

НЕЙРОННЫЕ СЕТИ В СИСТЕМАХ РАСПОЗНАВАНИЯ ТЕКСТА

Рассматриваются возможности использования свёрточных нейронных сетей для распознавания рукописных цифр.

ВВЕДЕНИЕ

Для решения задачи используется свёрточные нейронные сети с тремя скрытыми слоями. Обучение сети предлагается проводить с использованием базы рукописных цифр MNIST.

I. СТРУКТУРА СЕТИ

Для экспериментальной оценки качества работы свёрточной нейронной сети используется автоматизированная система распознавания рукописных цифр, в которой реализована модель свёрточной сети с тремя скрытыми слоями, без слоев подвыборки, но со смещениями рецептивных полей свёрточных нейронов не на один, а на два пикселя [1, 2]. На рис. 1 приведена структура реализованной сети.

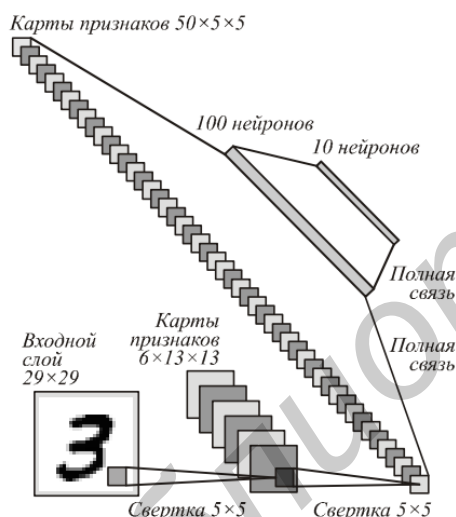


Рис. 1 – Структура свёрточной нейронной сети с тремя скрытыми слоями

Такой размер также обеспечивает достаточное наложение рецептивных полей друг на друга. В результате каждый слой свёртки уменьшает размер карты признаков с размера n до размера $(n - 3)/2$.

Входными данными нейронной сети являются изображения сегментированных рукопис-

ных цифр базы MNIST [3]. Размер обучающей выборки составляет 60000 символов, размер тестирующей выборки – 10000 символов.

II. ФУНКЦИЯ АКТИВАЦИИ. ОБУЧЕНИЕ

Функцией для скрытых слоёв сети был выбран гиперболический тангенс:

$$f(a) = A \cdot \tanh(S \cdot a)$$

где $f(a)$ – искомое значение элемента; a – взвешенная сумма сигналов предыдущего слоя; A, S – параметры активирующей функции.

Формулы функций ошибки приведены ниже:

$$E(w) = \frac{1}{2} \sum_{j=1}^p \sum_{k=1}^M (y_k^j - d_k^j)^2;$$

$$E(w) = - \sum_{j=1}^p \sum_{k=1}^M d_k^j \ln \frac{y_k^j}{d_k^j},$$

где p – размер обучающей выборки; j – номер обучающего примера; M – количество выходных нейронов; k – номер выходного нейрона; y_k^j – реальное значение сигнала выходного нейрона; d_k^j – ожидаемое значение.

Для ускорения работы алгоритма значения входных пикселей нормализуются по формуле:

$$y_i = \frac{x_i}{128} - 1,$$

где x_i – значение i -го пикселя изображения из базы; y_i – значение, подаваемое на вход сети.

Список литературы

1. Bishop, C.M. Neural Networks for Pattern Recognition – Oxford University Press, 1995. – P. 498
2. Garshin, A.A., Soldatova, O.P. An automated system of recognizing handwritten digits based on convolution neural networks // Certificate of a formal registration of computer software № 2010610988, the application № 2009616812 on December 1, 2009. Registered in the Register of Computer Programs February 1, 2010
3. LeCun, Y. The MNIST database of handwritten digits – <http://yann.lecun.com/exdb/mnist>.

Каханович Александр Иванович, магистрант кафедры информационных технологий автоматизированных систем Белорусского государственного университета информатики и радиоэлектроники, kakhanovichal@gmail.com.

Научный руководитель: Севернёв Александр Михайлович, доцент кафедры информационных технологий автоматизированных систем Белорусского государственного университета информатики и радиоэлектроники, кандидат технических наук, доцент.