_j(a))\} , \quad i \in K, \quad j \in J ,$$

где K – ИНС, J – число различных архитектур.

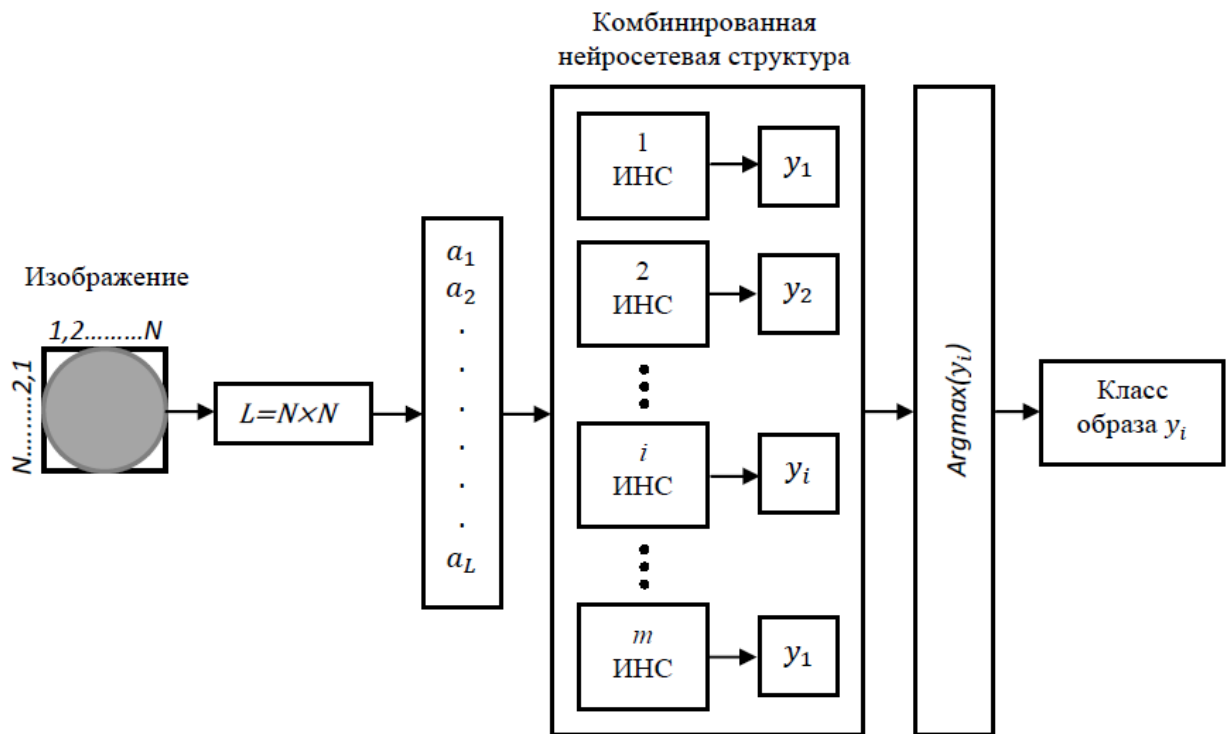


Рисунок 2 – Модель ИНС для классификации образов

Алгоритм обучения ИНС

В качестве алгоритма обучения выбрана процедура обратного распространения ошибки [9, 10].

Ошибка для обучающего примера P определяется по формуле

$$E_P = \frac{1}{2} \sum_{j=1}^L (y_j^N - d_j)^2, \quad (1)$$

где y_j^N – значение j -го выхода нейронной сети, N – номер выходного слоя, d_j – требуемое значение j -го выхода, L – число нейронов в выходном слое.

Полная ошибка

$$E = \sum_{P=1}^Z E_P,$$

где Z – число примеров в обучающей выборке, P – номер примера.

Скорректированное значение веса определяется по формуле

$$\Delta w_{ij} = -\eta \cdot \frac{\partial E}{\partial w_{ij}}, \quad (2)$$

где w_{ij} – значения веса, между i -ым и j -ым нейронами двух слоев; η – коэффициент скорости обучения равный $0 < \eta < 1$.

В начале процесса обучения значение параметра задается в диапазоне $0,5 < \mu < 1$. Далее для достижения минимума функции ошибки в процессе обучения значение параметра η уменьшается по следующей формуле

$$\eta_{t+1} = \alpha \cdot \eta_t,$$

где α – функция уменьшения μ – коэффициента скорости обучения, t – номер итерации.

Значение весового коэффициента на следующей итерации ($t+1$) определяется как сумма текущего и вычисленного корректирующего значений по формуле

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}(t),$$

Выходные значения для первого слоя нейронов определяются по формуле

$$S_1 = \sum_{i=1}^L w_{ij}^1 \cdot a_i,$$

где w_{ij}^1 – значения весов между входным вектором и первым слоем нейронов, a_i – значения входного вектора, L – число элементов входного вектора.

Выходные значения для последующих слоев нейронной сети определяются по формуле

$$S_j = \sum_{i=1}^m w_{ij}^n \cdot y_i^{n-1},$$

где y_i^{n-1} – выход нейрона на предыдущем слое, m – число нейронов в предыдущем слое.

Вычисление производной ошибки зависит от того, в каком слое находится нейрон, для которого корректируется рассматриваемый вес.

Производная ошибки для первого слоя

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial S_j} \cdot \frac{\partial S_j}{\partial w_{ij}} = \frac{\partial E}{\partial S_j} \cdot a_i. \quad (3)$$

Производная ошибки для последующих слоев

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial S_j} \cdot \frac{\partial S_j}{\partial w_{ij}} = \frac{\partial E}{\partial S_j} \cdot y_i^{n-1},$$

где n – номер слоя.

Сумма S_j влияет на общую ошибку только по отдельному нейрону y_i^n . Производная, обозначающая поправку входной суммы нейрона

$$\frac{\partial E}{\partial S_{ij}} = \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial S_j}.$$

Представим поправку через переменную:

$$\delta_j = \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial s_j} . \quad (4)$$

Учитывая, что ошибка для отдельного обучающего примера p определяется из формулы (1), получаем ошибку по нейрону y_j . Для выходного слоя сети поправку имеет вид

$$\frac{\partial E}{\partial y_j} = (y_j^N - d_j) . \quad (5)$$

Производная активационной функции нейрона

$$\frac{\partial y_j}{\partial s_j} = f'(s_j) . \quad (6)$$

Функция активации должна быть дифференцируемой и нелинейной. В качестве функции активации используется логистическая функция

$$f(s_j) = y_j = \frac{1}{1+e^{-\alpha s}} . \quad (7)$$

Производная логистической функции или по-другому сигмоида определяется по формуле

$$f'(s_j) = f(s_j) \cdot (1 - f(s_j)) = y_j \cdot (1 - y_j) = \frac{e^{-\alpha s}}{(1+e^{-\alpha s})^2} = \frac{1}{1+e^{-\alpha s}} \cdot \left(1 - \frac{1}{1+e^{-\alpha s}}\right) . \quad (8)$$

Таким образом, из формул (4-8) определяется поправка выходного слоя сети

$$\delta_j^N = \frac{\partial E}{\partial y_j} \cdot \frac{\partial y_j}{\partial s_j} = (y_j^N - d_j) \cdot f'(s_j) = (y_j^N - d_j) \cdot y_j^N \cdot (1 - y_j^N), \quad (9)$$

где N – номер выходного слоя.

Производная ошибки из (3) определяется по формуле

$$\frac{\partial E}{\partial w_{ij}} = (y_j^N - d_j) \cdot y_j^N \cdot (1 - y_j^N) \cdot y_i^{N-1} ,$$

Скорректированные значения весов определяются по формуле

$$\Delta w_{ij}^n = \eta \cdot \delta_j^n \cdot y_i^{(n-1)} , \quad (10)$$

где n – номер слоя.

Указанная в формуле (9) ошибка δ_j^N соответствует ошибке элемента выходного слоя сети. Однако ошибки элементов на предшествующих скрытых слоях не связаны непосредственно с целевыми выходными значениями. Весовые значения элементов в предшествующих скрытых слоях корректируются пропорционально их вкладу в величину ошибки последующего слоя, более близкого к выходу сети [10]. Вклад для элемента предшествующего выходному слою определяется от величины ошибки выходного

элемента и значения веса, соединяющего выходной и предшествующий элементы. Ошибка для элемента скрытого слоя определяется по формуле

$$\delta_j^n = f'(s_j) \cdot \sum_{k=1}^m \delta_k^{n+1} \cdot w_{kj}^{n+1},$$

где k – индекс элементов слоя посылающего ошибку, m – число нейронов в $(n+1)$ -ом слое.

Таким образом, вычисляются поправки для выходного слоя, через которые выражаются поправки для следующего скрытого слоя. Аналогичным образом вычисляются поправки и для остальных скрытых слоев более низких уровней.

Алгоритм имеет меньшую эффективность, если изменения значений различных весов ощутимо отличаются, то есть, отличаются производные по весам. Для уменьшения влияния этой проблемы вводится параметр α , сглаживающий слишком большое изменение корректирующих значений весов [10]. Кроме того параметр α позволяет преодолевать небольшие локальные минимумы [9]. Таким образом, из формулы (10) корректирующее значение веса находится по формуле:

$$\Delta w_{ij}^n(t+1) = \eta \cdot \delta_j^n \cdot y_i^{(n-1)} + \alpha \cdot \Delta w_{ij}^n(t),$$

где n – номер слоя.

Схема использованного алгоритма обучения с обратным распространением ошибок показана на рисунке 3.

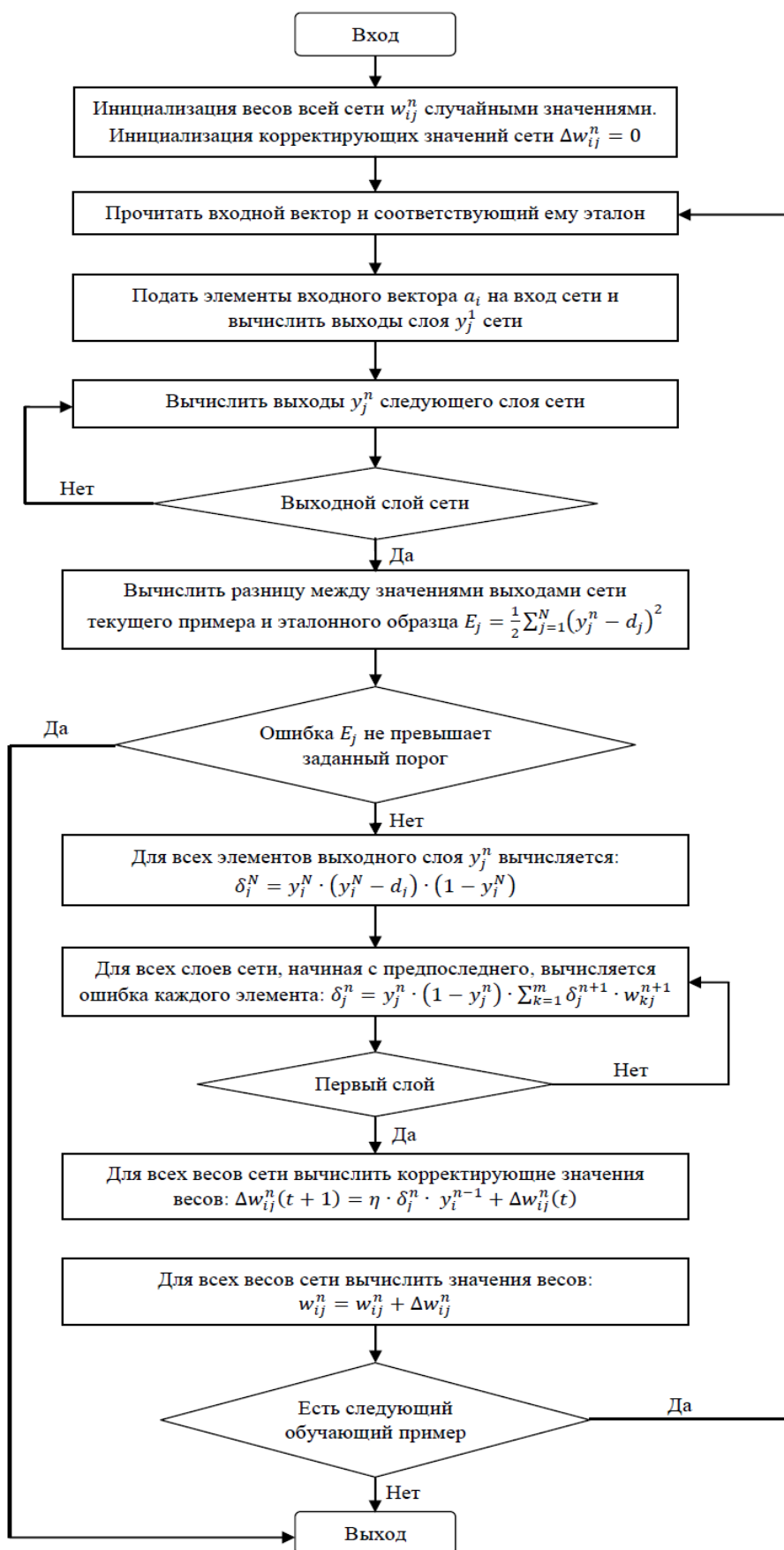


Рисунок 3 – Алгоритм обратного распространения ошибок

Результаты обучения и тестирования показаны на рисунках 4, 5.

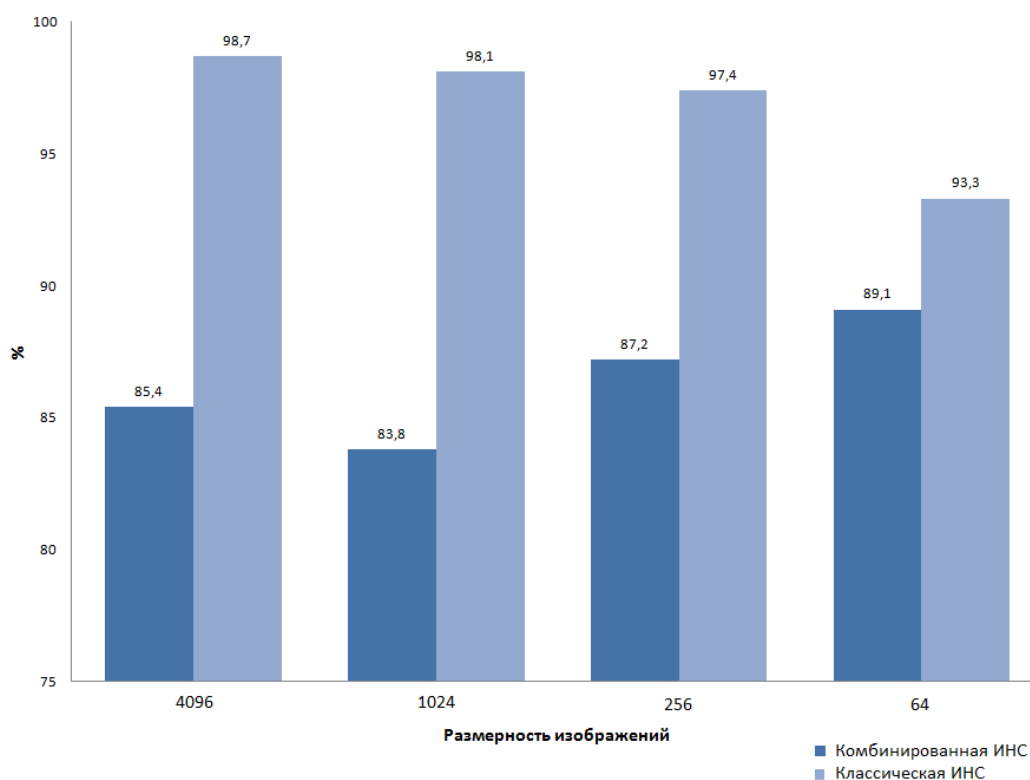


Рисунок 4 – Результаты обучения ИНС

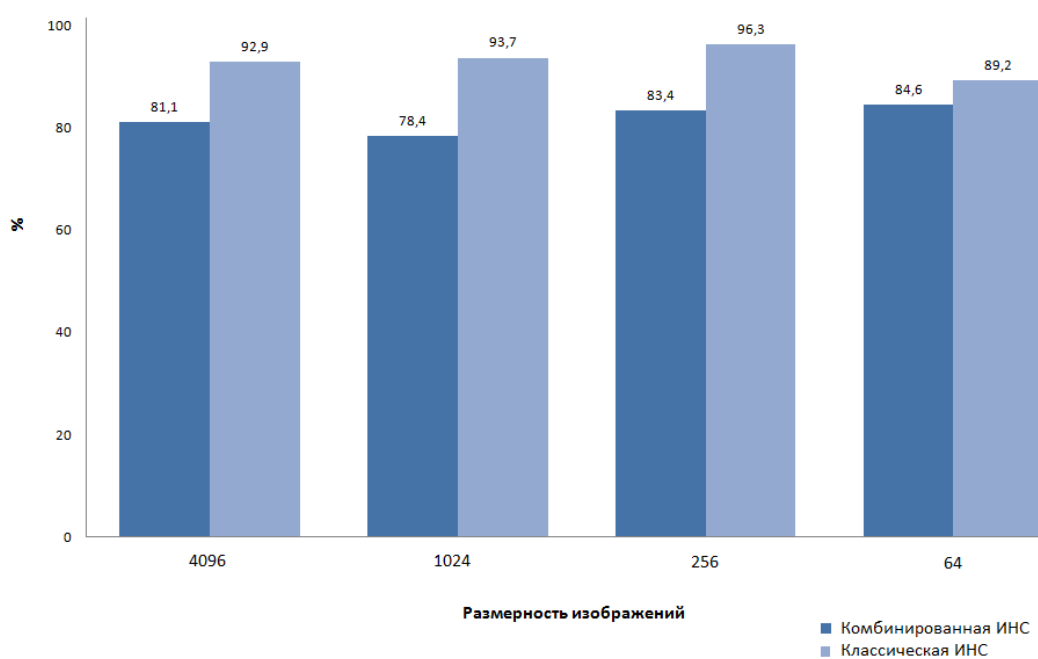


Рисунок 5 – Результаты тестирования ИНС

Более высокие результаты, как при обучении, так и при тестировании были получены с помощью модели на основе комбинированной структуры.

Литература

1. Томашевич Н.С., Томашевич Д.С., Галушкин А.И. Методы реализации инвариантности к аффинным преобразованиям при распознавании двумерных изображений// Информационные технологии. Приложение. 2001, № 1.
2. Srinivasa N., Jouaneh M. A neural network model for invariant pattern recognition// IEEE Transactions on signal processing. June 1992, Vol. 40, No. 6, pp. 1595-1598.
3. Widrow B., Winter R.G., Baxter R. A. Layered neural nets for pattern recognition// IEEE Transactions on acoustics, speech and signal processing. July 1988, Vol. 36, No. 7, pp. 1109-1118.
4. Томашевич Н.С. Нейросетевые алгоритмы выделения признаков и распознавания объектов на 2D изображениях инвариантные к аффинным преобразованиям// Труды 8-ой Всероссийской конференции «Нейрокомпьютеры и их применение». Москва. Россия. 2002, с. 185-187.
5. Fukumi M., Omatu S., Takeda F., Kosaka T. Rotationinvariant neural pattern recognition system with application to coin recognition// IEEE Transactions on neural networks. March 1992, Vol. 3, No. 2, pp. 272-278.
6. Fukumi M., Omatu S., Takeda F., Kosaka T. Rotationinvariant neural pattern recognition system estimating a rotation angle// IEEE Transactions on neural networks. May 1997, Vol. 8, No. 3, pp. 568-581.
7. Сосулин Ю.Г., Фам Чунг Зунг. Нейросетевое распознавание двумерных изображений// Радиотехника и электроника. 2003, т. 48, № 8, с. 969-978.
8. Сосулин Ю.Г., Фам Чунг Зунг. Инвариантное распознавание изображений комбинированной нейронной сетью// Нейрокомпьютеры: разработка и применение. 2003, № 8-9, с. 29-37.
9. Чубукова И.А. Data Mining. Учебное пособие для ВУЗов. – М.: Бином-Пресс, Бином. Лаборатория знаний, Интернет-Университет Информационных Технологий, 2008. – 382с.
10. Каллан Роберт Основные концепции нейронных сетей. Пер.с англ. – М.: Вильямс, 2001. – 287с.