

А. Е. Кирюшина

## Распознавание математических символов с использованием сверточной нейронной сети

Научный руководитель: к.т.н. И. П. Тищенко

*Аннотация.* Статья посвящена алгоритму распознавания математических символов в печатных и электронных документах. В качестве инструмента распознавания используется сверточная нейронная сеть. Также в статье представлены результаты распознавания.

*Ключевые слова и фразы:* сверточная нейронная сеть, греческий алфавит, OpenCV.

### Введение

Одной из отличительных особенностей текстового документа от математического является наличие в последнем соответствующего набора символов: знаков математических операций, букв греческого алфавита, специальных математических символов (знаки интеграла, скобки, знаки операций). Также для математической формулы характерно использование символов разного размера (индексы и степени). Некоторые формулы могут иметь в составе символы, расположенные на нескольких строках или по длине превышающие среднюю длину символа или пробела. В предыдущей работе [1] были описаны результаты распознавания функциональных символов, знака интеграла и скобок с использованием сверточной нейронной сети (далее СНС) [2]. Тестирование СНС проводилось на 10 классах символов. Процент распознавания был достаточно высоким 97.9%, причем на функциональных символах процент распознавания составляет 100 [3]. В качестве следующего этапа выбрано распознавание алфавита греческих символов. Как описывалось в предыдущих статьях [1, 3], в качестве исходных данных поступают изображения с формулами, выделенные из pdf документа.

## 1. Изменение размера изображения

На входе имеется изображение с формулой, которое сегментируется согласно алгоритму, описанному в статье [4]. Полученный набор изображений выделенных символов передается на блок изменения изображения. В качестве новых параметров высоты и ширины выбраны 32 пикселя  $\times$  32 пикселя. Число 32 выбрано для того, чтобы в первом и во втором сверточных слоях нейронной сети размер карт признаков был целым числом [3].

В качестве метода изменения размера изображения была взята функция `cvResize` из библиотеки OpenCV [5]. На рисунке 1 представлен результат работы данной функции:

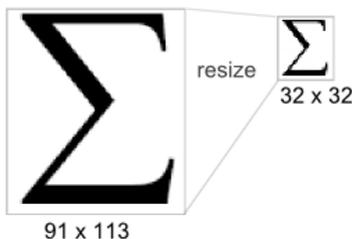


Рис. 1. Результат работы алгоритма `resize` из OpenCV

Далее изображения размером  $32 \times 32$  поступают на блок распознавания.

## 2. Архитектура СНС для распознавания греческих букв

СНС — нейронная сеть, имеющая многослойную архитектуру. Слои предназначены для возможности быстрого обучения и уменьшения общего значения функции ошибки [6].

На вход сети и каждого последующего слоя поступают изображения, представленные в виде двумерных массивов, или карты признаков. Модель СНС [6] представляет собой последовательность сверточных и вычислительных слоев. Сверточный слой состоит из нескольких карт признаков, на которых все нейроны имеют общие веса. Вычислительный слой предназначен для локального усреднения и подвыборки, что уменьшает разрешение карт признаков. Последний слой является классификационным, в котором количество нейронов равно

количеству распознаваемых классов. Каждый новый слой получает сигнал от локального рецептивного поля из предыдущего слоя. Таким образом, обеспечивается связь нейронов. В работе использовалась сеть без слоев подвыборки. Рассмотрим ее архитектуру.

Выбранная модель сети состоит из пяти слоев: входного, трех скрытых слоев свертки и классификационного выходного. Локальное рецептивное поле сверточных нейронов [7] имеет смещение на два пикселя и размер  $5 \times 5$ . Подробная архитектура сети представлена на рисунке 2:

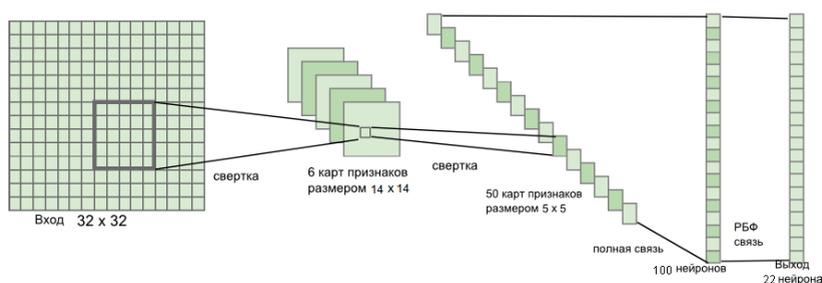


Рис. 2. Архитектура СНС

Входными данными нейронной сети являются сегментированные изображения греческих букв из математической формулы, размер каждого изображения  $32 \times 32$ . Таким образом, входной слой содержит 1024 нейрона. Размер обучающей выборки составляет 362 изображения, размер тестирующей выборки — 196 символов.

Второй слой, первый слой свертки, состоит из шести карт признаков размером  $14 \times 14$ . В данном слое находится  $14 \times 14 \times 6 = 1176$  нейронов. Каждый элемент признаков соединен с рецептивным полем  $5 \times 5$  из входного слоя. Из этого следует, что каждый элемент карты имеет 26 обучаемых весов. Таким образом, в первом сверточном слое находится  $1176 \times 26 = 30576$  синаптических связей и 156 обучаемых весов. Перейдем к описанию второго сверточного слоя.

Третий слой, второй слой свертки, состоит из 50 карт признаков размером  $5 \times 5$  и содержит  $5 \times 5 \times 50 = 1250$  нейронов. Каждый элемент карты имеет связь с шестью рецептивными полями размером  $5 \times 5$  из предыдущего слоя. Таким образом, во втором сверточном слое находится 7550 весовых коэффициентов и 188750 синаптических связей.

Данные сверточные слои используются для извлечения признаков из изображения. Каждый нейрон соединен со всеми нейронами из предыдущего слоя.

Четвертый слой, открытый, состоит из 100 нейронов. В этом слое содержится 125100 связей и столько же параметров для обучения.

Пятый слой, выходной, состоит из 22 нейронов, так как распознается 22 символа из алфавита. Также в последнем слое содержится 2222 связей и столько же весов.

Обучение построенной СНС производится с помощью метода обратного распространения ошибки [8]. Но в данном случае учитывается архитектура сети, а именно то, что веса одной карты признаков используются совместно.

### 3. Результаты распознавания

Выбранных значений параметров настройки сети было достаточно, чтобы получить довольно высокий процент распознавания. В таблице 1 представлены результаты распознавания тестовой выборки. Таблица 2 демонстрирует точность распознавания обучающей и тестовой выборки.

ТАБЛИЦА 1. Результаты распознавания греческого алфавита

Символ	% распознавания	Символ	% распознавания
$\alpha$	100%	$\pi$	100%
$\beta$	75%	$\rho$	100%
$\gamma$	100%	$\sigma$	100%
$\delta$	100%	$\varsigma$	100%
$\varepsilon$	100%	$\tau$	100%
$\zeta$	100%	$\upsilon$	100%
$\eta$	100%	$\phi$	100%
$\lambda$	100%	$\chi$	88.2353%
$\mu$	100%	$\psi$	100%
$\nu$	100%	$\omega$	100%
$\xi$	100%	$\varphi$	100%

ТАБЛИЦА 2. Результаты обучения сети

Точность распознавания обучающей выборки (362 символа), %	Точность распознавания тестовой выборки (196 символов), %	Количество эпох обучения
100	98.07	90

### Список литературы

- [1] Кирюшина А. Е. *Структурный анализ математических формул на основе метода выделения базовых признаков* // Тезисы докладов Всероссийской конференции с международным участием «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем». — Российский университет дружбы народов, Москва, апрель 2012, с. 127–129. ↑(document)
- [2] Солдатова О. П., Гаршин А. А. *Применение сверточной нейронной сети для распознавания рукописных цифр* // Компьютерная оптика. — Т. 34, 2010, с. 251–259. ↑(document)
- [3] Кирюшина А. Е. *Распознавание символов на основе метода базовых локальных признаков* // Сборник трудов Молодежной конференции «Научное развитие информационных технологий». — УГП имени А. К. Айламазяна, Переславль-Залесский, апрель 2012. ↑(document), 1
- [4] Кирюшина А. Е. *Структурный анализ математических формул и символов* // Сборник трудов Молодежной конференции «Научное развитие информационных технологий». — УГП имени А. К. Айламазяна, Переславль-Залесский, апрель 2011, с. 51–61. ↑1
- [5] Сайт библиотеки OpenCV, <http://opencv.willowgarage.com/wiki/>. ↑1
- [6] LeCun Y., Boser B., Denker J., Henderson D., Howard R., Hubbard W., Jackel L. *Backpropagation Applied to Handwritten Zip Code Recognition* // Neural Computation. — Т. 4, 1989, с. 541–551. ↑2
- [7] Определение рецептивного поля, [http://en.wikipedia.org/wiki/Receptive\\_field](http://en.wikipedia.org/wiki/Receptive_field). ↑2
- [8] Уоссермен Ф. *Нейрокомпьютерная техника: Теория и практика*: «Мир», 1992. ↑2

A. E. Kiryushyna. *Mathematic symbols recognition using convolution neural network* // Proceedings of Junior research and development conference of Ailamazyan Pereslavl university. — Pereslavl, 2013. — p. . (in Russian).

ABSTRACT. This paper describes the results of Greek letters recognition. To solve the problem the author used convolution neural network.

*Key Words and Phrases:* convolution neural network, Greek letters, OpenCV.